



**HAL**  
open science

# Nains sans géants : architecture décentralisée et services Internet

Francesca Musiani

► **To cite this version:**

Francesca Musiani. Nains sans géants : architecture décentralisée et services Internet. Sociologie. Ecole Nationale Supérieure des Mines de Paris, 2012. Français. NNT : 2012ENMP0043 . pastel-00795169

**HAL Id: pastel-00795169**

**<https://pastel.hal.science/pastel-00795169>**

Submitted on 27 Feb 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

École doctorale n° 396 : Economie, Organisations, Société

## Doctorat ParisTech

# THÈSE

pour obtenir le grade de docteur délivré par

**l'École nationale supérieure des mines de Paris**

**Spécialité " Socio-économie de l'innovation "**

*présentée et soutenue publiquement par*

**Francesca MUSIANI**

le 15 octobre 2012

**Nains sans géants**

***Architecture décentralisée et services Internet***

Directeur de thèse : **Cécile MÉADEL**

### Jury

**Mme Valérie BEAUDOUIN**, Directrice d'études, Télécom ParisTech

**M. Geoffrey C. BOWKER**, Professeur, University of California at Irvine

**M. Massimiano BUCCHI**, Professeur, Università degli Studi di Trento

**M. Alexandre MALLARD**, Maître de recherche, Mines ParisTech

**Mme Cécile MÉADEL**, Professeure, Mines ParisTech

**Mme Laurence MONNOYER-SMITH**, Professeure, Université de technologie de Compiègne

Rapporteur

Président

Examinateur

Examinateur

Directeur de thèse

Rapporteur

T  
H  
È  
S  
E



*Nani gigantum humeris insidentes.*

Bernard de Chartres (XIIème siècle)

*Sur les épaules d'un géant.*

Page d'accueil, *Google Scholar* (depuis 2004)



*Pour mes parents,*

*avec qui, il y a longtemps,*

*l'aventure parisienne a commencé.*



## REMERCIEMENTS

S'il s'agit le plus souvent d'une démarche solitaire, un travail de recherche de plusieurs années ne peut s'accomplir avec succès sans la contribution et l'apport d'un certain nombre de personnes. C'est tout particulièrement le cas de cette thèse, qui, depuis ses débuts, a pu bénéficier de l'insertion dans deux équipes de recherche, et de nombreuses autres collaborations qui m'ont apporté beaucoup tant sur le plan personnel que professionnel. Je souhaite donc exprimer ma reconnaissance aux nombreuses épaulés, de nains et de géants, sur lesquelles je suis montée au cours de ces quatre années de thèse.

Les premières, les plus patientes et présentes, sont celle de ma directrice de thèse Cécile Méadel. Ces quelques mots sont sans doute inadéquats à exprimer toute ma reconnaissance à ma « mère académique » pour les discussions éclairantes dans le bureau de la mezzanine ; pour les nombreux retours, toujours arrivés au bon moment et avec le sourire (le sien – et par conséquent le mien aussi !) ; pour la patience de corriger, couper et remettre en forme mes longues phrases d'Italienne ; et surtout, pour m'avoir écouté chaque fois que je le demandais – et m'avoir laissée libre de me débrouiller quand il le fallait. Un grand merci, enfin, pour ces quatre années de travail et de conversations ensemble, qui dépassent le cadre de la thèse. Je me réjouis par avance de pouvoir les continuer dans un futur proche.

Depuis son arrivée au CSI au début de ma deuxième année de thèse, Alexandre Mallard a été mon co-encadrant à tous les égards, sauf que formellement sur la première page de cette thèse. Un grand merci pour le soutien qu'il a toujours apporté à ce travail, pour les relectures attentives, les suggestions pointues et les discussions intéressantes – et pour les conversations sympathiques dans les couloirs du CSI. Et, bien sûr, pour être dans mon jury.

À n'en pas douter, les épaulés de Valérie Beaudouin, Geoffrey Bowker, Massimiano Bucchi et Laurence Monnoyer-Smith étaient déjà suffisamment chargés sans que je ne rajoute au fardeau 350 pages de thèse à lire et à évaluer. C'est d'autant plus un plaisir et un honneur pour moi qu'ils aient accepté de faire partie de mon jury. Je les en remercie sincèrement et je me réjouis par avance de nos discussions.

Le Centre de Sociologie de l'Innovation a été un environnement de travail idéal au cours de ces quatre années, et les échanges avec ses membres ont été un ingrédient fondamental de cette thèse. Merci à tous, et en particulier à Antoine Hennion, Fabian Muniesa et Dominique Linhardt, coordonnateurs de l'atelier doctoral du CSI, ainsi qu'aux collègues et amis doctorants qui, au fil des séances, ont commenté mes premiers jets de chapitre et mes articles. Merci à Madeleine Akrich pour m'avoir témoigné sa confiance en ma capacité de mener à bien ce travail, il y a quatre ans et à plusieurs reprises par la suite, et à Catherine Lucas pour le sourire intarissable et la grande efficacité qui ont toujours accueilli mes demandes administratives.

Au cours de cette dernière année, plusieurs collègues, amis et proches ont bien voulu répondre présent à mes sollicitations de relecture de chapitres. Un grand merci, pour leurs retours et pour avoir été une partie importante de ma vie de thésarde, à Romain

Badouard, Clément Combes, Mélanie Dulong de Rosnay, Jean-Marc Galan, Brice Laurent, Clément Mabi, Morgan Meyer, Valérie Schafer, Katharina Schlierf et Guillaume Sire. Je suis tout particulièrement reconnaissante à Joachim Bloche pour avoir eu le courage, pendant ma dernière ligne droite, de relire l'ensemble de mon manuscrit. Mention spéciale pour Enoch Peserico, à qui le début et la fin de cette thèse – la première idée de sujet, et le titre – doivent beaucoup, sans mentionner les nombreuses contributions et discussions pendant son élaboration.

Depuis ses débuts, ce travail a en partie été conduit dans le cadre de deux projets financés par l'Agence Nationale de la Recherche : Vox Internet (2008-2010) et ADAM (Architectures distribuées et applications multimédias, 2010-en cours). Il ne serait sans doute pas le même sans avoir pu bénéficier des interactions avec les membres de ces deux équipes. En particulier, je souhaite exprimer ma reconnaissance à Françoise Massit-Folléa pour toutes les opportunités que Vox a représenté pour moi, ainsi que pour son soutien et pour son amitié, qui ont largement dépassé le cadre du projet. Merci à Danièle Bourcier, Primavera De Filippi, Isabelle Demeure, Annie Gentès, François Huguet et Maya Bacache – ainsi qu'à Cécile, Alexandre et Mélanie – pour les échanges dans le cadre d'ADAM, qui ont grandement contribué à finaliser ce travail. Je me réjouis de poursuivre cette collaboration dans les mois à venir.

Le réseau qui a contribué à rendre cette thèse possible ne se limite pas à ces deux projets. Une pensée amicale supplémentaire à Valérie Schafer – les travaux menés en collaboration avec elle et Hervé Le Crosnier, depuis 2010, m'ont beaucoup apporté sur le plan personnel et ont donné une contribution importante à la problématique de cette thèse. La communauté de l'Association Internationale des Études et Recherches sur l'Information et la Communication (AIERI/IAMCR) a commenté mon travail au fil des années et des continents : merci en particulier à Claudia Padovani, Stefania Milan, Jeremy Shtern, Jo Pierson, Arne Hintz et Elena Pavan. L'Institut des Sciences de la communication du CNRS, ainsi que les conférences EASST, STSItalia, GigaNet et STGlobal, ont été des arènes de discussion utiles et stimulantes.

Cette thèse n'existerait bien sûr pas sans la disponibilité et l'enthousiasme des équipes de chercheurs, développeurs et entrepreneurs qui travaillent chaque jour avec les « nains » de l'Internet. Je leur suis très reconnaissante pour le temps qu'ils m'ont consacré – en particulier, je remercie l'équipe Tribler de m'avoir accueilli à Delft – et j'espère qu'ils trouveront ce travail utile. Merci également à tous les spécialistes du P2P qui ont accepté de répondre à mes questions dans la phase exploratoire de ce travail, en particulier à Olivier Auber, Gwendal Simon, Serge Abiteboul et, *last but not least*, Fabrice Le Fessant.

Enfin, une pensée reconnaissante à mes futurs collègues de l'Institut d'études diplomatiques de l'université de Georgetown, à Washington, DC. La confiance qu'ils ont accordée à ma capacité de terminer mon travail de thèse dans les délais prévus, et leur appréciation pour mon programme de recherche présent et futur, ont donné une motivation et une contribution importantes à la finalisation de ce travail. *I believe we can look forward to a great year.*

## TABLE DES MATIÈRES

<b>REMERCIEMENTS</b>	<b>7</b>
<b>CHAPITRE 1</b>	<b>13</b>
<b>INTRODUCTION GÉNÉRALE</b>	<b>13</b>
<b>EN GUISE DE PRÉLUDE...</b>	<b>14</b>
<b>SECTION 1. L'APPROCHE DÉCENTRALISÉE DES SERVICES INTERNET</b>	<b>18</b>
INNOVATION ET INTERNET, UNE QUESTION D'ARCHITECTURE ?	20
LES MARGES DU RÉSEAU, UNE RÉPONSE À L'AVENIR « NUAGEUX » DE L'INTERNET ?	23
LE P2P COMME OPPORTUNITÉ ET ALTERNATIVE	25
<b>SECTION 2. TROIS CAS POUR UN REGARD COMMUN</b>	<b>29</b>
UNE SOCIOLOGIE DE L'INNOVATION « EN TEMPS RÉEL »	30
INNOVATION ET INNOVATEURS DANS LE P2P	34
PLAN DE LA THÈSE	38
<b>CHAPITRE 2</b>	<b>41</b>
<b>INNOVATION ET INTERNET DISTRIBUÉ : UN REGARD INTERDISCIPLINAIRE</b>	<b>41</b>
<b>SECTION 1. CONTEXTE ET PRATIQUE(S) DE L'INNOVATION SUR INTERNET</b>	<b>43</b>
LES <i>INTERNET STUDIES</i> VERS UNE INTÉGRATION DES STS	43
(DÉ)STABILISATIONS : POUR UNE ÉTUDE DE L'INTERNET COMME INNOVATION « EN RÉSEAU »	46
DES USAGERS REPRÉSENTÉS AUX USAGERS PIONNIERS ET CO-CONCEPTEURS : L' « OUVERTURE » DE L'INNOVATION EN RÉSEAU	50
<b>SECTION 2. LE PAIR-À-PAIR, UN MODÈLE DE RÉSEAU INFORMATIQUE AUX USAGES MULTIPLES</b>	<b>55</b>
PARTAGE, DÉCENTRALISATION, AUTONOMIE	56
QUESTIONS DE « COUCHES », QUESTIONS DE TERMINOLOGIE	57
DROIT(S) EN RÉSEAU : SÉCURITÉ, <i>PRIVACY</i> ET SURVEILLANCE	70
<b>SECTION 3. DE L' « INTÉRÊT DES TUYAUX » : ÉTUDIER LES ARCHITECTURES</b>	<b>79</b>
LES ARCHITECTURES DE L'INTERNET ET LA DÉCOUVERTE DU « TRAVAIL INVISIBLE »	80
QUELLES ARCHITECTURES POUR LE FUTUR INTERNET ?	83
<b>EN GUISE DE SYNTHÈSE</b>	<b>88</b>
<b>CHAPITRE 3</b>	<b>89</b>
<b>P2P « DÉCENTRALISÉ » : S'ÉLOIGNER DU SERVEUR POUR LE REPENSER</b>	<b>89</b>
<b>INTRODUCTION</b>	<b>89</b>
<b>SECTION 1. INTERNET DÉCENTRALISÉ, ENTRE COOPÉRATIONS ET TENSIONS</b>	<b>92</b>
USENET, LE MODÈLE DÉCENTRALISÉ ORIGINEL	93
<i>DOMAIN NAME SYSTEM</i> , L'HYBRIDE SOUPLE	95
<b>SECTION 2. LA DIFFUSION MONDIALE D'INTERNET ET SES CONSÉQUENCES SUR L'ARCHITECTURE DU RÉSEAU</b>	<b>98</b>
LE PASSAGE AU MODÈLE CLIENT/SERVEUR	98
LA RUPTURE DU PARADIGME DE COOPÉRATION	99
LA FIN DU RÉSEAU OUVERT	103
LA DISPONIBILITÉ DE BANDE PASSANTE : DES CONNEXIONS ASYMÉTRIQUES	105
<b>SECTION 3. ENTRE ANCIENS MODÈLES ET INNOVATION : LA « RENAISSANCE » DU P2P</b>	<b>106</b>

CRÉATION ET PUBLICATION DE L'INFORMATION	106
LES DÉFIS POUR LA « SÉCURISATION » DES RÉSEAUX	107
QUELLE DÉCENTRALISATION ?	108
<b>SECTION 4. AUX MARGES DU RÉSEAU</b>	<b>110</b>
ENTRE DÉCENTRALISATION ET PASSAGE À L'ÉCHELLE, DANS UN ENVIRONNEMENT FRAGILE	110
PARTAGER DES RESSOURCES À LA PÉRIPHÉRIE DE L'INTERNET	111
OPTIMISER SANS STABILISER	113
<b>EN GUISE DE CONCLUSION : LE BAR SICILIEN ET LES SYSTÈMES DISTRIBUÉS</b>	<b>115</b>
<b>CHAPITRE 4</b>	<b>117</b>
<b>FAROO, LE MOTEUR DE RECHERCHE « RÉPARTI SUR LES UTILISATEURS »</b>	<b>117</b>
<b>INTRODUCTION</b>	<b>117</b>
<b>SECTION 1. PASSER À L'ÉCHELLE : LE CHOIX DE DÉCENTRALISER LA RECHERCHE D'INFORMATION EN LIGNE</b>	<b>122</b>
UN POSITIONNEMENT ÉCONOMIQUE ET « ACADÉMIQUE » DIFFICILE	126
CE QUE « FAIT » L'ARCHITECTURE D'UN MOTEUR P2P	129
<b>SECTION 2. EXPLORER LES NIVEAUX DE DISTRIBUTION</b>	<b>132</b>
LA FORCE DE LA COLLABORATION, OU LE <i>CRAWLER</i> DISTRIBUÉ	133
LA « SAGESSE DES FOULES », OU LA DÉCOUVERTE DISTRIBUÉE D'INFORMATION	135
MOINS DE COÛTS D'INFRASTRUCTURE, OU L'INDEX DE RECHERCHE DISTRIBUÉ	138
LE « POUVOIR DES DÉCLICS », OU LE CLASSEMENT DISTRIBUÉ PAR LE WEB IMPLICITE	140
PAS DE « SUPER-PAIRS », OU UN AMORÇAGE DISTRIBUÉ	142
« COMME DES CELLULES », OU LA MISE À JOUR DISTRIBUÉE	143
<b>SECTION 3. LES « SIX NIVEAUX DE DISTRIBUTION » ET LA CENTRALITÉ DE L'USAGER</b>	<b>146</b>
QUI VIENT EN PREMIER, L'ŒUF (L'INDEX) OU LA POULE (LES USAGERS) ?	148
LE TÉLÉCHARGEMENT DU CLIENT P2P, UNE BARRIÈRE À L'ENTRÉE ?	150
« LÉGALITÉ » DU DISPOSITIF, ENTRE RESPONSABILITÉ INDIVIDUELLE ET PARTAGÉ	154
<b>SECTION 4. LIER PERSONNALISATION ET VIE PRIVÉE : LE P2P COMME « RECHERCHE CENTRÉE SUR L'USAGER » ?</b>	<b>158</b>
DU « DOCUMENT-CENTRE » À L' « USAGER-CENTRE »	159
L'ÉCONOMIE DE L'ATTENTION ET LA RECHERCHE « SOCIALE »	161
DE LA « <i>PRIVACY BY POLICY</i> » À LA « <i>PRIVACY BY ARCHITECTURE</i> »	165
<b>EN GUISE DE SYNTHÈSE</b>	<b>171</b>
<b>CHAPITRE 5</b>	<b>175</b>
<b>WUALA, OU L'« ART DU COMPROMIS » ENTRE DISTRIBUÉ ET CENTRALISÉ</b>	<b>175</b>
<b>INTRODUCTION</b>	<b>175</b>
<b>SECTION 1. UN STOCKAGE EN RÉSEAU P2P</b>	<b>178</b>
REDONDANCE : FAÇONNER UN STOCKAGE P2P PERMANENT	182
« FRAGMENTS ENCRYPTÉS » ENTRE LOCAL ET GLOBAL, UNE PROMESSE DE SÉCURITÉ ?	184
CONFIANCE ET FERMETURE DU CODE SOURCE : LA SÉCURITÉ « PAR OBSCURITÉ » EN DÉBAT	187
<b>SECTION 2. STOCKAGE « SOCIAL » ET PRIVACY</b>	<b>194</b>
MOT DE PASSE ET RESPONSABILITÉ DE L'USAGER	196
LIENS DIRECTS ET « CLÉ D'AMITIÉ »	198
CONFIDENTIALITÉ DES DONNÉES ET ALLOCATION DE RESSOURCES	200
<b>SECTION 3. ESPACE DISQUE CONTRE ESPACE EN LIGNE : LA RECHERCHE D'UN MODÈLE SOCIO-ÉCONOMIQUE DURABLE POUR LE STOCKAGE DISTRIBUÉ</b>	<b>203</b>
AU COMMENCEMENT ÉTAIT L'APPLICATION BUREAU... QUI RENCONTRA LE WEB	204

TROC ET « QUALITÉ DE SERVICE »	206
DERRIÈRE LE NŒUD, L'USAGER : MOTIVATION ET RESPECT DES RESSOURCES	210
« <i>OUT THERE</i> » OU « <i>IN THERE</i> » ? LE NUAGE D'USAGERS ET LA CULTURE DU PARTAGE	215
ENTRE FRAGMENTS HÉBERGÉS ET SURVEILLANCE COLLABORATIVE : LA « LÉGALITÉ » DU <i>CLOUD P2P</i>	219
<b>SECTION 4. QUELLE DÉCENTRALISATION ? LE RÔLE CHANGEANT DES SERVEURS ET DES PAIRS</b>	<b>225</b>
DES NŒUDS-USAGERS QUI DEVIENDRONT AUTO-SUFFISANTS...	227
...OU DES USAGERS-RESSOURCES, « OPTIMISATION » DES SERVEURS ?	229
<b>EN GUISE DE SYNTHÈSE</b>	<b>234</b>
<b>CHAPITRE 6</b>	<b>239</b>
<b><u>DE TRIBLER À P2P-NEXT, OU LA CONSTRUCTION DU PEER-TO-PEER COMME VALEUR POUR L'EUROPE</u></b>	<b>239</b>
<b>INTRODUCTION</b>	<b>239</b>
<b>SECTION 1. UN PROJET UNIVERSITAIRE POUR UN PARADIGME DE P2P « SOCIAL-BASED »</b>	<b>245</b>
DE ABC À TRIBLER : LE « SOCIAL IMMÉDIAT »	245
PAIRS OU « BUDDIES » : DES SOLUTIONS SOCIALES POUR DES DÉFIS TECHNIQUES	248
<i>BUDDYCAST</i> , OU COMMENT LE « SOCIAL » SE TRADUIT EN ALGORITHME	252
SOCIAL ET <i>FAIR-PLAY</i> : LA BANDE PASSANTE COMME DEVISE	257
<b>SECTION 2. « <i>OUR PRODUCT OUT THERE</i> » : DU SOCIAL À LA VIDÉO À LA DEMANDE</b>	<b>260</b>
UNE INTERFACE ÉPURÉE, UNE ARCHITECTURE QUI RESTE « SOCIALE »	261
LA VIDÉO À LA DEMANDE DÉCENTRALISÉE ET L'OUVERTURE DE PARTENARIATS R&D	265
<b>SECTION 3. <i>P2P-NEXT</i> : VERS UN RÉSEAU EUROPÉEN DE TÉLÉVISION DISTRIBUÉE</b>	<b>269</b>
UNE COALITION COMPLEXE POUR UN OBJET CONTROVERSÉ	272
RESTER TRIBLER DANS <i>P2P-NEXT</i> : GÉRER LA « MISE-EN-GLOBALITÉ » D'UN PROJET LOCAL	276
CONSTRUIRE UNE « LÉGALITÉ » DU P2P ENTRE LICENCES LIBRES ET DROIT EUROPÉEN	278
<b>SECTION 4. DE PROBLÈME À VALEUR : LE <i>PEER-TO-PEER</i> DANS L'EUROPE</b>	<b>282</b>
LE P2P, UNE ARCHITECTURE POLITIQUE ET ÉCONOMIQUE POUR LA TÉLÉVISION EUROPÉENNE	283
VERS UNE CONTRIBUTION À L'ENCADREMENT JURIDIQUE DE LA TÉLÉVISION VIA P2P	286
UNE « VALEUR », L'ARCHITECTURE DÉCENTRALISÉE ?	289
<b>EN GUISE DE SYNTHÈSE</b>	<b>293</b>
<b>CHAPITRE 7</b>	<b>299</b>
<b><u>CONCLUSION GÉNÉRALE</u></b>	<b>299</b>
<b>SECTION 1. <i>NAINS, GÉANTS</i> ET L'« ALTERNATIVE » POUR LES SERVICES INTERNET</b>	<b>302</b>
L'ALTERNATIVE P2P : UNE QUESTION D'USAGES, DE TECHNIQUE ET DE « <i>BUZZ</i> »	303
REDÉFINIR LA DÉCENTRALISATION	304
« ÉLIMINATION DES INTERMÉDIAIRES », NOUVELLES INTERMÉDIATIONS	306
<b>SECTION 2. VERS UN P2P SOCIAL : REFORMATER LES RÉSEAUX, REPENSER L'USAGER</b>	<b>309</b>
L'INTRODUCTION DU « PARADIGME SOCIAL » DANS LE P2P	310
<i>USER FIRST</i> ? DE LA FACILITATION DES ÉCHANGES À LA CAPTURE DES AFFINITÉS	312
DU SOCIAL ET DES RESSOURCES : VERS UNE CENTRALITÉ DE L'USAGER	312
<b>SECTION 3. LE P2P : OBJET DU DROIT, PRODUCTEUR DE DROITS</b>	<b>314</b>
LA « LÉGALITÉ » DU P2P : AU DELÀ D'UN USAGE, POUR DES NOUVEAUX USAGES	315
LA PROTECTION « PAR LA TECHNIQUE » DE LA VIE PRIVÉE, ENTRE CONTRÔLE ET RESPONSABILITÉ	317
<b>OUVERTURE. DES SERVICES INTERNET EN P2P À L'« INTERNET ALTERNATIF » : UN ENJEU DE</b>	

<b>GOUVERNANCE</b>	<b>320</b>
INNOVATION ET P2P, UNE ÉTUDE DE GOUVERNANCE DE L'INTERNET ?	320
POUR UNE ÉTUDE DE L' « INTERNET ALTERNATIF »	321
<b><u>BIBLIOGRAPHIE</u></b>	<b><u>323</u></b>
<b><u>TABLEAUX</u></b>	<b><u>347</u></b>
<b><u>IMAGES</u></b>	<b><u>348</u></b>

# CHAPITRE 1

## INTRODUCTION GÉNÉRALE

*« Oh, tu t'occupes de peer-to-peer ! Tu ne pourrais pas me conseiller un bon endroit où aller chercher des films, maintenant qu'on a fermé Megaupload ? »*

*Depuis que, le 19 janvier 2012, le Federal Bureau of Investigation états-unien a disposé le blocage DNS<sup>1</sup> de dix-huit sites de streaming et de téléchargement direct de fichiers vidéo – parmi lesquels comptent le populaire Megaupload et son « cousin » Megavideo – je ne compte plus le nombre de fois que des amis, des proches et des simples connaissances m'ont posé la question ci-dessus.*

*Cette question me rappelle à chaque fois les nombreux articles de presse qui, au cours des deux ou trois dernières années, ont résumé la tendance de certaines pratiques de consommation numérique à se déplacer vers d'autres arènes que le P2P dans les termes suivants : « une baisse des échanges peer-to-peer face à la montée en puissance du streaming<sup>2</sup> ».*

*Ou encore, elle me rappelle les nombreuses remarques ironiques, dans le genre « fais attention à toi, tu travailles sur des choses dangereuses », qui m'ont été faites lorsque le très controversé projet de loi français préconisant une « réponse graduée » aux échanges illicites sur les réseaux s'est traduit dans une Autorité tout aussi controversée<sup>3</sup>, dans des dispositifs de surveillance et traçage des réseaux P2P d'une efficacité douteuse mais d'une valeur symbolique sûre, et dans un mécanisme techno-juridique de « désincitation » à l'échange sur les réseaux.*

---

<sup>1</sup> Le blocage par nom de domaine (ou blocage DNS pour *Domain Name System*) consiste à rendre inopérant le système de nommage utilisé pour localiser des ordinateurs et des services en ligne. Voir F. Musiani, « SOPA et Megaupload : quelle responsabilité pour les 'intermédiaires' de l'Internet ? », ADAM, 30 janvier 2012 [<http://adam.hypotheses.org/1188>]

<sup>2</sup> Voir par exemple O. Chicheportiche, « Le P2P ne fait plus recette, le streaming explose », ZDnet.fr, 13 octobre 2009 [<http://www.zdnet.fr/actualites/le-p2p-ne-fait-plus-recette-le-streaming-explose-39709235.htm>] et G. Belfiore, « Internet : P2P en baisse, streaming en hausse », Clubic, 19 février 2009 [<http://www.clubic.com/actualite-258912-traffic-internet-2008-2009.html>].

<sup>3</sup> La Haute Autorité pour la diffusion des œuvres et la protection des droits sur Internet (Hadopi), <http://hadopi.fr/>

## EN GUISE DE PRÉLUDE...

Qu'ont en commun mes amis lorsqu'ils recherchent leur prochain « Eden du cinéma », les auteurs des articles sur la corrélation entre *peer-to-peer* et streaming, et les collègues qui me font des remarques sur l'Hadopi ? Une conception du P2P comme « technologie de partage illégal », et très souvent, une confusion entre cette technologie de réseau et les usages qu'elle permet, héberge, préconise. Au cours de mes quatre années de thèse, j'ai appris à sourire aimablement avant de répondre que, en fait, je ne souhaite pas vraiment m'occuper de tout cela dans mon travail.

Au cours de ces quatre mêmes années, j'ai aussi appris à quel point une telle réponse de ma part est à son tour une simplification. Certes, le *peer-to-peer* (souvent traduit en français par « pair-à-pair » et plus souvent abrégé en P2P, abréviation que j'utiliserai par la suite) peut, en principe, être défini de manière assez simple d'un point de vue technique. Il s'agit d'un modèle de réseau informatique structuré de manière décentralisée, afin que les communications ou les échanges qui y ont lieu se fassent entre nœuds dotés d'une responsabilité égale dans le système. Les participants au réseau mettent à disposition une partie de leurs équipements et ressources informatiques (capacité de calcul, espace de stockage, bande passante) ; accessibles de manière directe par les pairs, ces ressources partagées sont nécessaires au bon fonctionnement du service offert par le réseau. La dichotomie entre un serveur, fournisseur de ressources, et les clients demandeurs de ressources, caractéristique du modèle *client-serveur*, est remplacée par une situation où tous les pairs hébergent ou fournissent la ressource, et tous les pairs la demandent.

Pourtant, pour un très grand nombre d'utilisateurs de l'Internet – depuis la rencontre entre le P2P et le grand public propulsée par le logiciel de partage de fichiers Napster, en 1999 – cette technologie est *de facto* synonyme de téléchargement (illégal) de contenus culturels. Dans cette optique, elle est en effet considérée comme une alternative au streaming vidéo, à son tour identifié avec des sites Web comme Megaupload ou, au mieux, YouTube – plutôt que d'être considérée comme une plateforme sur laquelle on peut « construire » du streaming vidéo et bien d'autres applications<sup>4</sup>. Enfin, pour une très grande majorité d'utilisateurs, de journalistes, de personnalités politiques, Hadopi est un projet de loi d'abord, une loi ensuite, et finalement un système juridique et technique, dont le but est de « contrer le P2P ». Les objets, les démarches de conception et développement, les collectifs et les usages dont il sera question dans cette thèse ne peuvent pas complètement faire l'économie des visions normatives fortes qui se confrontent et s'affrontent autour du P2P, qu'il

---

<sup>4</sup> Si le streaming vidéo correspond dans l'imaginaire d'une très grande majorité d'utilisateurs à des solutions centralisées proposées par des grandes plateformes, comme YouTube ou DailyMotion, ou même nombre de sites au statut légal plus douteux tel Megaupload, le streaming vidéo en P2P est déjà largement utilisé, et plusieurs projets de recherche étudient actuellement les moyens d'améliorer sa qualité de service (parmi eux, Tribler, le « protagoniste » du sixième chapitre de cette thèse). Ce système est plus particulièrement à l'œuvre dans le domaine du P2PTV. Les applications P2PTV sont destinées à redistribuer des flux (*streams*) vidéo en temps réel sur un réseau P2P ; chaque utilisateur, quand il télécharge un flux vidéo, permet simultanément aux autres utilisateurs de le télécharger, contribuant ainsi à la bande passante totale disponible.

s'agisse de l'identification de cette technologie avec le piratage de contenus culturels et intellectuels, ou à l'extrême opposé, de son déploiement dans une arène démocratique, égalitaire et collaborative accessible à tous : comme on aura l'occasion de voir tout au long de la thèse, mes objets d'étude sont informés par ces visions, discours, narratives, et les informent en retour.

Si elle ne veut certes pas négliger pas la puissance d'agir<sup>5</sup> de ces visions normatives, cette thèse ne souhaite cependant pas être une contribution ultérieure aux débats sur le droit d'auteur et sur la dialectique pillage/partage auxquels le P2P semble désormais être « naturellement » associé. Ce travail prend comme son point de départ la caractéristique de base du P2P en tant que modèle de réseau informatique : ce qui est, on l'a anticipé, le fait de permettre des échanges de données efficaces et directs entre des nœuds égalitaires. Égalitaires en termes de leur mise à disposition de ressources techniques à l'ensemble du système, et de la responsabilité qui leur est attribuée dans son fonctionnement.

Cette thèse s'intéresse donc au développement et à l'appropriation de services basés sur Internet<sup>6</sup> dont la conception intègre un choix de design particulier : le fait de déléguer la responsabilité et le contrôle de la gestion des données et des flux aux « marges » ou à la périphérie de ces systèmes en réseau. Les opérations nécessaires au bon fonctionnement des systèmes, et au fait qu'ils fournissent correctement les services auxquels ils sont destinés, dépendent techniquement de ceux que j'appelle les « nains » du réseau : les utilisateurs, leurs terminaux, leurs ressources informatiques, mobilisées de manière agrégée pour servir un objectif commun.

---

<sup>5</sup> C'est ainsi que Serge Proulx traduit une des notions fondatrices de la sociologie des sciences et des techniques, l'*agency*, en soulignant que cette traduction met en relief à la fois la dimension collective de l'agir et le « pouvoir des acteurs de formuler une critique effective et ainsi, de *faire dévier* le cours de la reproduction du monde social » (Proulx, 2009).

<sup>6</sup> Bien qu'étant sensible à la polysémie, à l'ambiguïté et à la complexité du terme et du concept de « service », sur lequel beaucoup a été écrit en sociologie des organisations, du travail et de la communication, ainsi qu'en socio-économie de l'innovation (Callon, Méadel & Rabeharisoa, 2002 ; Callon, Millo & Muniesa, 2007), je ne considère pas cette thèse comme l'endroit le plus approprié pour rentrer dans les débats autour de cette notion. J'adopte ici une définition basique de service Internet, celle de mise à disposition, au moyen de l'Internet, d'une capacité de communication, d'accès à l'information, de mise en relation ou de calcul. Voir aussi une définition juridique des « services en ligne », proposée par Alain Bensoussan, sur <http://www.alain-bensoussan.com/avocats/services-en-ligne-definition/2008/09/09?lang=fr>.

Un autre débat que je vais laisser de côté dans ce travail, tout en étant consciente de son importance pour certains acteurs (par exemple le projet de multilinguisme pour l'Internet *Net.lang*, <http://net-lang.net/>), porte sur la capitalisation du mot « Internet » et sur le rajout d'un article déterminant qui le précède. Je me limite à souligner ici que mon usage de la majuscule et de l'article pour me référer à l'Internet au cours de ce travail renvoient à certains attributs du « réseau des réseaux » qui me semblent importants dans le contexte de cette thèse : notamment, qu'on est en train de parler du système global de réseaux IP interconnectés, considéré dans sa totalité – et que celui-ci se différencie d'une variété d'autres réseaux internet : « [...] *internet is simply a contraction of the phrase interconnected network. However, when written with a capital 'I,' the Internet refers to the worldwide set of interconnected networks. Hence, the Internet is an internet, but the reverse does not apply. The Internet is sometimes called the connected Internet.* » (Parziale *et al.*, 2006 : 4).

Pour autant, la thèse ne touche que de manière secondaire au type de service qui est le plus souvent associé à l'architecture P2P, le partage de fichiers. Elle souhaite s'intéresser à la rencontre entre le choix de développer une architecture technique en P2P et des usages tels que la recherche d'information, le réseautage social, le visionnage de vidéos, le stockage de fichiers en ligne. Il s'agit d'usages qui nous sont très familiers dans notre pratique quotidienne d'internautes et d'utilisateurs de services en ligne, sous le nom de Google, Facebook, YouTube, Dropbox – les « géants » des technologies de l'information, basés sur une architecture de réseau client/serveur qui préconise une dichotomie clairement identifiable entre un serveur fournisseur de ressources, et des clients qui en sont demandeurs. Ce travail explore des dispositifs qui, tout en répondant à ces mêmes nécessités d'usage – recherche, réseautage, stockage – ont en commun un aspect d'architecture technique original par rapport à leurs célèbres contreparties centralisées : tous sont basés sur des technologies de réseau en P2P.

En suivant et en essayant de clarifier le « ballet entre programmeurs, logiciels et utilisateurs » (Abbate, 2012) qui construit la décentralisation dans les services Internet, cette thèse explorera les implications socio-politiques de l'approche distribuée et décentralisée à l'architecture technique des services Internet, en faisant l'hypothèse qu'une telle approche aux « couches inférieures » de ces systèmes questionne, ou peut questionner, les usages qu'ils servent, les dynamiques qui y ont lieu, les démarches qu'ils comportent.

Cette introduction comporte deux sections. La première section introduit la problématique de la thèse et les principaux questionnements qui guident le travail d'analyse mené dans les chapitres suivants. Après avoir introduit les éléments fondateurs qui définissent une architecture de réseau décentralisée – notion qui sous-tend l'ensemble des cas d'étude – je discute comment l'utilisation d'une approche d'architecture décentralisée est envisagée par nombre de projets et entreprises comme voie possible pour parer certaines difficultés actuelles de gestion du réseau, et pour proposer un modèle d'organisation différent pour les services basés sur Internet. J'introduis également les travaux qui, en puisant dans l'histoire de l'Internet et en discutant de ses futurs, s'intéressent aux relations entre architecture technique et innovation, et explorent les manières dont les modifications d'architectures ont non seulement des causes mais aussi des conséquences économiques, politiques et sociales. Enfin, cette section soulève la question de la constitution, avec la technologie P2P, d'une *opportunité* sociale, politique et économique pour les services basés sur Internet, pour arriver à formuler les deux grandes questions de recherche que sous-tendent ce travail et leur articulation en groupes de questionnements spécifiques, qui ont trait au partage de compétences entre acteurs, à la formation de dynamiques collectives à la redéfinition des droits des utilisateurs, aux modèles économiques qui émergent avec le distribué.

La deuxième section détaille la démarche méthodologique adoptée dans cette thèse, que porte un regard analytique sur le développement et les premières appropriations de trois services Internet décentralisés. La section introduit les manières dont la thèse analyse l'architecture et les agencements techniques développés dans ces trois cas,

pour voir comment se constituent des dispositifs qui se donnent pour ambition une décentralisation partielle ou intégrale. Ce travail est présenté comme une étude des « couches inférieures » que les applications P2P sous-tendent, au moyen d'une sociologie de l'innovation « en temps réel », une sociologie des architectures de réseau, et une sociologie de la matérialité des dispositifs comme source de droit(s). Enfin, cette section introduit les trois cas d'étude qui constituent la partie centrale de cette thèse – Faroo, un moteur de recherche P2P, Wuala, une plateforme de stockage distribué, et Tribler, un logiciel de streaming vidéo décentralisé – et les choix méthodologiques qui en ont guidé la sélection, avant d'introduire le plan général du reste de la thèse.

## SECTION 1. L'APPROCHE DÉCENTRALISÉE DES SERVICES INTERNET<sup>7</sup>

L'architecture technique qui sous-tend l'Internet actuel – à la fois celle de l'Internet global, et celle des réseaux, systèmes, services qui le peuplent – n'est pas statique, pas plus qu'elle ne s'est imposée d'elle-même grâce à une supériorité technique intrinsèque. L'histoire du « réseau des réseaux » est celle d'une évolution constante, qui répond à une logique de normalisation de fait, liée aux modifications des usages, en particulier à leur massification, et à un ensemble de choix non seulement techniques mais économiques, politiques, sociaux. Tout comme l'architecture de l'Internet et des services Internet a fait l'objet de controverses par le passé, elle est actuellement soumise à de nombreuses tensions, tandis que l'on discute de ses futurs et que, après en avoir reconnu le statut de mécanisme de régulation politique, on commence à en reconnaître pleinement l'importance en tant que levier de développement et de contraintes économiques (van Schewick, 2010), tout en soulignant son aptitude « *by design* » (Braman, 2011) au changement et aux modifications.

Le développement de services basés sur des architectures de réseau décentralisées, distribuées, P2P s'affirme actuellement comme un des axes importants de transformation à moyen terme dans les modes de communication et de gestion des contenus numériques. Co-présents dans les mots des acteurs, comme on verra, ces trois termes – décentralisation, distribution et pair à pair – seront utilisés, tour à tour, dans la thèse pour indiquer des architectures de réseau préconisant que la responsabilité des échanges ou des communications se trouve aux marges ou à la périphérie du système, et que l'ensemble des ressources du système ne se trouve pas dans un même endroit physique, mais est réparti dans plusieurs machines. Leurs définitions et leurs « différences » respectives seront précisées au cours de la thèse ; notamment, le deuxième chapitre sera l'occasion d'explorer plus dans le détail la variété de définitions techniques qui correspondent à ces termes et leur socle commun, et le troisième chapitre explorera l'agencement de la décentralisation et recentralisation partielle des réseaux dans nombre d'exemples « historiques », comme Usenet ou le Domain Name System.

Il convient toutefois de préciser d'ores et déjà que, s'il n'existe pas de définition univoque de ces architectures, on peut identifier certains éléments communs que ces différents termes sous-tendent. Premièrement, les multiples unités de calcul qui composent le réseau (souvent appelées nœuds, dotées chacune de mémoire locale, et communiquant par échanges) ; deuxièmement, la réalisation des objectifs de l'application au moyen d'un partage de ressources (qui peuvent être de plusieurs

---

<sup>7</sup> Cette section doit beaucoup aux discussions et aux travaux menés en collaboration avec Alexandre Mallard et Cécile Méadel dans le cadre du projet ADAM (Architectures distribuées et applications multimédias – <http://adam.hypotheses.org>), financé par l'ANR dans le programme Contenus et Interactions (CONTINT), qui se poursuit depuis novembre 2010 ; elle doit également beaucoup à des travaux menés en collaboration avec Valérie Schafer et Hervé Le Crosnier [Musiani & Schafer, 2011 ; Schafer, Le Crosnier & Musiani, 2011].

types), pour servir un *but commun* (e.g. résoudre un problème de calcul à large échelle) ou des *besoins individuels* au moyen d'une *coordination* de ressources partagées. Troisièmement, la tolérance à la défaillance de nœuds individuels et le conséquent manque de point unique d'échec ; quatrièmement, un passage à l'échelle souple. Enfin, la modification constante de la structure du système (la topologie et latence du réseau, le nombre d'ordinateurs qui y sont connectés), et l'attribution à chaque nœud d'une vision incomplète ou limitée du système. En première approximation, une architecture distribuée n'est pas nécessairement décentralisée (il peut y avoir délégation au réseau distribué d'une tâche qui est ensuite recentralisée), tandis qu'une architecture décentralisée est nécessairement distribuée.

Si le concept de décentralisation est en quelque sorte inscrit dans le principe même de l'Internet – et notamment dans l'organisation de la circulation des flux – son urbanisme actuel semble n'intégrer ce principe que de manière limitée, et certaines de ses faiblesses ont pu apparaître récemment. Par exemple, à la lueur du succès retentissant qu'ont connu certains services comme les *social media* : alors que chaque internaute est devenu, au moins potentiellement, un consommateur mais aussi un distributeur et un producteur de contenus multiples, cet urbanisme conduit à centraliser des masses considérables de données dans certaines régions de l'Internet, alors même qu'elles ont vocation à être rediffusées aussi vite dans de multiples places d'un réseau désormais pleinement globalisé.

Le mode actuel d'organisation de ces services et de la structure du réseau qui rend possible leur fonctionnement, avec ses points de passage obligés, ses carrefours plus ou moins contraints, ses emmagasinages performants, soulève nombre de questions en termes à la fois d'une utilisation optimisée des ressources de stockage, et de la fluidité, rapidité et efficacité des échanges électroniques. S'y ajoutent des interrogations sur la sécurité des échanges et la stabilité du réseau. Une série de dysfonctionnements (voir les doutes que les *black-out* répétés de *Twitter* ont soulevé quant à sa durabilité) et de pannes aux conséquences globales (voir, début 2008, la paralysie mondiale de YouTube et des réseaux pakistanais suite à un routage de requêtes de type *Border Gateway Protocol*<sup>8</sup>, visant à limiter la diffusion de certains contenus dans ce pays), attirent l'attention sur des questions de sécurité et de protection des données qui sont inhérentes à la structure actuelle de l'Internet.

Le recours à des architectures de réseau décentralisées et à des formes d'organisation distribuées pour les services Internet est donc envisagé par nombre de projets, entreprises, services, comme voie possible pour parer certaines difficultés de gestion du réseau, de ses localités à sa globalité, dans une perspective d'efficacité, de sécurité et de développement durable numérique (pour une meilleure utilisation des ressources).

---

<sup>8</sup> Le *Border Gateway Protocol* (BGP) est le protocole qui supervise les décisions de routage sur l'Internet.

## INNOVATION ET INTERNET, UNE QUESTION D'ARCHITECTURE ?

Dans le récent ouvrage *Internet Architecture and Innovation*, Barbara van Schewick souligne que des décennies de recherches sur les processus d'innovation ont permis de mieux comprendre comment les changements du droit, des normes, des prix influencent le contexte économique de l'innovation et les décisions des innovateurs, et vice-versa. Mais il manque encore, selon elle, une compréhension fine des relations entre architecture technique et innovation, au vu du cloisonnement des disciplines qui a pu faire considérer les architectures techniques comme des artefacts « pertinents seulement pour les ingénieurs » (van Schewick, 2010 : 2-3). Rien n'est pourtant plus vrai dans le cas de l'Internet, comme l'a souligné Lawrence Lessig : l'architecture du réseau des réseaux est certes du code, de la technique – le logiciel et les équipements qui définissent le cyberspace – mais elle entraîne et comporte (*embeds*) certains principes, contribue à définir les termes d'utilisation de cet espace, et influence ce qu'il est possible de faire dans cet espace. Ces caractéristiques et possibilités constituent des préconditions au développement du réseau des réseaux ; si certaines architectures invitent l'innovation, d'autres la freinent (Lessig, 1999, 2000).

En précisant que l'architecture ne se réduit pas qu'à une « autre manière » de parler de la technologie, certains travaux récents analysent, dans cette lignée, comment celle-ci renvoie à l'idée d'une pluralité normative pour l'Internet qui intervient, sans discontinuité, sur les usages, le droit et la technique, et contribue finalement à la régulation (Brousseau, Marzouki & Méadel, 2012 ; Massit-Folléa, Méadel & Monnoyer-Smith, 2012). En effet, en choisissant de placer l'intelligence non pas au cœur du réseau Internet mais à sa périphérie, aux deux bouts de la communication (*end-to-end*), ses concepteurs ont assuré un développement continu d'innovations, d'expérimentations et de développements de produits dont la longévité est arbitrée par le succès public (Schafer, Le Crosnier & Musiani, 2011 : 40).

Pour analyser les technologies de l'information et de la communication et les médias « en réseau », l'Internet en particulier, les *software studies* et les *critical code studies* ont récemment relevé le défi de l'interdisciplinarité (Fuller, 2008), en s'appuyant sur un ensemble de travaux en sociologie des sciences et des techniques qui se sont attachés à explorer les qualités sociales et politiques des infrastructures (Star, 1999 ; Star & Bowker, 2002). En outre, certains auteurs, croisant l'informatique, la sociologie, le droit et les *Science & Technology Studies* (STS), montrent qu'une approche méthodologique et une étude innovantes des architectures, intégrant les liens entre celles-ci et les pratiques, est possible (Agre, 2003 ; Elkin-Koren, 2006 ; Braman, 2011). Les modifications d'architectures ont non seulement des causes mais aussi des conséquences économiques, politiques et sociales, et c'est bien de ces allers-retours constants et de ces enjeux dont il est question lorsque les modèles sous-tendant les réseaux, la gestion des flux, les emplacements et les traitements des données sont mis en débat.

Ecosystème en évolution constante, sous l'effet du développement du Web, de ses usages commerciaux et grand public, des (re)compositions de la chaîne de valeur Internet, le réseau des réseaux pose, depuis le développement d'Arpanet à la fin des années 1960, des questions de durabilité des modèles techniques et économiques qui

sous-tendent la gestion de ses flux, jusqu'à questionner son statut de « bien collectif » (Mounier, 2002). Aux origines de l'Internet, le principe de décentralisation gouverne la circulation des transmissions. Ses principaux usagers, les chercheurs, ont conçu des modes de diffusion des données et de facturation extrêmement pragmatiques et parfois techniquement un peu précaires : diffusion des *News* de sites en sites par inondation<sup>9</sup>, tables de routage complexes, facturation des usages en constante évolution et approximative reposant, dans le cas français comme européen, sur le bon vouloir de centres pionniers et d'une communauté d'intérêt. Au cours des années 1980, le « modèle Internet » est doté d'une architecture ouverte, distribuée, qui n'hésite pas à revendiquer le *best effort*<sup>10</sup>, la perte ou la redondance de paquets d'informations, qui privilégie le *running code*<sup>11</sup> plutôt que des protocoles fondés sur l'œcuménisme et la robustesse technique développés au sein des organismes de normalisation traditionnels.

L'Internet voit ses principes fondateurs peu à peu érodés à partir des années 1990, au profit d'outils permettant davantage de gestion des flux et de sécurité. Les analogies, notamment spatiales, permettant de décrire l'évolution de l'Internet ne manquent pas : passage d'un modèle horizontal à un modèle vertical, d'un modèle *bazar* à un modèle *cathédrale* (Raymond, 1997), « minitélisation » ou « broadcastisation » de l'Internet (Schafer & Thierry, 2012), marche forcée vers un « Internet civilisé », balkanisation, transformation de *Netville* en agglomérations de banlieues aseptisées. Autant d'images suggérant que l'Internet, interconnexion de plus de 40 000 réseaux différents<sup>12</sup>, est placé sous une pression considérable face à la volonté croissante de gestion et de contrôle de ses flux – et aux reconfigurations de son modèle générées par son propre succès.

C'est notamment l'introduction du World Wide Web en 1990 (Berners-Lee, 2000) qui conduit à la diffusion d'architectures du type dit client-serveur : le navigateur *Web* et d'autres applications, qui émergent dans les premières phases de la commercialisation de l'Internet, se basent sur un protocole où le client amorce une connexion à un serveur connu, en téléchargement des données et se déconnecte.

---

<sup>9</sup> Protocole de diffusion d'information utilisé aussi bien pour transmettre du contenu que pour le calcul distribué ou parallèle, où chaque nœud transmet l'information à tous ses voisins en une seule étape de temps. (Voir chapitre 2, section 2, « *flooded request* »).

<sup>10</sup> Le *best effort* est un modèle d'acheminement de données dont le principe est de garantir que chaque élément ayant un rôle dans l'acheminement fait le maximum, « tout ce qu'il peut », pour amener ces données de la source à la destination, et ne donne aucune garantie sur la qualité de l'acheminement ni même sur son succès (Voir chapitre 3, section 2).

<sup>11</sup> Part de la célèbre phrase de David Clark, pionnier de l'Internet et de l'Internet Engineering Task Force (IETF), « *We believe in rough consensus and running code* » (on croit au consensus approximatif et à la programmation évolutive), à signifier que l'IETF s'intéresse notamment à la mise en place de systèmes agiles, pouvant être implémentés et changés rapidement.

<sup>12</sup> D'après les estimations de Tony Bates, Philip Smith, Geoff Huston, membres de l'IETF, dans le CIDR (*Classless Inter-Domain Routing*) Report, un ensemble de statistiques sur l'allocation des adresses IP, qui incluent des informations sur le nombre de réseaux autonomes présents sur l'Internet [<http://www.cidr-report.org/as2.0/>, version du 1<sup>er</sup> août 2012].

L'organisation des services de recherche d'informations, depuis Netscape jusqu'au modèle Google, en passant par les portails (*Yahoo!*, AOL, MSN) contribue également à la diffusion de ce modèle. En même temps, des sites comme eBay et Amazon contribuent à l'explosion du commerce électronique ; l'introduction en bourse de Netscape en 1995, suivie par d'autres start-up (eBay, Amazon) et l'augmentation de la valeur du NASDAQ (valeurs technologiques) marque la fin des années 1990, avant l'éclatement de la bulle spéculative en 2000 et la crise boursière qui s'ensuit. Les services Internet qui prennent pied au cours des premières années 2000, comme les réseaux sociaux, les outils de messagerie, les applications de stockage de contenus numériques, se basent sur des modèles techniques et économiques dans lesquels les utilisateurs demandent des services à de puissants serveurs qui stockent de l'information et/ou gèrent le trafic.

Même si le trafic sur l'Internet fonctionne toujours sur le principe de la distribution généralisée, il co-existe donc avec une logique d'organisation des services qui prévoit une concentration autour de serveurs ou centres de données qui habilitent l'accès aux contenus. L'augmentation importante des *spams*, et la présence toujours plus importante de protocoles gourmands en bande passante compromettant le bon fonctionnement de l'Internet en tant que « ressource commune » (Ostrom, 1990), amènent à des interrogations sur la nature coopérative du réseau (Mueller, 2012), tandis que naissent des mesures de gestion qui en érodent la symétrie. Les gestionnaires de réseau déploient, aux points de contact entre les réseaux internes et l'Internet extérieur, les pare-feu – dispositifs qui filtrent les paquets, et choisissent quel trafic laisser passer et à qui nier l'accès. L'augmentation des adresses IP dynamiques, que les fournisseurs de bande passante trouvent utiles pour le déploiement permanent des services Internet, forcent les applications décentralisées à construire des répertoires dynamiques des hôtes auxquels se connecter. Enfin, la popularité de la *Network Address Translation* (ou « traduction d'adresse réseau ») amène à une situation où les véritables adresses des hôtes ne sont plus seulement instables pour ces applications : elles ne sont plus accessibles.

Face au grand mouvement d'*enclosure* qui étend sans cesse le niveau de protection des contenus, des régulations privées se substituent à une puissance publique gérant des droits de propriété intellectuelle par l'intermédiaire de la loi (Boyle, 2003). Parallèlement, le *best effort* recule au profit d'une prioritarisation, à la fois « conditionnelle » et « active » (Curien & Maxwell, 2011 : 43-44), de la part des opérateurs de réseau. Si la première – une optimisation *ad hoc* des réseaux selon ses exigences à un moment donné – ne suscite que peu de réactions, puisqu'elle permet de résoudre des problèmes de congestion, la seconde est davantage critiquée : en plaçant des paquets volontairement en tête de file ou au contraire en discriminant et ralentissant certains d'entre eux, les fournisseurs d'accès à Internet posent la question de la neutralité du « réseau des réseaux », actuellement source de vives controverses (Marsden, 2010 ; Curien & Maxwell, 2011 ; Schafer, Le Crosnier & Musiani, 2011).

La concentration autour de serveurs donnant accès aux services et contenus, le nombre grandissant d'applications consommatrices de bande passante, la confiance croissante dont les utilisateurs investissent le « réseau des réseaux » pour des applications et usages qui entraînent de nouveaux critères de sécurité, résultant en

pare-feu qui divisent l'Internet en réseaux ou régions qui s'autoalimentent et se suffisent à eux-mêmes, ou encore les changements d'échelle de l'Internet, créent des défis de durabilité. De nombreux acteurs du secteur IT, à commencer par les entreprises, s'interrogent sur le mode d'organisation des services et la structure du réseau, et identifient des problèmes spécifiques qui mettent à l'épreuve l'architecture Internet dans son scénario de « concentration » actuel. La recherche d'alternatives au mode dominant d'organisation des services et de la structure du réseau – en quête d'efficacité et de durabilité – se poursuit. En cherchant les meilleures solutions, certains se retournent vers l'Internet d'il y a quinze ou vingt ans – aux qualités persistantes d'une ancienne technologie qui replonge dans la topologie de l'Internet pré-commercial, mettant à profit les « marges » du réseau : le P2P.

### LES MARGES DU RÉSEAU, UNE RÉPONSE À L'AVENIR « NUAGEUX » DE L'INTERNET ?

L'architecture en pair à pair désigne un modèle de réseau structuré afin que les communications ou les échanges se fassent entre nœuds qui ont la même responsabilité dans le système. La dichotomie entre les serveurs, fournisseurs de ressources, et les clients demandeurs de ressources, caractéristique du modèle *client-serveur*, est remplacée par une situation où tous les pairs ont la ressource et tous les pairs la demandent (Schollmeier, 2001). L'architecture P2P embrasse le principe de décentralisation en utilisant le réseau d'une manière différente par rapport aux applications client-serveur. Dans cette architecture, les usagers demandent des services à des serveurs de capacité limitée ; par conséquent, l'addition de clients peut ralentir le transfert de données pour tous les usagers<sup>13</sup>.

Dans l'architecture P2P, les usagers n'utilisent pas seulement des ressources pour accéder aux contenus (de la bande passante, de l'espace de stockage, de la capacité de calcul) mais ils en fournissent également – si la demande à laquelle le système doit répondre augmente, la capacité totale du système augmente aussi. Les systèmes P2P peuvent également présenter des avantages de stabilité et d'endurance, car la nature distribuée du système en améliore la force et empêche son invalidation totale en cas d'échec d'un des nœuds. L'efficacité du P2P en tant que modèle de distribution est strictement liée au partage de la capacité de calcul et de bande passante entre tous les composants du système, ce qui change la structure de distribution et l'allocation des coûts en augmentant l'utilisation de bande passante au niveau du réseau, plutôt qu'à celui du serveur. Par ailleurs, les vulnérabilités potentielles des systèmes en P2P résident surtout dans les codes non sécurisés, susceptibles d'autoriser l'accès à distance aux ordinateurs qui forment le réseau, et même de le compromettre. La sécurité et les mécanismes de vérification de fichiers (tels que la fonction de hachage<sup>14</sup>, la vérification des fragments qui vont recomposer le fichier, et une variété

---

<sup>13</sup> Et conduire, dans des situations limite, à des scénarios de déni de service.

<sup>14</sup> La fonction de hachage est une fonction qui, à partir d'une donnée fournie, calcule une *empreinte* servant à identifier rapidement, bien qu'incomplètement, la donnée qui a été à son origine. Cette fonction sert, en informatique, à rendre plus rapide l'identification des données : calculer l'empreinte d'une donnée ne coûte qu'un temps négligeable par rapport au calcul de la donnée initiale.

de méthodes de cryptage) sont par conséquent les domaines dans lesquels les applications basées sur le P2P évoluent actuellement le plus.

Au cours de leur histoire « grand public », encore relativement brève, les réseaux P2P ont été considérés presque exclusivement comme une menace pour l'industrie des contenus numériques. L'usage principal de ces réseaux par le public étant le partage non autorisé de fichiers musicaux ou vidéo, le problème des droits de propriété intellectuelle, du droit d'auteur notamment, s'est imposé en tant que cadrage médiatique et politique prédominant des réseaux P2P et de leurs usages (Elkin-Koren, 2006). Cette réputation s'est tout particulièrement forgée avec l'avènement, au tournant des années 2000, de pratiques d'échange de contenus à l'échelle mondiale, impliquant des millions d'utilisateurs. Le cas le plus emblématique est celui de l'explosion, en 1999, du service Napster ; avec ses soixante millions de « partageurs », ce service a permis le partage des fichiers entre plusieurs nœuds. Mais la présence d'un serveur central visant à faciliter l'identification des contenus, point de passage obligé de chaque échange et cible facile pour les « ayant droit », a finalement causé sa perte. Une seconde génération de systèmes de partage de fichiers en P2P, utilisée par exemple par *Gnutella*, a dépassé la nécessité d'un serveur central et démontré les avantages et la robustesse d'un système décentralisé. Les systèmes de troisième génération sont hybrides et constituent un mélange des deux précédents ; ils ont été massivement employés par *KaZaA* et d'autres applications de partage de fichiers (Beuscart, 2002 ; Farchy, 2003 ; Laflaquière, 2005).

L'importance que le débat autour des violations du droit d'auteur, facilitées par ces échanges, a revêtue dans les médias et l'opinion publique s'est reflétée dans les priorités de recherche, à la fois pour les projets à caractère technique/informatique et les recherches en sciences économiques et sociales. Pour autant, une telle conception ne saurait résumer à elle seule la philosophie des systèmes en P2P : des auteurs comme la juriste Niva Elkin-Koren, ou l'informaticien et sociologue David Hales soulignent de façon originale que la signification politique et technique de ces dispositifs serait à chercher ailleurs, notamment dans un ensemble de « *vertus* » (Elkin-Koren, 2006), des qualités d'efficacité, de stabilité et d'endurance, qui tiennent à la fois du technique, du social, de l'économique et du légal.

En premier lieu, la transmission directe de données entre les machines d'un réseau décentralisé, avec éventuellement un principe de fractionnement des fichiers, est susceptible de promouvoir une meilleure efficacité des échanges de contenus, une plus grande liberté et éventuellement l'émergence de nouveaux principes organisationnels, sociaux et légaux. Au-delà de ces bénéfices, significatifs du point de vue de l'économie des échanges, on a également souligné que les systèmes P2P peuvent fournir des solutions spécifiques pour la protection des libertés personnelles (Wood, 2010) ou l'émergence de processus de décision alternatifs et des environnements participatifs (Elkin-Koren & Salzberger, 2004) rendus possibles par l'échange direct de contenus entre les différents nœuds du réseau (Hales, 2006). Les implications sont multiples, en termes de performance technique, mais aussi pour redéfinir des concepts tels que la sécurité et la *privacy*, reconfigurer les emplacements des données et des échanges, les frontières entre l'utilisateur et le réseau, les outils à disposition. En somme le P2P peut avoir un impact sur l'attribution, la

reconnaissance et la modification de droits entre utilisateurs et fournisseurs des services.

Si les modèles P2P offrent des nouvelles pistes d'exploration et de maintien des équilibres au sein de l'écologie Internet, les grands acteurs du marché IT actuel posent des conditions. Un « *P2P turn* » est-il encore possible, alors que le *cloud computing*, l'« informatique dans les nuages », semble prendre le dessus, renforçant donc un modèle économique et technique dans lequel l'utilisateur final sollicite de puissants centres de serveurs, qui stockent l'information et gèrent le trafic sur le réseau ?

Bien que la définition même de *cloud* fasse actuellement l'objet de vives controverses (une revue spécialisée a réuni plus de vingt définitions différentes du concept<sup>15</sup>), ce modèle indique généralement que le fournisseur propose l'infrastructure physique et le produit logiciel, abritant ainsi à la fois les applications et les données dans un lieu éloigné de l'utilisateur (le fameux « nuage », *cloud* en anglais) et interagit avec ce dernier grâce à une interface client. On s'achemine dans ce cas vers un modèle de déportation sur des serveurs distants de traitements informatiques traditionnellement localisés sur le poste utilisateur, et d'un agrandissement de la distance entre le déroulement des opérations nécessaires au fonctionnement du service, et l'équipement personnel/local des clients de ce service.

Toutefois, le nuage décentralisé est aussi envisageable, et il serait conçu pour répartir la puissance de calcul et les ressources du nuage entre les terminaux de tous les utilisateurs/contributeurs. C'est le souhait exprimé par Eben Moglen, l'inspirateur du projet de réseau social décentralisé Diaspora\*. Le chercheur et activiste note que, en l'état, le *cloud* n'a profité qu'aux grands acteurs : son développement a favorisé une situation où « des serveurs ont gagné [davantage de] liberté. Liberté de bouger. Liberté de louvoyer ; de combiner et de diviser, de ré-agréger et d'utiliser toute sorte d'astuces. Les serveurs ont gagné en liberté. Les clients n'ont rien gagné » (Moglen, 2010). Il s'agit là d'une affirmation très tranchée et peut-être d'une simplification excessive : le modèle du *cloud*, même sans en envisager une version décentralisée, apporte bien quelque chose aux clients/utilisateurs finaux – l'accès permanent au service, en mobilité (Mowbray, 2009). Il reste cependant intéressant de noter que la décentralisation pourrait être une réponse possible à l'avenir « nuageux » de l'Internet, susceptible de relier les grandes inconnues de ce modèle – notamment qui est responsable du traitement des données et des flux, et de quelles manières – aux machines des utilisateurs.

## LE P2P COMME OPPORTUNITÉ ET ALTERNATIVE

Telle est donc la problématique dans laquelle se situe cette thèse : dès que l'on

---

<sup>15</sup> « Twenty-One Experts Define Cloud Computing », *Cloud Computing Journal*, 2008 (<http://cloudcomputing.sys-con.com/node/612375/print>).

agrandit la focale rivée sur le partage de fichiers, les appropriations du P2P soulèvent la question de la constitution, avec cette technologie, d'une *opportunité*<sup>16</sup> sociale, politique et économique pour les services basés sur Internet. Cette thèse est également une exploration de la décentralisation comme *alternative* pour ces services, dans la mesure où la norme actuelle pour les services hébergés par le « réseau des réseaux » est la concentration autour de serveurs donnant accès aux contenus. Ce travail a l'intention de contribuer à une connaissance des problématiques spécifiques que posent les architectures de réseau décentralisées du point de vue des usages, de la diffusion des technologies, des modèles économiques, des droits et normes expérimentés ou éprouvés, des contenus distribués.

La thèse s'articule autour de deux questions de recherche principales ayant trait aux implications socio-politiques de l'approche distribuée et décentralisée de l'architecture technique des services Internet. Qu'est-ce que dessine une architecture de réseau décentralisée du point de vue de l'articulation des acteurs et des contenus, de la répartition de responsabilités, de l'organisation du marché et de la capacité à exercer du contrôle, des formes d'existence et des rôles d'entités telles que les nœuds du réseau, les usagers, les unités centrales ? Sous quelles conditions un réseau qui répartit la responsabilité de son fonctionnement à ses marges, et qui suit un modèle non hiérarchisé ou hybride, peut-il se développer dans l'Internet d'aujourd'hui ?

Ce travail tente d'apporter quelques éléments de réponse à ces questions en suivant les développeurs de trois services Internet, construits sur un modèle de réseau décentralisé – un moteur de recherche, un service de stockage et une application pour le streaming vidéo –, les collectifs d'usagers pionniers qui se développent avec ces services, et ponctuellement, les arènes politiques où l'on discute de l'organisation et de la gouvernance de l'Internet à moyen et long terme. Au fil des pages et des cas d'étude qui constituent la partie centrale de la thèse, on verra ces deux questions s'articuler de trois manières spécifiques.

Quels nouveaux partages de compétences entre fournisseurs de service, opérateurs des réseaux et utilisateurs ces applications sont-elles susceptibles d'engendrer ? On verra en effet comment les applications P2P instaurent des configurations nouvelles dans l'univers des pratiques de coopération et de communication : l'accès aux usages suppose en contrepartie que soient mises à disposition du collectif ou du service des ressources propres à l'utilisateur, ou considérées comme telles jusqu'ici. En analysant les nouvelles formes d'engagement de l'utilisateur et des autres acteurs concernés par la mise en œuvre des services, on cherche à comprendre les modalités de leur viabilité technique et économique, mais aussi comment sont réglées des questions « classiques » associées à la gestion des contenus numériques, comme celle de la sécurité des données ou de la *privacy*.

Quelles sont les nouvelles dynamiques collectives, à l'échelle du réseau, qui sont susceptibles de se mettre en place avec le développement du P2P ? Ces services

---

<sup>16</sup> Dans une acception peu « classiquement française » mais de plus en plus mobilisée par les acteurs eux-mêmes aussi bien que dans la littérature, je donne au mot « opportunité » la signification de « possibilité » ou « occasion », dérivée de l'anglais *opportunity*.

proposent en effet, selon des configurations variées, l'aménagement de nouvelles relations entre le local et le global. On assiste à la mise en opération, par exemple, d'un service de stockage distribué qui est basé sur la fragmentation des fichiers, et leur dissémination dans l'ensemble du réseau d'utilisateurs ; ou d'un autre qui conceptualise un moteur de recherche « par affinités », dans lequel l'utilisateur est défini par une capacité à établir un lien avec des savoirs qu'il détient, du point de vue de sa localisation propre. Ces applications posent des questions inédites du point de vue de l'articulation entre l'individu et le collectif, dans l'espace du réseau ou dans l'espace géographique. Comment du global se recrée-t-il à partir de l'éclatement du local ? Quelle visibilité sur le collectif ces systèmes offrent-ils à un utilisateur qui devient un nœud actif, au sens fort du terme, et non simplement un participant connecté au réseau global ? Quelles sont les nouvelles inégalités numériques qui sont susceptibles de s'instaurer dans ces situations ?

Quelles sont les formes de régulation émergentes de ces technologies, du point de vue des droits de propriété, du droit à la vie privée et à la confidentialité des données personnelles, ou encore des droits d'accès aux contenus ? Les frontières de la « légalité » de contenus, usages, plateformes sont-elles définies et configurées de manière assez souple pour respecter et accompagner l'innovation ? De la même façon que les réseaux P2P « de première génération » ont mis au défi les dispositifs de la propriété intellectuelle classique, cette thèse explore les manières dont les architectures décentralisées et P2P pour les services Internet contribuent à une redéfinition de notions comme celle d'auteur et celle de contributeur, à une reconfiguration des droits des utilisateurs, ainsi qu'à des déplacements de la frontière entre usages privés et usages publics.

Enfin, la thèse touchera de manière plus ponctuelle à la recherche de modèles économiques pour une technologie qui est en principe strictement associée à l'économie de la gratuité : quels modèles d'affaires sont susceptibles de prendre en charge le développement de ces applications ? Un premier enjeu est celui de l'allocation optimale d'une ressource qui est en partie « publique ». La question de l'allocation des ressources de stockage, et indirectement de la fluidité des échanges électroniques, si elle ne se pose pas encore avec acuité, s'inscrit dans le moyen terme. Le développement de services basés sur des architectures décentralisées peut être analysé comme une stratégie de contournement de ces problèmes d'allocations. La question est dès lors de déterminer quels sont les acteurs qui ont un intérêt à développer ces types d'applications : est-ce les développeurs/entrepreneurs, qui voient dans l'architecture décentralisée une opportunité de baisser les barrières à leur entrée dans un secteur où la concurrence est aigüe ? Les utilisateurs, porteurs d'une demande de fluidité, ou de sécurisation, ou de confidentialité des échanges ? Ou encore, les régulateurs et les pouvoirs publics, qui identifient des opportunités dans le fait de mener une politique d'encouragement à ces stratégies alternatives ? Un second enjeu est celui du développement durable de l'écologie Internet : le choix d'une approche qu'un des acteurs définira de « force brute<sup>17</sup> », c'est à dire la copie de l'intégralité des contenus circulants dans un système sur des machines sous contrôle direct d'une firme, ou d'une approche où ceux-ci sont repartis sur les terminaux des

---

<sup>17</sup> Entretien avec Georg Schneider, développeur du moteur de recherche décentralisé Faroo.

utilisateurs, n'est pas sans conséquences pour le type, la quantité et la qualité des ressources informatiques et économiques mobilisées par les différents acteurs concernés.

## SECTION 2. TROIS CAS POUR UN REGARD COMMUN

La démarche méthodologique adoptée dans cette thèse propose de porter un regard détaillé sur le développement et les premières appropriations de trois services Internet décentralisés, qui constituent mes cas d'étude. Tout en répondant à des nécessités d'usages diverses (moteur de recherche, stockage de données, streaming vidéo), ces projets et applications ont en commun un aspect d'architecture technique original par rapport aux applications plus répandues servant les mêmes usages: tous sont basés sur des technologies de réseau en P2P.

La thèse analyse l'architecture et les agencements techniques développés dans ces trois cas, pour voir comment se constituent des dispositifs qui se donnent pour ambition la décentralisation – de manière au moins partielle ou hybride, et dans un cas, intégrale. Ce travail suit les représentations que les innovateurs dans le domaine des services Internet en P2P se font de leurs usagers, quels sont les collectifs actuels et futurs, les pratiques, les formes d'organisations qu'ils envisagent et qu'ils tendent à faire exister en lien avec les technologies qu'ils élaborent. Au moyen des observations de cas et de dispositifs, des analyses des contenus stockés, échangés, recherchés, des entretiens avec les chercheurs et les entrepreneurs responsables des projets, mais aussi avec les parties prenantes impliquées dans le développement de l'architecture technique de l'Internet, j'essaie de retracer et de suivre les choix, les façons de faire, les transformations et modifications des services Internet décentralisés, qui en déterminent les formes et caractéristiques actuelles, ainsi que leur évolution dans le temps.

Si la partie centrale de cette thèse est donc constituée par une analyse monographique détaillée, l'un après l'autre, des trois cas sélectionnés, ce travail se termine par un regard commun porté sur ces trois projets, et sur la pluralité des enjeux que leur analyse a contribué à dégager. Vu le nombre limité de cas, ce travail ne prétend pas à l'exhaustivité ; tâche dont un travail de cartographie des projets et expérimentations en matière d'architecture distribuée, mené en parallèle et à la suite de mon terrain<sup>18</sup>, a par ailleurs révélé toute la difficulté, au vu de l'importance du travail de définition, de classification et de catégorisation qu'elle implique.

Pourtant, la thèse souhaite, par ce regard commun, essayer d'échapper à l'écueil classique des suivis d'innovation au cas par cas. Il s'agit d'éviter que les indications fournies par ces analyses ne restent pas éparpillées dans des projets indépendants, et de permettre à une vision d'ensemble des réactions et des pistes d'appropriation de ces technologies de donner lieu à un repérage de convergences dans les retours de ces expériences – retours porteurs de développements ultérieurs, et d'applications possibles.

---

<sup>18</sup> Il s'agit de la tâche 2 du projet ADAM, « Exploration des projets et expérimentations en matière d'architecture distribuée ». L'objectif de cette tâche est de procéder à un inventaire, une exploration et une systématisation des connaissances disponibles concernant les technologies de réseau de type architecture distribuée, avec l'idée que cette variété de termes et d'architectures correspondantes n'a pas été encore très bien perçue par les sciences sociales, alors même qu'elle n'est pas sans conséquences sur la variété de leurs possibles implications sociales et économiques.

Dans le but d'essayer d'obtenir une vision commune des pistes d'appropriation de la technologie P2P, l'enquête empirique à l'œuvre dans cette thèse répond donc à une démarche d'identification de technologies et d'usages qui sont, pour reprendre l'expression de Bruno Latour, *en train de se faire* (Latour, 1987). La sociologie des techniques « en temps réel » que l'on expérimente ici nous semble une méthode viable afin d'appréhender des situations variables aux dimensions diverses, et de tirer des conclusions sur leurs possibles développements et applications. D'une part, parce que le terrain que l'on cherche à explorer est un monde fortement concurrentiel, qui connaît des innovations permanentes et rapides, ainsi qu'un fort renouvellement des approches technologiques explorées (Delmas-Marty, 2012 : 14)<sup>19</sup> ; d'autre part, et peut-être plus fondamentalement, parce que ces projets, comme d'autres dans le domaine du numérique, cassent la vieille opposition entre le *technology push* et le *demand pull* (Douthwaite, 2002).

#### *LE P2P, INNOVATION EN CONSTRUCTION*

Il ne s'agit en effet ni de retrouver, dans l'élaboration ou le façonnage des services en P2P observés, des éléments de compréhension des usages attendus ou prescrits ; ni de décrire les besoins, cultures ou valeurs opérant dans le domaine pour aboutir à des préconisations sur la forme et les propriétés des dispositifs techniques. Les

---

<sup>19</sup> Il est d'ailleurs nécessaire de préciser ici quelques termes relatifs aux différents stades de développement et publication d'un logiciel (*software release life cycle*), termes qu'on retrouvera souvent au fil des chapitres suivants. Voir aussi [Humble & Farley, 2010].

La phase de *pre-alpha* se réfère à toutes les activités exercées au cours du projet de logiciel avant la phase de test. La phase *alpha* du cycle est la première phase de test du logiciel ; elle est généralement fermée au public et limitée éventuellement à des testeurs pressentis par les développeurs eux-mêmes (on parle de *closed alpha*), car le logiciel est encore instable et pourrait causer des plantages ou des pertes de données. Parfois, l'alpha est disponible publiquement (comme bonus, ou pré-commande, *public alpha*), dans ce cas, les développeurs poussent plus tôt dans leur démarche vers une plus grande stabilité, pour améliorer la pertinence des essais de ces testeurs pionniers.

La phase *beta* commence lorsque les fonctionnalités du logiciel sont complètes ; un logiciel dans cette phase comporte encore, généralement, de nombreux bugs et des problèmes de vitesse et de performance. Cette phase marque généralement la première mise à disposition du logiciel à l'extérieur de l'organisation qui l'a développé, et l'accent est mis sur la réduction de ces problèmes de fonctionnement ; elle incorpore donc souvent des tests d'utilisabilité. Les développeurs publient soit une version beta fermée (*private/closed beta*) ou ouverte (*public beta*) ; la première est destinée à un groupe restreint de personnes, sur invitation, tandis que la deuxième est ouverte à un grand groupe, ou toute personne intéressée. Les utilisateurs/testeurs rapportent les bugs qu'ils trouvent, et suggèrent parfois de nouvelles fonctionnalités supplémentaires qu'ils estiment devraient être disponibles dans la version finale.

Enfin, la publication définitive (*release*) du logiciel est précédée par un *release candidate*, une version beta avec le potentiel d'être un produit final.

architectures décentralisées qui sous-tendent les dispositifs observés dans cette thèse font partie et témoignent justement d'un monde où la coupure entre usager et concepteur est souvent brouillée, et la place des technologies transformée ; où l'utilisateur intervient étroitement dans les choix technologiques, et la technologie provoque des usages imprévus ; en bref, l'architecture du réseau transforme aussi étroitement les pratiques qu'elle est transformée par elles. Mon choix de traiter le récit « avec l'observateur », à la première personne et d'un ton *embedded* sur le terrain, cherche d'ailleurs à refléter dans le style d'écriture cette hypothèse fondatrice de mon travail. Je cherche ici à faire un récit qui se donne à découvrir, avec des strates de transformation qui s'effacent successivement, mais que l'on retrouve dans la technologie dès lors que le récit vient leur donner sens.

Cette thèse met donc en œuvre des enquêtes auprès de trois services Internet en P2P afin de constituer, « en temps réel » et de façon spécifique d'abord, transverse ensuite, aux applications concernées, un lieu de capitalisation sur des technologies et des usages qui sont en train de se faire. Ce travail souhaite ainsi contribuer à appréhender dans leur globalité des questionnements sous-jacents, qui engagent tout autant la construction de la « démocratie numérique » (Vanbreemersch, 2009 ; Cardon, 2010), le façonnage d'un Internet plus égalitaire et horizontal, que la viabilité économique des innovations produites dans ce domaine. Ces questionnements renvoient aux nouvelles compétences que les dispositifs attribuent à leurs usagers, aux problèmes de gestion techno-juridique des contenus qu'ils posent, aux dynamiques d'échange et de partage communautaires spécifiques qu'ils instaurent ou sont susceptibles d'instaurer, aux formes de gouvernance qu'ils supposent ou tendent à faire exister.

### *UNE SOCIOLOGIE DES ARCHITECTURES DE RÉSEAU*

En utilisant cette approche, je souhaite par ailleurs dépasser, et souligner les limites, de la démarche qui est aujourd'hui la plus courante quand on prend le P2P comme objet d'étude : même lorsqu'elles se focalisent sur des formes d'organisation en P2P (par exemple en étudiant la coopération dans les communautés en ligne), les recherches en sciences sociales ont jusqu'ici contribué à la tendance qui réduit le P2P aux usages qu'il rend possible, un d'entre eux en particulier : le partage (illégal) de fichiers. Cette thèse veut donc souligner l'importance d'étudier le lien entre la façon dont les applications prennent forme et leurs influences possibles sur les pratiques, les droits et les relations sociales. Mon travail se veut un exemple d'étude des « couches inférieures » que les applications P2P sous-tendent, une analyse de leur incidence sur les types d'échange qui y ont lieu et sur les caractéristiques de leurs utilisateurs – afin de faire ressortir les façons dont les attributs de la technologie sont susceptibles d'informer pleinement des questions en effet cruciales pour les usages, tels que les traitements et les emplacements physiques des données, la gestion des ressources de calcul, l'extraction des informations. Il est d'autant plus important d'éclaircir ces questions car, du fait de l'« invisibilité » de ces couches inférieures du réseau, les utilisateurs n'ont souvent pas une connaissance directe de leur influence sur les usages.

En m'appuyant sur des auteurs que je discuterai plus largement dans le prochain chapitre (Agre, 2003 ; Elkin-Koren, 2002, 2006, 2012 ; DeNardis, 2009 ; van Schewick, 2010 ; Braman, 2011), cette démarche me voit donc attribuer une attention particulière à un aspect de la technologie P2P tout à fait « discret » et même invisible aux yeux des utilisateurs: leur architecture. Que nous apprennent la mise en place et le formatage de liens, nœuds, points de passage obligés, protocoles de propagation de l'information – en un mot, les architectures – des applications basées sur la technologie P2P, et des opportunités qu'elles présentent ? Je m'intéresse à la structuration des architectures en faisant l'hypothèse que cette forme de distribution questionne les procès et les usages, dans l'idée qu'une analyse ayant comme point de départ le P2P en tant que architecture technique d'applications diverses – plutôt que le partage de fichiers, ou un autre usage spécifique fait de cette technologie parmi plusieurs possibles – est un instrument plus approprié pour comprendre les changements présents et futurs favorisés par le P2P et les acteurs sociaux qui le développent, utilisent, réglementent. En prenant les architectures, artefacts transparents pour l'usager par choix délibéré de leurs créateurs, comme objets d'étude, ma démarche est informée par les travaux en *Science and Technology Studies* (STS) sur les infrastructures en tant que systèmes socio-techniques en constante évolution, informés non seulement par des éléments physiques invisibles à l'utilisateur final, mais aussi par des facteurs tels que l'organisation sociale et la connaissance (Star, 1999 ; Star & Bowker, 2002).

#### *LA MATÉRIALITÉ DES DISPOSITIFS COMME SOURCE DE DROIT(S)*

Finalement, l'approche adoptée dans cette thèse se doit de prêter une attention spéciale à l'articulation entre droit(s) et technologies P2P. Il s'agit d'éclaircir les manières dont l'agencement collaboratif de ressources informatiques, la localisation ou la globalisation de flux de données, le filtrage ou la prioritarisation de types de trafic contribuent à façonner des portraits et des pratiques d'utilisateurs dotés de certains droits, et en sont façonnés en retour.

Cette dimension techno-juridique est désormais structurante pour comprendre l'avenir de services qui ont vocation à organiser la circulation à grande échelle d'une diversité de contenus, et qui peuvent dans certains cas passer rapidement du stade de projet embryonnaire au stade de service utilisé par des milliers d'utilisateurs. Une attention particulière à cette dimension se justifie d'autant plus pour des applications comme celles du P2P, qui tendent à reconfigurer le statut des frontières séparant l'usager du reste du réseau, et qui exigent de sa part une prise en compte plus ou moins importante de la matérialité des dispositifs qui l'y rattachent.

Ce volet de l'analyse sous-tend donc trois objectifs. Il s'agit en premier lieu de mener à bien l'analyse de la « légalité » des services pris en considération, dont l'architecture évolue constamment et dont la loi écrite a souvent du mal à suivre les rythmes trop rapides et les directions imprévisibles de changement. Pour autant, la notion de légalité et l'analyse du caractère « légal » de ces systèmes n'est pas

comprise ici comme une exploration de la *conformité* des dispositifs aux lois existantes, dans une acception plus classiquement juridique. Au centre de ce travail sont plutôt le suivi et l'explicitation de comment l'architecture interroge et formate les dispositifs (tels que, comme on le verra, le mot de passe ou les « clés d'amitié ») pour repartir de manières différentes ou novatrices l'allocation des droits, entre utilisateurs, fournisseurs de service, opérateurs de réseau, fournisseurs d'accès ; et de comment, symétriquement, les usages et les procès viennent modifier la formulation et l'exécution des droits ouverts aux usagers.

L'adoption de cette perspective entraîne deux autres implications. D'une part, ce travail essaye de fournir des outils nécessaires au dépassement d'une analyse des rapports entre le droit et la technologie qui se focalise trop souvent sur un seul aspect : le fait que les technologies émergentes posent des défis aux régimes légaux existants, et créent un besoin de réforme de ces mêmes régimes légaux. Finalement, il s'agit de mettre en œuvre une étude de la matérialité des dispositifs comme source de droit et de droits : que les objets et ressources inscrits et produits avec le P2P puissent eux-mêmes être conçus et traités en tant que systèmes de définition et protection des droits de l'utilisateur des services Internet.

#### *IMAGINAIRES ET ÉPREUVES, CONCEPTEURS ET UTILISATEURS*

La thèse articule deux entrées complémentaires dans l'analyse de la construction des services décentralisés, qui vont croiser les trois ensembles de questionnements posés à chacun des cas – le P2P comme forme particulière d'innovation en construction et en temps réel, le façonnage réciproque des architectures P2P et des procès/usages, et la matérialité des dispositifs comme source de droit.

La première entrée, et celle qui a le plus de place dans ma démarche, s'adosse au suivi des concepteurs. Il s'agit d'identifier leurs stratégies dans la construction des technologies et le montage des modèles d'affaires associés, et également de comprendre les types de valeur ou de culture auxquels ils se réfèrent dans les usages qu'ils projettent. L'expérience montre, dans de nombreux contextes qui dépassent sans doute le P2P et les services Internet, que les imaginaires techniques jouent un rôle important dans la manière dont se forme la légitimité scientifique, mais aussi politique, culturelle et sociale des projets. En même temps, je cherche à questionner la dimension plus idéologique ou utopiste – dans la mesure où ceux-ci se disent porteurs d'une idée d'Internet plus déconcentré et dispersé – des projets en question. Ce caractère utopiste est pris parfois comme objet pour essayer de montrer en quoi il entraîne des façons de faire, explique des choix, trouve des échos. Dans cette lignée, on attribue une attention particulière à l'observation des passages, transformations, négociations, modifications des objets et moments d'épreuve des projets et dispositifs, au delà des « programmes » établis.

La seconde entrée s'intéresse au rôle actif joué par les usagers et les communautés qu'ils forment. Cette entrée se réfère notamment à deux hypothèses que suggèrent les recherches de sociologie ou d'économie concernant la contribution des usagers aux innovations, que les usagers soient enrôlés au cœur de la conception des dispositifs

techniques (cas des « usagers-innovateurs », Akrich, 1998, von Hippel, 2005) ou qu'ils interviennent en aval de cette phase. D'une part, les technologies de l'information et de la communication (TIC) produisent des objets qui, au moment de leur invention et aujourd'hui, de plus en plus, aux moments de leur appropriation, sont ouverts et laissent une large part d'évolution, d'interprétation et de réorganisation par les utilisateurs. D'autre part, parce que les TIC renvoient à une dynamique d'interaction fine entre les pratiques culturelles, les représentations sociales et les propositions de contenu véhiculés sur les supports. Les médias comme dispositifs à la fois techniques, sémiotiques et sociaux engagent toujours une vision de l'*homo communicans* et ont ainsi un impact direct sur les formes d'interaction et les productions intellectuelles. Ces impacts, qui rétroagissent parfois directement sur les contenus techniques des services, sont d'autant plus importants que ces services sont associés explicitement, ou donnent prise implicitement, à des dynamiques communautaires. A cet égard, la durée du travail de terrain (presque deux ans, de l'été 2009 au printemps 2011), a permis non seulement de voir s'opérer la maturation technique mais aussi d'observer le rôle dynamique joué par les usagers pionniers dans la mise en œuvre des applications.

## INNOVATION ET INNOVATEURS DANS LE P2P

Il suffit d'un regard rapide à SourceForge, le plus grand « dépôt de code source » existant sur l'Internet, pour s'en rendre compte : ce ne sont pas les expérimentations avec des architectures de réseau décentralisées qui manquent dans le monde du logiciel libre et de l'open source, ni dans la recherche universitaire<sup>20</sup>. Pourtant, si le logiciel libre et le P2P ont entamé dans le passé nombre de collaborations fructueuses – au point que les deux termes sont souvent, de manière imprécise, compris et utilisés de manière interchangeable<sup>21</sup> – les projets auxquels j'ai fini plus particulièrement par

---

<sup>20</sup> Une recherche « P2P » sur SourceForge donne 42 pages de projets et téléchargements possibles, chacune d'entre elle comporte une vingtaine de résultats (<http://sourceforge.net/directory/?q=P2P>).

<sup>21</sup> Beaucoup d'outils en P2P, en particulier les premiers grands systèmes de partage de fichiers, sont effectivement nés au sein des communautés de logiciel libre et de l'open source et en ont, à leur tour, facilité le développement et l'organisation, dans une démarche éthique commune de partage de ressources, de gestion consensuelle et sans centre, d'attribution d'importance au choix et à la liberté de l'utilisateur. Pourtant, un nombre important d'applications, sous-tendant une technologie P2P et servant des usages variés, sont à ce jour partiellement ou complètement propriétaires. Comme a souligné en 2000 le développeur Dave Winer, « *The P in P2P is People* » : c'est-à-dire, ce qui est important dans les réseaux P2P, ce sont les gens (Winer, 2000). Ce commentaire souligne en quoi la connexion entre le développement d'applications P2P et le mouvement open source est significatif : les projets open source s'organisent autour de groupes de travail décentralisés, qui s'auto-gèrent et sont eux-mêmes rendus possibles par des technologies Internet en P2P. Si le P dans P2P est celui de « People », note Tim O'Reilly, ceux qui veulent travailler dans l'espace P2P ont d'importantes leçons à tirer, non seulement des technologies facilitant la création de communautés qui s'auto-organisent, mais aussi des cadres organisationnels développés afin de gérer ces communautés (O'Reilly, 2000). L'open source n'est pas simplement déterminé par un ensemble de licences pour la distribution des logiciels, mais, à un niveau plus profond, par un ensemble de techniques pour un développement collaboratif et global de logiciels. C'est là que, en effet, que la boucle entre l'open source et le P2P se boucle, comme l'avait déjà montré un des moteurs de la première communauté open source, Usenet :

m'intéresser dans le cadre de mon travail sont, dans deux cas sur trois, totalement ou partiellement propriétaires ; et le troisième est né en milieu universitaire, mais il a été, au cours de sa brève histoire, bien plus « saisi » que ses semblables par des acteurs politiques et du secteur privé. Il convient de noter cet aspect en tant que premier choix méthodologique de taille adopté dans cette thèse : il s'agit d'un choix de cas qui n'est pas représentatif du nombre de projets de développement libre par rapport au nombre de projets propriétaires dans le domaine du P2P, en général, mais qui a pu me donner, pour cette raison et pour d'autres que je précise par la suite, des « prises » originales sur le terrain par rapport à des projets qui restent exclusivement dans le domaine de la recherche universitaire ou des communautés open source. En m'intéressant à la viabilité technique, économique et politique de l'alternative décentralisée pour les services Internet, je ne pouvais qu'être amenée à privilégier ces arènes – bien que plus rares, et certes minoritaires – où des embryons de modèles d'affaires et de mise en marché des dispositifs étaient présents. Cela entraîne également la présence d'une communauté d'utilisateurs/clients pionniers, autre importante prise sur le terrain dont j'ai cherché à m'assurer l'accès dans la phase de sélection des cas d'étude.

### *LES CAS D'ÉTUDE*

Les trois cas qui constituent la partie centrale de cette thèse sont donc représentatifs d'une catégorie toute particulière de dispositifs décentralisés. Il s'agit de projets qui souhaitent proposer des alternatives à des services en ligne occupant une place importante dans la vie quotidienne des usagers de l'Internet, et ce sur la base d'architectures décentralisées ou en P2P : les usages concernés touchent différentes opérations classiques de gestion et d'utilisation de contenus – recherche, stockage, partage. Les systèmes décentralisés en question se positionnent donc par rapport à des services fournis par des « grands acteurs » de l'Internet comme Google, Facebook, Dropbox, Picasa. En mettant à profit le potentiel du P2P et de la décentralisation, ils visent à répondre aux mêmes exigences que ces grands services du point de vue de l'utilisateur final (qui continuera par exemple à rechercher des mots, à constituer des réseaux d'amis ou à partager des photos), mais en étant construits sur une plateforme technologique différente. Un deuxième choix méthodologique a donc été de représenter, autant que possible, ces différents types de service à l'utilisateur final. Trois cas ont finalement été retenus, en croisant les critères d'enquête qui dérivent de ma problématique et la disponibilité pratique des terrains pour une exploration soutenue de ma part.

---

un système qui, sans contrôle central, copie des fichiers entre ordinateurs, et dans lequel chaque site participant au réseau sauvegarde une copie des messages postés dans les forums, et se synchronise périodiquement avec ses pairs. Les « labels » open source et P2P indiquent donc tous les deux, généralement, des technologies ou des communautés permettant aux gens de s'associer librement, de manière directe, et ils sont souvent parmi les incubateurs d'innovation les plus prometteurs. Ces similitudes, pourtant, ne devraient pas emmener à traiter le P2P et l'open source comme étant tout à fait similaires, ce que montrent plusieurs cas récents – comme Faroo et Wuala, deux des trois cas analysés dans cette thèse.

Faroo<sup>22</sup> est une start-up anglo-germanique proposant un moteur de recherche décentralisé à plusieurs niveaux de l'architecture technique, et visant à placer les affinités et préférences des usagers au centre des interrogations. Né en 2007, l'un des rares projets de moteur de recherche P2P à avoir dépassé le stade de *thought experiment* ou de travail de recherche universitaire, Faroo est piloté par une petite équipe interdisciplinaire comprenant trois ingénieurs, une psychologue et une linguiste, et relie une communauté de près de deux millions d'utilisateurs/pairs. Ses développeurs peuvent compter sur un « noyau dur » d'utilisateurs pionniers, particulièrement intéressés par le modèle décentralisé intégral que sous-tend le moteur et par les manières novatrices de traiter la recherche « sociale » et la gestion des données personnelles que cette architecture comporte. Les développeurs proposent par ailleurs, de façon intéressante, une version propriétaire de leur logiciel, ce qu'ils « justifient » par le fait que la spécificité de leur système réside dans la technologie qu'ils développent, plutôt que dans les coûts de ce développement, et la publication de leur logiciel comme de l'open source serait faire cadeau de cet avantage à la concurrence<sup>23</sup>. Cet ensemble de caractéristiques, uni à l'intérêt que l'équipe a manifesté pour ma démarche – précondition nécessaire de l'accès au terrain – rendait Faroo un cas d'étude idéal pour l'exploration des alternatives décentralisées à Google et Yahoo!.

Wuala<sup>24</sup> est une start-up née en Suisse, et rachetée dans un deuxième temps par un important fabricant français de matériel de stockage, proposant un système distribué de stockage de fichiers qui comporte également des fonctionnalités de réseautage social. Créée, elle aussi, en 2007, au sein d'un scénario foisonnant d'applications de stockage en ligne reposant sur des serveurs « classiques » ou des plates-formes en nuage, l'équipe de développeurs de Wuala – une quinzaine de personnes en 2010, tous ingénieurs – travaille depuis lors à une plateforme de stockage hybride, mêlant centralisation et décentralisation, et préconisant un système sophistiqué de « troc » de son espace disque en échange d'un espace de stockage permanent sur le réseau d'utilisateurs. Autre exemple de client P2P propriétaire, également doté d'une base d'usagers importante et engagée dans la discussion et le façonnage des aspects décentralisés du système, le logiciel Wuala m'a semblé un terrain propice à l'exploration du « compromis » entre une approche largement dépendante d'une plateforme de serveurs, et une approche d'architecture distribuée et décentralisée – pour servir un usage, le stockage en ligne, dont Dropbox est actuellement le représentant centralisé de premier plan.

Enfin, Tribler<sup>25</sup> est un projet de recherche universitaire lancé en 2007 à l'université

---

<sup>22</sup> <http://www.faroo.com>

<sup>23</sup> Voir le chapitre 4 et <http://www.faroo.com/hp/p2p/faq.html#opensource>: « *Open Source is perfect when competing by a cost advantage with a commercial product on the same technological level (Linux, OpenOffice). But it's not a good idea to hand over your technological advantage to a monopoly, when competing with its free service with enormous brand power* ».

<sup>24</sup> <http://www.wuala.com/>

<sup>25</sup> <http://www.tribler.org/>

technique de Delft, aux Pays Bas. Ses développeurs proposent un logiciel qui facilite la recherche et le partage de contenus privés entre utilisateurs, et qui sera assez rapidement doté d'une fonction de streaming vidéo. Inspiré du protocole de communication BitTorrent, Tribler est censé combiner son mécanisme sous-tendant avec des dynamiques communautaires et sociales. L'équipe Tribler est à présent le partenaire principal du projet P2P-Next<sup>26</sup>, financé par l'Union Européenne afin de développer un service de télévision *on-demand*. Les portes du laboratoire Tribler m'ont été ouvertes par la rencontre avec David Hales, membre de l'équipe en 2009 et 2010, un sociologue et informaticien intéressé par les manières dont les communautés P2P arrivent à se construire de manière durable. Au vu de son statut hybride entre projet universitaire, développement commercial et partenariat public/privé, et des mobilisations « politiques » novatrices du P2P auxquelles il a donné lieu, le projet Tribler était une occasion rare pour observer la problématisation politique et technique du P2P en tant que modèle alternatif pour les services audiovisuels par Internet, ainsi qu'à sa légitimation comme « valeur » dans le contexte européen.

### *LES MATÉRIAUX EMPIRIQUES*

L'enquête qualitative conduite dans cette thèse se base sur plusieurs types de matériaux empiriques. Il s'agit en premier lieu de documents mis à ma disposition par les équipes de développeurs ou par les firmes : documentation scientifique, notes de publication<sup>27</sup>, et, dans les cas de Faroo et Tribler, documents de travail intermédiaires destinés à un usage interne par l'équipe. Deuxièmement, des documents en ligne disponibles à une consultation publique, dont les sources ou les auteurs sont encore une fois les équipes et les firmes : bases de données, sites web, blogs, archives des listes de diffusion. Le cœur de mon travail d'analyse est cependant constitué par les entretiens avec les chercheurs et les entrepreneurs participants aux projets, mais aussi avec d'autres parties prenantes dans le développement du P2P (acteurs de l'*open source*, économistes). Les nombres d'entretiens (en personne en France, en Italie, en Suisse et aux Pays-Bas, ainsi que par téléphone et par e-mail) par cas d'étude sont les suivants : Wuala 8, Faroo 6, Tribler 9<sup>28</sup>, auxquels se somment 7 entretiens sur le thème des réseaux sociaux en P2P, 5 au sujet des moteurs de recherche en P2P, 5 sur le streaming vidéo en P2P, et 12 autres entretiens sur d'autres questions inhérentes, à différents titres, au développement d'applications P2P ou décentralisées. J'ai également effectué une semaine d'observation participante auprès du groupe Tribler, à l'université technique de Delft, au printemps 2010.

L'apport des utilisateurs pionniers des dispositifs P2P au développement et à la mise

---

<sup>26</sup> <http://www.p2p-next.org/>

<sup>27</sup> *Release notes*, documents distribués à l'occasion et à l'appui de la publication d'un logiciel – généralement quand celui-ci est encore dans sa phase de développement (du *pre-alpha* au *beta public*).

<sup>28</sup> À des fins narratives et/ou pour des raisons de confidentialité, les noms utilisés dans les chapitres 4, 5 et 6 sont fictifs.

en opération de ces systèmes a été étudié, lui aussi, de plusieurs façons. Si, de manière ponctuelle, j'ai pu bénéficier d'indications quantitatives recueillies par les firmes et par les développeurs, je me suis particulièrement concentrée sur les débats en ligne, à la fois entre utilisateurs, et entre usagers et développeurs.

Outre que sur les blogs et les forums utilisateur mis à disposition par les firmes, nombre de discussions intéressantes et éclairantes pour ma démarche ont eu lieu dans plusieurs autres arènes de discussion en ligne, non gérées directement par les développeurs, et spécialisées dans la recension et l'évaluation des nouveaux produits logiciels paraissant sur le marché. Ces plateformes sont depuis quelques années des points de référence certains pour le milieu des start-up (et des investisseurs) en informatique – et y être discuté est un signal de notoriété certaine (bien que, souvent, temporaire) pour son produit ou service. J'introduis ici quelques unes de ces plateformes, qui m'ont été très utiles pour l'ensemble des trois cas.

ReadWriteWeb<sup>29</sup>, fondé en 2003 par le néo-zélandais Richard MacManus, est un blog consacré à l'actualité des technologies Internet, Web en particulier. Il en couvre l'actualité et se distingue par l'accent mis non seulement sur le technique, mais aussi sur les usages et leur impact sur les médias et la société. Depuis octobre 2008, il comporte une édition francophone. Ars Technica<sup>30</sup>, fondé en 1998 par Ken Fisher, qui voulait en faire la publication de référence pour ceux qu'il appelait les *alpha geeks* – technologues et professionnels du secteur IT – est désormais un site de blogging réputé sur nombre de questions différentes concernant les technologies de l'information. Il garde un ethos de « technologie comme forme d'art », qui plaît à bien de lecteurs à la double casquette de programmeur et utilisateur. The Register<sup>31</sup>, bien reconnaissable à son ton caustique, est un magazine en ligne de nouvelles technologiques, fondé en 1994 en Grande-Bretagne. Ses sarcasmes ont pour cible principale les grands acteurs de l'industrie du logiciel.

S'ajoutent enfin à ces trois sites plusieurs autres plateformes, qui ne proposent généralement pas des articles complets, mais des recensions plus schématiques sur une variété de produits IT. En sont des exemples TheNextWeb, CNET, SmallNetBuilder, ReviewSaurus, et Beta<sup>32</sup>, site de recensions d'applications en version d'essai.

## PLAN DE LA THÈSE

La thèse comporte six chapitres après cette introduction.

---

<sup>29</sup> <http://www.readwriteweb.com/>

<sup>30</sup> <http://arstechnica.com/>

<sup>31</sup> <http://www.theregister.co.uk/>

<sup>32</sup> <http://thenextweb.com/>, <http://news.cnet.com/>, <http://www.smallnetbuilder.com/>, <http://www.reviewsaurus.com/> et <http://beta.seadvd.com> respectivement.

Le deuxième chapitre discute un ensemble de travaux qui, centrés sur les études de l'Internet comme objet technique, politique, économique et social, et sur les processus d'innovation qui le sous-tendent, proposent à l'attention du lecteur une pluralité de perspectives qui aideront, au fil des chapitres suivants, à « faire parler » l'enquête de terrain. Dans un premier temps, j'explicité les jalons conceptuels qui guident l'analyse des phénomènes étudiés dans cette thèse, ancrés dans trois domaines principaux : les *Internet studies*, la sociologie de l'innovation technique et la sociologie des usages des TIC. Ensuite, j'aborde un ensemble de travaux dont la lecture m'a été plus spécifiquement suggérée par ce que je découvrais sur le terrain sur les réseaux P2P comme systèmes techniques complexes et objets d'étude interdisciplinaire par excellence: il s'agit à la fois de travaux en sciences informatiques sur les réseaux distribués, et en sciences juridiques sur des questions de sécurité et de *privacy* de l'information « en réseau ». Enfin, je discute un ensemble de travaux au croisement de l'informatique, de l'économie et du droit, qui se réunissent autour d'un intérêt pour l'architecture du réseau Internet comme dispositif politique, juridique et économique.

Le troisième chapitre s'attache à explorer la « décentralisation » dans les dispositifs P2P analysés. Il propose d'abord une réélaboration de l'histoire du P2P, pour montrer comment, en cherchant les meilleures façons de contourner les problèmes spécifiques où l'architecture de l'Internet est durement mise à l'épreuve par les scénarios de « concentration » actuels, les développeurs se retournent vers l'Internet d'il y a quinze ou vingt ans: le premier Internet qui avait été créé en tant que système de communication entre machines de statut égal, partageant des ressources l'une avec l'autre. Comme le montrent l'histoire du premier Internet et l'histoire en train de se faire du « nouveau » P2P, une architecture P2P est le résultat d'un compromis compliqué entre décentralisation, présence d'intermédiaires, passage à l'échelle, organisation autonome des terminaux/utilisateurs, et résistance aux pannes. Un modèle technique distribué peut signifier un avantage économique pour l'entreprise à ses premiers pas, ou pour un utilisateur qui, en permettant à ses pairs de tirer profit de ses propres ressources matérielles, voit sa qualité de service améliorée; tandis que le serveur ou le point central du réseau se voit saisi, mis à l'épreuve, réapproprié, transformé.

Le quatrième chapitre explore les contraintes, principalement techniques mais aussi de marché, qui s'exercent sur un dispositif dont les concepteurs veulent pousser la décentralisation jusqu'au bout, pour introduire une version de P2P « intégral » dans la recherche d'information en ligne. Pour ce faire, il analyse plusieurs moments clés du développement de Faroo, start-up anglo-allemande née en 2005 et se proposant d'implémenter, sur la base d'une architecture technique en *peer-to-peer*, un moteur de recherche « réparti sur les utilisateurs » composé de nœuds-usagers qui, en lançant des interrogations de recherche, partagent en même temps des informations sur eux-mêmes, leurs préférences et leurs affinités avec les autres usagers. Le chapitre se concentre plus particulièrement sur trois aspects : la mise-en-opération des « six niveaux de distribution » qui constituent l'ossature du service ; l'implémentation socio-technique du compromis entre personnalisation et *privacy* ; le rôle de l'utilisateur et de sa machine dans le système, en termes de mise à disposition de ressources matérielles et de calcul, et de contribution à une « ressource informationnelle »

distribuée.

Le cinquième chapitre se tourne vers l'histoire de Wuala, une start-up suisse proposant un service de stockage en ligne communautaire et distribué. Il s'agit d'une application dont les créateurs ont travaillé, selon les moments, avec l'hybridation entre différents modèles d'architecture technique, avec la représentation de ces architectures auprès de leur public de clients potentiels, et avec la volonté de ces clients/usagers de s'engager dans un « jeu » de partage de ressources aux nombreuses implications techniques, juridiques et sociales. En suivant, après Faroo, une deuxième histoire de ce que la décentralisation de l'architecture technique change aux relations entre développeurs, entre utilisateurs, entre les deux, ce chapitre explore une approche hybride au P2P; il montre comment se construit le compromis entre une approche largement dépendante d'une plateforme de serveurs, et une approche d'architecture distribuée et décentralisée, où le fonctionnement du dispositif repose sur la mise à disposition de ressources informatiques de la part des usagers.

Le sixième chapitre montre comment les négociations qui ont parsemé le développement d'un prototype de logiciel P2PTV – Tribler, un projet universitaire de communauté sociale P2P créée en 2006 aux Pays-Bas – ont contribué à la problématisation politique et technique du peer-to-peer en tant que modèle alternatif pour les services audiovisuels par Internet, ainsi qu'à sa légitimation comme « valeur » dans le contexte européen. A cette fin, le chapitre suit les manières dont le dispositif est sorti de ses frontières académiques, pour saisir des institutions ou des espaces de collaboration entre recherche publique et secteur privé ; comment l'équipe de ses développeurs a par moments capitalisé sur son étiquette P2P, et a parfois dû s'en défendre, pour proposer une relecture juridique d'une technologie aux usages politiquement controversés ; comment ses créateurs ont proposé un modèle technique et économique à la fois différenciant et risqué, dans un contexte où le principal « client » est l'Union européenne avec ses puissants moyens financiers et sa machine bureaucratique et juridique très articulée.

Enfin, le septième et dernier chapitre détaille les conclusions de cette thèse, présentées autour de trois axes. Le premier axe a trait à la construction et la production du P2P comme alternative pour les services Internet; le deuxième axe suit le façonnage, la fabrication et la stabilisation de collectifs d'usagers et de formes de lien social sous-tendus par l'architecture P2P; le troisième axe discute le P2P appliqué aux services Internet comme objet du droit et producteur de droits.

## CHAPITRE 2

### INNOVATION ET INTERNET DISTRIBUÉ : UN REGARD INTERDISCIPLINAIRE

Un très récent ouvrage faisant le point sur vingt ans de sociologie des usages des technologies de l'information et de la communication (TIC) a judicieusement souligné comment une synthèse ou systématisation générale du champ des études de l'Internet, au vu de la prolifération et de la variété des travaux qui se revendiquent de cette appartenance, est devenue presque impossible (Denouël & Granjon, 2011). Dresser un cadre exhaustif de ces travaux serait encore plus difficile dans une thèse à l'allure interdisciplinaire comme veut l'être celle-ci : les démarches d'innovation et de développement que j'ai observées au prisme des *Internet studies*, de la sociologie de l'innovation technique et de la sociologie des usages des TIC m'ont amenée à m'intéresser à des littératures qui, tout en étant en principe centrées autour du même objet, restent pour la plupart assez disjointes les unes des autres si on raisonne, par exemple, en termes de références croisées. C'est le cas par exemple des apports des sciences informatiques sur les réseaux distribués, des productions des juristes sur des questions de *privacy*, surveillance et sécurité informationnelle, et des récents travaux sur l'architecture technique de l'Internet comme paradigme économique.

Mon ambition dans ce chapitre est bien plus modeste qu'une tentative de synthèse des études de l'Internet qui non seulement risque de rester stérile mais dont je sortirais, à n'en pas douter, perdante. Entreprendre une revue de littérature centrée sur les études de l'Internet comme objet technique, politique, économique et social, et les processus d'innovation qui le sous-tendent, devient une tentative de rassembler une pluralité de perspectives, de tisser et retrouver les liens qui aideront à « faire parler » le terrain au fil des chapitres qui suivent – tout en cherchant à échapper aux écueils que le grand travail bibliographique et généalogique nécessaire à la tâche risque de provoquer. C'est l'objectif des trois sections qui composent ce chapitre.

Dans un premier temps, je vais essayer d'explicitier les fondements du regard que j'ai porté sur le *peer-to-peer* « alternatif », en tant que processus d'innovation et activité technique. Mon but sera de proposer à l'attention du lecteur les jalons conceptuels

que sous-tendent l'analyse des phénomènes étudiés dans cette thèse, ancrés dans trois domaines principaux : les *Internet studies*, la sociologie de l'innovation technique et la sociologie des usages des TIC. L'intégration dans les études de l'Internet des *Science & Technology Studies (STS)*, et de l'attention particulière que ces approches prêtent à la contextualité et la pratique, est encore à ses débuts. Cependant, elle ouvre la voie à la découverte du travail souvent « invisible » de l'innovation sur l'Internet, et fournit les outils pour analyser ce travail en tant que partie intégrante de la manière dont se construit le « réseau des réseaux » en tant qu'artefact social, économique et politique.

Dans une deuxième partie, je discute une série de travaux qui, au risque de la simplification, je serais tentée de définir comme « orientés objet », dont la lecture m'a été plus spécifiquement suggérée par ce que je découvrais sur le terrain sur les réseaux P2P comme systèmes techniques complexes et objets d'étude interdisciplinaire par excellence. Il s'agit en premier lieu de travaux en sciences informatiques sur les réseaux distribués, parfois mobilisés par les acteurs eux-mêmes ; j'aborderai ensuite des travaux écrits par des juristes, qui visent à préciser et explorer les questions de sécurité et de *privacy* de l'information « en réseau ». Mon but sera d'essayer de montrer qu'établir des liens entre ces deux groupes de travaux peut contribuer à éclairer le dense substrat technique qui est nécessaire précondition à mes analyses, tout en montrant clairement que le P2P ne peut pas se résumer à ce seul substrat technique.

La dernière section sera l'occasion de rassembler et discuter un ensemble de travaux au croisement de l'informatique, de l'économie et du droit réunis autour d'un commun intérêt pour l'architecture du réseau Internet comme dispositif politique, juridique et économique. Il s'agit de travaux qui ont profondément marqué mon approche consistant à rechercher dans les « couches inférieures » d'un réseau certaines des raisons à la base de choix de développement, de réappropriations par les usagers, de formes d'organisation et de régulation. Ces travaux explorent comment les trajectoires d'innovation actuelles de l'Internet suggèrent que des formes spécifiques de distribution et décentralisation de l'architecture (ou leur manque) impactent procédures, pratiques et usages ; ils seront mobilisés au cours de cette thèse pour montrer les influences réciproques entre le développement de l'architecture et de l'organisation technique des services P2P alternatifs et la définition de leurs usagers, en tant que consommateurs, partageurs, producteurs et « gestionnaires » d'informations personnelles et de contenus numériques.

## SECTION 1. CONTEXTE ET PRATIQUE(S) DE L'INNOVATION SUR INTERNET

Cette section s'attache à expliciter les fondements du regard que je porte sur le *peer-to-peer* « alternatif » en tant que processus d'innovation et activité technique. Mon but sera de proposer à l'attention du lecteur les jalons conceptuels qui sous-tendent l'analyse des phénomènes étudiés dans cette thèse, ancrés dans trois domaines principaux : les *Internet studies*, la sociologie de l'innovation technique et la sociologie des usages des TIC. En particulier, je souhaite souligner comment le « *STS turn* » dans les *Internet studies*, avec son attention particulière prêtée à la contextualité et la pratique, ouvre la voie à la découverte et à l'explicitation du travail souvent invisible ou peu visible de l'innovation sur et avec Internet, et fournit les outils pour analyser ce travail en tant que partie intégrante de la manière dont se construit un Internet politique, économique, juridique, « social ».

Les auteurs et les notions discutés dans cette section, bien que mobilisant des paradigmes plus ou moins différents, se situent au croisement des intérêts de la sociologie de l'innovation pour les objets techniques et le rôle fondamental qu'ils tiennent dans l'édification du social, et de la sociologie des usages pour les pratiques sociales qui naissent autour de la manipulation et de l'appropriation d'un certain nombre de technologies émergentes. Cela dans le cadre d'un intérêt plus spécifique pour les TIC, qui prend de l'ampleur d'abord dans le cadre de la sociologie des médias, puis avec le développement académique des sciences de l'information et de la communication : comme l'a souligné Michel de Fornel, du fait qu'une TIC « ne suppose pas seulement l'action d'un individu mais l'action conjointe et continue de deux (ou plus) individus qui acceptent de coopérer pour créer et maintenir un espace interactionnel partagé », cette technologie « n'est donc pas un simple outil qui prolonge la perception en donnant un accès à un espace mais un artefact interactionnel » (de Fornel, 1994 : 126). Les travaux au carrefour de ces intérêts connaîtront un véritable boom avec l'arrivée de l'Internet, puis du Web, auxquels suivra l'établissement d'un domaine d'étude interdisciplinaire « centré objet », les *Internet studies*, à la fin des années 90.

### LES *INTERNET STUDIES* VERS UNE INTÉGRATION DES STS

Il serait difficile de nier que non seulement l'Internet mais aussi l'ensemble de recherches qui le prennent pour objet, regroupées sous l'étiquette *Internet studies*<sup>33</sup>, sont « *all about networks* » (Flyverbom, 2011 : 165). On pourrait aussi soutenir que tel est également le cas pour le champ des STS, en premier lieu la théorie de l'acteur-réseau (*Actor-Network Theory*, ANT) et ses évolutions successives : c'est ce que suggère John Monberg quand il souligne que « les deux champs sont organisés autour d'un réseau (*nexus*) sociotechnique, plutôt qu'autour d'une formation disciplinaire » (Monberg, 2005).

---

<sup>33</sup> Qui reste largement utilisée en français dans sa forme anglaise (Jouët, 2011 : 47).

Internet est une technologie, ou en ensemble de technologies, qui continue à provoquer à la fois fascination et inquiétude. A partir d'une tradition de recherche sociologique et anthropologique sur ses implications sociales (Shields, 1996; Hine, 2000; DiMaggio *et al.*, 2001; Woolgar, 2002), un champ d'études de l'Internet a pris forme pour explorer et comprendre la complexité du social, du juridique, des relations politiques et économiques à la base du développement de l'Internet. Barry Wellman affirme que le milieu des années 1990 correspondent à l'ère « préhistorique » de ce champ d'études, car les chercheurs se sont depuis détournés d'une analyse principalement documentaire et mettent désormais davantage l'accent sur l'approche critique des relations entre Internet et société (Wellman, 2004 : 127). L'attention des chercheurs a été tout particulièrement retenue par les manières dont l'Internet est géré et gouverné, les autorités formelles et informelles qui coordonnent et assurent son fonctionnement, les groupes sociaux qui le peuplent. Leurs perspectives sont informées par des différents ensembles d'hypothèses ontologiques, épistémologiques et méthodologiques se traduisant en différentes idées de ce qu'est l'Internet, qui le mobilise, qui en est affecté, comment il est représenté et « justifié », et comment on peut l'appréhender et connaître.

En dépit de leurs différences, les littératures comprises dans les *Internet studies* se proposent généralement de démontrer qu'une variété d'acteurs et d'enjeux s'entrecroisent autour de cette technologie, de sa gestion et gouvernance. Qu'il s'agisse d'études du rôle critique d'un ou plusieurs aspects du « réseau des réseaux » vis-à-vis de la politique internationale (Baird, 2002; Johnson et al., 2004; Mueller, 2002) ou des manières dont s'organisent les communautés d'utilisateurs de l'Internet (Hamelink, 2000; Klein, 2002), la plupart des écrits dans le champ des *Internet studies* ont adopté un regard sur ces dynamiques et espaces – les arènes de discussion des décideurs politiques d'un côté, le milieu « privé » des usagers comme des individus et groupes de l'autre – qui tend à les considérer comme deux domaines aux frontières assez rigides et clairement identifiables. Tarek Cheniti souligne comment cette « fracture spatiale » est souvent due à l'incapacité, de la part de ces auteurs, à conceptualiser la relation entre ces deux domaines en dehors des cadres théoriques et des vocabulaires de la mise en application du droit et de l'exercice de la citoyenneté. En conséquence, les chercheurs dans ce domaine tendent à s'inspirer des paradigmes théoriques et des outils conceptuels des sciences juridiques et politiques, et les approches qu'ils suggèrent se rapportent à une perception des décideurs politiques et juridiques comme acteurs soit utiles, soit « nuisibles » par rapport à l'Internet (Cheniti, 2009).

Si jusqu'à présent, relativement peu a été écrit en STS sur le thème spécifique des études d'Internet, il existe déjà une importante littérature qui interroge les idées préconçues sur la nature et les implications de l'organisation de la technologie en des domaines variés, en soulignant le caractère non accompli, incertain, et contingent des systèmes socio-techniques (Grint et Woolgar, 1997). Ainsi, lorsque les auteurs STS décrivent et évaluent de manière critique la façon dont les régimes de gouvernance sont fabriqués (Bucchi, 2006), ils mettent en doute les revendications de l'exhaustivité dans la théorie politique avec ses hiérarchies implicites, rationalités et intentionnalités. Les travaux dans le domaine STS peuvent donc donner des outils pour déconstruire et critiquer ; ils proposent une autre façon de penser les principes

sur lesquels repose la littérature des *Internet studies*. Par exemple, Irwin observe que la recherche STS peut être fructueuse pour décrire « la constitution même de l'ordre social et naturel dans des contextes spécifiques de gouvernance » (Irwin, 2008 : 586).

Très récemment, Mikkel Flyverbom a fourni un complément nécessaire et bienvenue à ce courant, en montrant de manière convaincante comment l'étude des politiques globales autour de systèmes et objets complexes – de l'environnement aux nanotechnologies, jusqu'à l'Internet, son cas d'étude – dispose d'un instrument méthodologique innovant dans la boîte à outils STS (Flyverbom, 2011). Les concepts mêmes qui ont interrogé et quelque peu ébranlé la sociologie et la science politique « classiques » au cours de ces dernières années – en premier lieu le concept à la fois célèbre et redouté de *multistakeholderism*, qui inspire le Forum sur la Gouvernance de l'Internet (Malcolm, 2008) – trouvent des équivalents dans les notions de « forums hybrides » (Callon, Lascoumes & Barthe, 2009) et d'« assemblages socio-techniques » (Caliskan & Callon, 2010).

En suivant Michel van Eeten, il faut également remarquer que l'attention aux manières dont les objets, les sujets et autres éléments se constituent, se positionnent, entrent en relation et s'associent représente une clé pour ouvrir l'étude de l'Internet, comme objet politique et global, aux « activités du monde réel » : en se concentrant exclusivement sur les entités et organisations engagées dans des discussions explicites sur la gouvernance mondiale de l'Internet, on ne parvient souvent pas à inclure dans le cadre des politiques globales de l'Internet un certain nombre de problèmes qui jouent un rôle crucial dans l'élaboration et la régulation des manières dont l'Internet fonctionne, tels que les accords d'interconnexion entre les fournisseurs de services Internet, les arrangements de routage et filtrage de contenus, ainsi que le contrôle de phénomènes comme le spam, le *phishing* et les *botnets* (van Eeten, 2009). Avec des ouvrages tels que *Protocol Politics* (DeNardis, 2009), l'utilité de la boîte à outils STS dans l'exploration des phases de développement technique et standardisation du « réseau des réseaux » comme arènes de discussion, réponse et mise en œuvre des grandes questions socio-politiques de l'Internet, devient de plus en plus évidente.

Dans cette acception, et en mettant donc en relief le besoin de prêter une attention particulière à la contextualité et la pratique, le « STS turn » dans les *Internet studies* peut ouvrir la voie à la découverte du « travail invisible » de l'innovation sur Internet, souvent peu ou sous-représenté, et fournit les outils pour analyser ce travail en tant que partie intégrante de la manière dont se construit le « politique » dans l'Internet. Il s'agit de mettre l'accent sur les pratiques qui créent la gestion et la gouvernance de l'Internet comme une réalité vivante, et en déterminent les manières d'opérer, travailler, résister, fonctionner. En parallèle, il s'agit aussi de ne pas assumer valeurs et rationalités déployés par les praticiens de l'Internet comme indicateurs des manières dont ceux-ci perçoivent le monde, mais de les problématiser comme ressources et catégories qu'ils déploient dans certaines circonstances pour créer et soutenir certaines configurations – organiser leur monde.

Les histoires racontées dans les chapitres qui composent le reste de cette thèse sont peuplées – on en discutera plus largement dans le chapitre suivant – de développeurs, d'ordinateurs et d'utilisateurs qui existent et opèrent « aux marges du réseau ». Elles ont lieu dans des environnements techniques et « économiques » transitoires et souvent fragiles, au milieu de fréquentes connexions et déconnexions, se reposant sur des protocoles de transport et systèmes d'exploitation souvent différents, parfois incompatibles, toujours complexes. Dans ce cadre, le courant de perspectives sur l'innovation technique qui fait des moments de stabilisation et déstabilisation d'un dispositif un outil méthodologique clé s'avère d'une utilité particulière pour comprendre la genèse et l'évolution des systèmes techniques qui nous intéressent ; on fait référence, tout particulièrement, à la sociologie de la traduction et aux évolutions qui ont suivi sa première apparition au début des années 1980.

La sociologie de la traduction a été présentée et élaborée au fil des années et à travers de nombreuses publications par les trois chercheurs qui sont généralement considérés comme les fondateurs de la théorie de l'acteur-réseau (ANT), Michel Callon, Bruno Latour et John Law (Callon, 1986, 1991, 2006 ; Latour, 1987, 2006 ; Law, 1992). L'ANT se propose de défier les catégories et distinctions *a priori* (ou *taken-for-granted*) et appelle à une analyse des manières pratiques au moyen desquelles les éléments « humains et non-humains » se relient les uns aux autres et se stabilisent (Latour, 2005). Les analyses informées par l'ANT cherchent à démontrer comment des phénomènes tels que la « société », les « organisations », le « politique » et le « technique » consistent en fait d'associations hybrides et instables, et stabilisations temporaires se basant sur des « traductions » (Callon, 1986 ; Latour, 1987). S'intéressant aux objets techniques en voie de constitution, à leur état instable et labile, le but de l'ANT est d'appréhender l'innovation en tant que processus plutôt qu'objet ou artefact ; il s'agit de capturer les manières dont des entités spécifiques se relient entre elles et se stabilisent (ou pas).

La prise en compte de cette perspective dans l'organisation du travail – par exemple d'une équipe de développeurs informatiques, comme ce sera souvent le cas au cours des chapitres suivants – conduit par exemple à éviter d'attribuer à cette entité un ensemble de caractéristiques entièrement définissables *a priori*, pour se concentrer sur le processus d'organisation et les manières dont celui-ci influence et s'entremêle avec le système développé ; en bref, pour observer comment une constellation d'éléments se rassemble et s'éloigne au fil du temps et au moyen de négociations constantes (Czarniawska, 2008). Le terme d'acteur-réseau qui donne son nom à l'ANT répond d'ailleurs à l'idée que, selon la perspective de l'observateur et le moment de l'observation, l'innovateur peut apparaître à la fois comme un (acteur) et comme un multiple (un réseau) : la dimension de collectif est mise en valeur à la fois pour les acteurs sociaux et pour les objets techniques, pouvant être traités comme des réseaux et sous-réseaux formés par une variété d'entités.

Une étude de comment un « modèle » de l'Internet se développe, s'organise, se déploie et se positionne par rapport à d'autres modèles possibles ou plus répandus,

commence donc par remplacer une sociologie « du social » par une sociologie « des associations » (Latour, 2005). Une approche relationnelle et orientée vers la pratique se met en place, en opposition à la fois au déterminisme techno-scientifique<sup>34</sup> et au socioconstructivisme<sup>35</sup> ; celle-ci avait été définie lors des premiers pas de l'ANT en tant que « co-évolution de l'objet et de sa société » (Akrich, Callon & Latour, 1988 : 19), et préconise un principe de symétrie généralisée selon lequel on doit pouvoir trouver des explications à la fois techniques et sociales « à *tout ce* que l'innovation concerne » (Callon & Latour, 1986 : 21).

Étant donné qu'on y décrit des innovations naissantes ou en train de faire leurs premiers pas, les analyses présentées dans cette thèse s'appuient tout particulièrement sur un autre point central du modèle de lecture de l'innovation proposé par l'ANT : le rôle de premier plan joué par l'innovateur lui-même, par sa capacité d'appréhender les réseaux à « géométrie variable » qui l'entourent, et de les mettre à profit.

Deux aspects sont tout spécifiquement mis en relief dans le travail de l'innovateur. D'une part, sa capacité d'arriver à connaître des différents facteurs et acteurs qui peuvent influencer d'une manière ou d'une autre le projet d'innovation, et d'en mener à bien une analyse exhaustive : ceux-ci peuvent varier des modèles économiques et techniques ayant sous-tendu des précédentes innovations, jusqu'à la nature et la variété des acteurs qui s'y intéressent, ou encore, comment ceux-ci sont organisés. La complexité socio-technique à laquelle l'innovateur se confronte dans sa démarche de conception l'amène à considérer un nombre toujours croissant d'objets et d'acteurs : l'entrée toujours plus détaillée dans le contenu technique de sa démarche entraîne souvent des consultations avec un nombre croissant et de plus en plus variés d'acteurs, et symétriquement, « plus il découvre sur le réseau socio-technique environnant l'objet, plus il doit l'analyser de différents points de vue » (Vinck, 2003 : 8). Il s'agit d'une véritable analyse socio-technique dont l'innovateur se charge : comme le souligne John Law, celui-ci se fait « ingénieur hétérogène », sociologue lui-même, en fonction de son habileté à comprendre et organiser à son compte acteurs et objets – ou pour reprendre la terminologie de l'ANT, « humains et non-humains » (Law, 1987).

D'autre part, le succès de l'innovation sera strictement lié à la capacité de l'innovateur d'« intéresser » et « enrôler » d'autres acteurs autour de sa vision de l'innovation et sa manière de l'articuler : il s'agit de former un réseau d'alliances susceptible d'acquiescer éventuellement une position dominante de soutien à l'innovation (Callon, 1986a). Au croisement de l'approche STS et de la théorie des mouvements sociaux, les travaux centrés sur les mouvements scientifiques et intellectuels (*scientific/intellectual movements*, Frickel, 2004 ; Frickel & Gross, 2005) et sur les mouvements « orientés technologie » et « orientés produit » (*technology- and product-oriented movements*, Hess, 2005) se sont plus particulièrement

---

<sup>34</sup> La perspective déterministe en sociologie des sciences et des techniques suppose d'un côté que le changement technique est un facteur indépendant de la société et autonome ; d'autre part, que le changement technique provoque un changement social.

<sup>35</sup> Posture préconisant la construction sociale des objets techniques, qui consiste à faire du technique un « produit » du social : par rapport au déterminisme technique, il s'agit d'une sorte d'« autre extrême », qui veut ramener tous les phénomènes techniques à des causes purement sociales.

concentrés sur les manières dont des mobilisations hybrides d'innovateurs, consommateurs, utilisateurs et entrepreneurs autour d'objets techniques et scientifiques complexes peuvent avoir un rôle déterminant dans la construction et la diffusion de formes alternatives de culture matérielle. En se concentrant, dans le cadre de l'Internet, sur le mouvement open source – et en le comparant avec des dynamiques semblables autour des énergies renouvelables et de la thérapie nutritionnelle – David Hess souligne notamment que l'innovation assume dans ces cas les connotations d'un « conflit centré sur l'objet », au cours duquel les processus d'incorporation et transformation modifient le design de la technologie ou du produit, la diversification entre les technologies et les produits augmente, et la variété de positions entre les acteurs-alliés s'affine tout en servant un objectif commun (Hess, 2005 : 516).

Si ces travaux peuvent parfois renvoyer à l'idée que le travail de l'innovateur en réseau consiste à « préparer » la société afin que l'innovation puisse y être reçue (Latzko-Toth, 2010), en modifiant l'environnement dans lequel le dispositif se situe aussi bien que le dispositif lui-même, d'autres auteurs, notamment la « deuxième vague » de l'approche constructiviste, ont contribué à mettre l'accent sur la dimension de « co-construction », à la fois des usagers et des technologies, que ce travail représente (Oudshoorn & Pinch, 2003).

Le travail de l'innovateur inclut donc l'inscription dans le dispositif technique des représentations qu'il se fait de ses usagers potentiels (Flichy, 1995 ; Monnoyer-Smith, 2008). La conception d'une innovation technique revient à configurer ses usagers (Woolgar, 1991), c'est-à-dire à intégrer dans le dispositif une définition représentée de l'identité des usagers ainsi que différentes contraintes susceptibles d'intervenir lors de l'utilisation (Millerand, 1998). Comme on le verra plus dans le détail par la suite, l'ouverture de ce modèle à la sociologie des usages étendra l'analyse à la place des utilisateurs comme acteurs à part entière dans le façonnage de l'objet technique, en explorant la capacité de ceux-ci, aussi bien que des innovateurs/premiers concepteurs, de construire à la fois le dispositif, ses usagers, et leurs représentations.

Dans le projet théorique de l'ANT figure également au tout premier plan une autre notion sur laquelle s'appuient mes analyses, celle d'épreuve. En effet, souligne Michel Callon, c'est lorsque des débats ou des controverses ont lieu autour d'un objet, d'un fait, d'une innovation scientifique ou technique, que ceux-ci sont visibles et transparents pour le sociologue, et ouvrent la voie à un travail d'exploration et d'éclaircissement pour les acteurs eux-mêmes ; ce sont ces moments qui permettent d'avoir une prise sur la genèse et l'évolution de l'innovation (Callon, 1986). Il s'agit donc de repérer et de mettre en relief les moments critiques dans le développement des dispositifs ; les points controversés qui éclairent les manières dont les acteurs et les objets sont mobilisés par rapport à ces dispositifs ; les occasions et situations où leur articulation est discutée et mise en question – en somme, les épreuves de la solidité du réseau.

Le rôle central des controverses comme outil conceptuel – au même titre que l'épreuve – pour penser le développement d'innovations avec l'ANT a non seulement contribué à la définition de celle-ci comme une « sociologie des controverses », mais a aussi profondément informé les développements de ses « descendants » ou parents

proches. En premier lieu, l'approche des mondes sociaux (Strauss, 1978 ; Clarke, 1991 ; Clarke & Star, 2008), parfois qualifié d' « écologique » (Star & Griesemer, 1989 ; Fujimura, 1992) en considération de sa préoccupation centrale consistant à rendre compte de la perspective de chaque acteur au sein d'un « monde » commun, un groupe d'acteurs partageant un même engagement dans une certaine activité, ainsi que les ressources pour mener à bien cette activité, et construisant en même temps une idée commune sur la façon dont elle doit être menée (Clarke, 1991 : 131).

Si, d'un côté, on retiendra du modèle écologique l'incitation à ne pas se renfermer dans une vision de l'activité de l'innovateur trop stratégique et stratégiste (voir de « campagne militaire » en quête d'alliés, comme souligne Joan Fujimura en 1992), ainsi que la conception de l'innovation comme activité partagée et multiple, il s'agit également de retenir ici une des contributions centrales de cette approche à l'étude de l'innovation, la notion d'objet-frontière<sup>36</sup>. Proposée pour la première fois par Susan Leigh Star et James Griesemer en 1989, cette notion désigne une entité servant d'interface aux perspectives de différents acteurs et mondes sociaux, suffisamment « plastique pour s'adapter aux besoins locaux et aux contraintes des divers groupes qui l'utilisent », capable d'exister en même temps dans plusieurs mondes sociaux, tout en satisfaisant aux « exigences informationnelles » de chacun et en étant suffisamment robuste pour maintenir une identité commune au cours de ces adaptations (Star et Griesemer, 1989). Au fil des chapitres suivants, on retrouvera des entités qui peuvent se qualifier d'objets-frontières sous plusieurs aspects. D'un côté, on verra comment ils agissent en tant que facilitateurs de négociations et de coopération au cours des débats entre innovateurs et usagers pionniers, voire entre les innovateurs eux-mêmes, en contribuant à éclairer et clarifier les reconfigurations et réappropriations de l'innovation (*ibid.*, 1989). D'un autre côté, ils contribuent à faire émerger la dimension de pratique de l'innovation au fil du temps et de l'espace, comme souligne Guillaume Latzko-Toth :

Si le concept d'objet-frontière contribue à éclairer les processus de traduction à l'œuvre dans l'activité scientifique et technique, il invite aussi à un renversement de perspective sur les fins et les moyens de l'innovation : les artefacts ne constituent pas (toujours) le but de l'activité technique, ils en sont aussi (et peut-être le plus souvent) les points d'appui, les charnières qui lui permettent de s'articuler et de se perpétuer *en tant que pratique* (Latzko-Toth, 2010).

En conclusion de cette section, le moment est venu de clarifier mon usage d'une notion que, sans m'y attarder, j'ai déjà introduite : celle de dispositif (socio-technique). Comme on verra et on aura l'occasion de préciser, la littérature en sciences informatiques se réfère au P2P en tant que réseaux et « systèmes » – terme qui revient fréquemment dans les mots des acteurs interrogés au cours de cette thèse, développeurs et ingénieurs en très grande majorité<sup>37</sup>. Par rapport à une vision

---

<sup>36</sup> À laquelle a récemment été dédié un numéro de la *Revue de l'anthropologie des connaissances* (Trompette et Vinck, eds. 2009).

<sup>37</sup> On retrouve également dans le vocabulaire des francophones parmi ces acteurs le mot « machine » pour désigner l'ordinateur, généralement un terminal domestique ; ce mot fait écho, mais ne se

relativement stabilisée, d'entité aux frontières définies et dotée de la cohérence et la solidité que le mot « système » suggère, la notion de dispositif contribue dans ce contexte à faire ressortir une dimension de mise-en-relation et d'articulation, et à concevoir ces systèmes<sup>38</sup> comme ensemble d'éléments reliés dans un réseau hétérogène, comprenant « des discours, des institutions, des aménagements architecturaux, des décisions réglementaires, des lois, des mesures administratives, des énoncés scientifiques, des propositions philosophiques, morales » (Foucault, 1994 [1977] : 299).

En vocabulaire ANT classique, le dispositif socio-technique<sup>39</sup> est défini comme un « assemblage d'actants humains et non humains où les compétences et les performances sont distribuées » (Akrich et Latour, 1992 : 259) que le travail d'innovation consiste à faire exister. Il convient aussi de noter l'accent mis par des contributions récentes sur la « puissance d'agir » (Proulx, 2009) des objets, que cette notion permet d'inclure dans l'analyse pour une appréciation plus fine de la dimension collective de l'agir : « les dispositifs font des choses. Ils articulent des actions; ils agissent ou en font agir d'autres » (Muniesa, Millo & Callon, 2007). Plusieurs auteurs ont d'ailleurs souligné l'aptitude de la notion de dispositif, dans ses mobilisations plus récentes par les chercheurs en sciences sociales – tout particulièrement par ceux qui s'intéressent aux dispositifs d'information et de communication – à élargir une focale auparavant rivée sur les rapports de pouvoir et le contrôle social, les « connotations normatives et disciplinaires perçues chez Foucault » (Beuscart & Peerbaye, 2006 : 6-7), pour sous-tendre « une réflexion qui prend sens au sein de différentes formes de tensions et de médiations qui s'articulent et interagissent dans des dispositifs médiatiques, numériques ou communicationnels en constante évolution : télévision, internet, téléphone mobile, outils collaboratifs » (Appel, Boulanger et Massou, 2010).

#### DES USAGERS REPRÉSENTÉS AUX USAGERS PIONNIERS ET CO-CONCEPTEURS : L' « OUVERTURE » DE L'INNOVATION EN RÉSEAU

La troisième et dernière partie de cette section introduit brièvement les débats existants en littérature sur le statut de l'utilisateur dans le processus d'innovation, et sur la relation entre les usages et l'évolution des dispositifs techniques d'information et de communication. Encore une fois, loin de prétendre à l'exhaustivité, il s'agit ici de fournir quelques jalons conceptuels sous-tendant les parties de nos analyses qui se concentrent sur l'intervention des utilisateurs dans le processus d'innovation. On a déjà vu comment cette intervention peut s'explicitier dans les phases de création et

---

superpose pas complètement à l'usage qu'en fait Nicolas Dodier pour désigner des entités compactes et capables d'un fonctionnement autonome (Dodier, 1995 : 14).

<sup>38</sup> Ainsi que leurs « couches inférieures », substrats, infrastructures (voir Bowker et Star, 1999 ; Star, 1999 et la section 3 de ce chapitre).

<sup>39</sup> La notion co-existe avec celle de « réseau socio-technique », ou encore d'« agencement socio-technique ».

développement par des processus de représentation – la mise-en-dispositif de l'idée que le développeur, ou l'ingénieur, se fait de ses usagers (Flichy, 1995), ce qui se prolonge souvent par l'implication de certains d'entre eux dans des mécanismes d'essai, d'expérimentation et de *test* pouvant être conduits en laboratoire ou, ce qui arrive fréquemment dans le cas des services Internet, directement « en réseau » au moyen de versions 'beta' d'essai (Bardini, 1996). Il s'agit ici de dire quelques mots sur la manière dont s'est opéré un déplacement dans l'analyse des utilisateurs comme acteurs à part entière dans le façonnage de l'objet technique : un important ensemble de travaux explore désormais la capacité des utilisateurs, aussi bien que des innovateurs/premiers concepteurs, de construire à la fois le dispositif, ses usages, et leurs représentations.

Deux questions sont au cœur de la sociologie des usages des TIC, une tradition de recherche sur la « matérialité du social » qui, selon Julie Denouël et Fabien Granjon, est en partie une spécificité française<sup>40</sup> et n'a jamais été au centre des préoccupations de la raison sociologique jusqu'à ce que la question de l'appropriation des TIC dans une société devenant « technologique dans son ensemble » ne tende *de facto* à s'imposer comme une problématique centrale (Denouël & Granjon, 2011 : 9). D'une part, l'objectivation du social au moyen des objets techniques destinés à l'échange d'information et à la communication ; d'autre part, la médiation sociotechnique, le principe selon lequel l'étude de l'édification du social ne peut faire abstraction des éléments sociotechniques – tout comme la description des objets techniques ne peut faire sans la mise en actes de « savoirs proprement sociaux ».

Josiane Jouët observe que l'intérêt de ce courant de recherches, marqué dans ses premiers temps par un bricolage théorique important autour de références et traditions variées, est de « proposer des modèles d'interprétation des usages qui croisent la spécificité des processus de communication avec différentes dimensions de l'innovation technique et de l'innovation sociale » (Jouët, 2000 : 513-514). Le socle commun à ces modèles d'interprétation est le fait qu'ils naissent comme critiques d'une tradition de pensée longtemps prévalente autour des usages, et s'y opposent parfois entièrement. Cette perspective est centrée sur trois points : une nette séparation entre la figure du concepteur/innovateur et les usagers ; la séquentialité des moments qui composent le processus d'innovation, qui voit l'objet d'abord créé et

---

<sup>40</sup> Fabien Granjon rappelle que « Le courant de recherche que l'on qualifie, sans doute par facilité, de "sociologie des usages" prend ses racines au début des années 80, notamment au sein du Centre national d'étude des télécommunications (Cnet, devenu Orange Labs), un pôle de recherche alors public, au sein duquel Patrice Flichy va prendre la direction d'un laboratoire pluridisciplinaire baptisé UST pour "Usages sociaux des télécommunications". C'est dans cet environnement d'ingénieurs dédié à la production d'innovations technologiques, que vont être lancées des études d'usages sur la téléphonie ou sur les premiers réseaux télématiques (le Minitel), avec des contributions de chercheurs comme Josiane Jouët, Chantal de Gournay, Yves Toussaint, etc. Par ailleurs, Patrice Flichy et Paul Beaud, cofondent en 1983 la revue *Réseaux*, laquelle est encore aujourd'hui la principale tribune académique des études d'usages » et Julie Denouël renchérit : « La sociologie des usages des TIC est en partie une originalité française, liée à la politique de R&D de l'opérateur historique France Télécom. Il crée des outils et des services avec évidemment un intérêt certain pour en analyser les usages. » Entretien avec Hubert Guillaud, « Julie Denouël et Fabien Granjon : Les usages en question », *Internet Actu*, 15 juin 2011, <http://www.internetactu.net/2011/06/15/julie-denouel-et-fabien-granjon-les-usages-en-question/>

modifié par ses créateurs, puis reçu par les utilisateurs ; enfin, l'idée que, lors de la présentation aux utilisateurs<sup>41</sup> d'une nouvelle technologie, la variété des possibilités d'usage de cet objet se voit considérablement limitée par les balises inscrites dans l'objet lui-même.

Si la conception dichotomique entre concepteur et utilisateur et des « marges de manœuvre réduites » de ce dernier a longtemps perduré, cela tient probablement à la « matérialité rigide » et la stabilisation des TIC « classiques » (du magnétoscope aux premiers ordinateurs domestiques) avant leur découverte par les utilisateurs, et donc, à l'opacité de ces objets qui rendait la réappropriation des prescriptions d'usage un phénomène marginal (Latzko-Toth, 2010 : 13). Les travaux de Michel de Certeau consacrés aux « formes d'inventivité des pratiques ordinaires, aux 'arts de faire' et aux tactiques qui fondent le processus d'appropriation et permettent aux individus de se constituer 'un propre' » (Jouët, 2011 : 52), en bref, aux activités créatrice de l'usager (de Certeau, 1980), ouvriront la voie aux premiers travaux sur les modalités d'appropriation de l'innovation technique, conjuguées à des pratiques communicationnelles témoignant de l'autonomisation des individus et de l'émergence de nouvelles formes de social (Jouët, *ibid.*)

De son côté, le modèle de l'acteur-réseau avait mis en relief depuis le début les limites d'une conception linéaire du processus d'innovation, pour souligner les opérations de traduction plurielles, les aller-et-retours, mises en œuvre par les concepteurs au cours du processus d'innovation, strictement liées aux diverses représentations que ceux-ci se font de leurs usagers potentiels (Callon & Latour, 1986). Cependant, comme on a vu, le centre du modèle restait dans un premier temps l'innovateur et concepteur, les usagers étant une composante parmi d'autres du réseau d'acteurs qu'il s'agit d'intéresser et engager dans le projet d'innovation. C'est en tout premier lieu à Madeleine Akrich qu'on doit les travaux d'extension du modèle ANT à la prise en compte des usages comme partie intégrante du processus d'innovation (Akrich, 1992, 1993, 1998). Dans cet ensemble de travaux perdure par ailleurs une séparation bien définie des rôles du concepteur et de l'usager, tout comme la linéarité de leurs interventions : ils se concentrent en effet sur l'intervention des usagers sur des dispositifs « déjà constitués » (Akrich, 1998), leur appropriation, modification, extension par des utilisateurs actifs mais qui n'entrent pas en relation directe avec les concepteurs – en bref, qui restent en dehors du cycle des aller-et-retours de la conception.

Le caractère très intégré des utilisateurs dès les premières phases du processus d'innovation et la forte continuité entre concepteurs et usagers constituent probablement, d'ailleurs, une des spécificités des TIC par rapport à d'autres innovations. Dès les premières années du vingtième siècle, la radio et le

---

<sup>41</sup> De manière peu surprenante, souvent définis dans ce contexte en tant qu'utilisateurs « finaux » ; notion qui a une signification et des conséquences juridiques précises en informatique (se déclarer utilisateur final d'un bien signifie accepter le processus ayant conduit à la création de ce bien et accepter les conséquences liées à l'existence et la décadence du bien en question ; dans le cas d'un bien et donc d'un logiciel, au sens de la propriété intellectuelle, l'utilisateur final accepte toutes les clauses d'une éventuelle licence), et qui veut ici renforcer l'idée que les utilisateurs disposent d'une marge de manœuvre réduite dans leur appropriation de la technique, par rapport à la liberté de manœuvre qu'ont eu les concepteurs dans la phase de création.

phonogramme, dont le programme d'action a été très largement modifié par rapport aux intentions de leurs concepteurs, en ont été des exemples (Méadel, 1994). Notamment dans le cas des médias « en réseau » – dont on a pu observer depuis leurs débuts une plus grande malléabilité et plasticité, ainsi que l'immédiateté d'appropriation commençant souvent dès les premiers pas de la conception (Abbate, 1999) – des travaux successifs se sont attachés à démontrer la nature « pervasive » de l'intervention des utilisateurs dans le processus d'innovation, en explorant les moments de ce processus où les frontières entre l'utilisateur et le concepteur perdent en clarté, ou la modification de ces rôles au fil du temps, ou encore leur cohabitation chez un même acteur (Oudshoorn & Pinch, 2003; Boczkowski et Lievrouw, 2008 ; Lindsay, 2003). Valérie Beaudouin et Christian Licoppe soulignent une autre caractéristique propre aux usages des TIC en réseau, et la portée de ses implications :

« l'importance toute particulière sur internet des ajustements entre production et réception. Ainsi, a-t-on pu montrer que les moteurs de recherche finissaient par avoir une identité propre par le biais du profil de leurs utilisateurs et des requêtes qui leur étaient adressées, que les sites personnels avaient des tonalités spécifiques en fonction de leur lieu d'hébergement ou encore que les secteurs marchands transformaient la mise en scène de leurs espaces de vente au fil des interactions qui se nouaient avec leurs clients et leurs visiteurs. Ces ajustements réciproques entre la production et la réception permettent l'émergence de territoires marqués et une organisation sociale du réseau. » (Beaudouin & Licoppe, 2002).

Certaines contributions invitent, par ailleurs, à ne pas écraser sous cette problématisation de la « créativité de l'utilisateur » les différences existantes entre les producteurs et les utilisateurs d'une innovation, tout comme on reconnaît les différences entre les innovateurs : cela équivaldrait à contourner une analyse critique des contraintes sociales et culturelles qui influencent pourtant les rapports usager-technologie (Oudshoorn & Pinch, 2003).

Cette perspective permet d'appréhender nombre d'innovations ayant contribué au paysage des TIC tel qu'on le connaît aujourd'hui – du WiFi communautaire aux pratiques de *blogging*, jusqu'au P2P – en observant les besoins, non perçus au cours des phases de conception et re-conception, qui entraînent des détournements de fonctionnalités par les évolutions que les usagers apportent aux services et techniques proposés. Le « pouvoir » de l'utilisateur sur le produit se construit et peut être observé à travers son interprétation des technologies à sa disposition, leur réinvention et sa participation à leur conception. Les innovations répondant à ce modèle dit de « l'innovation ascendante », dont Eric von Hippel est le principal porteur depuis plusieurs années (von Hippel 1986, 2001, 2005), partagent la caractéristique de « partir » des usagers et de se diffuser sur le réseau en organisant la coopération entre eux : cela a été le cas des étudiants américains bricoleurs et passionnés de musique qui sont à l'origine de la conception des premiers systèmes P2P (Cardon, 2006). Se dessinent dans ce modèle des profils d'utilisateur qu'on n'aura pas de difficulté à reconnaître dans les pages qui suivent, du *lead user* – l'utilisateur expert qui n'hésite pas à bricoler des dispositifs naissants pour en tirer un plus grand bénéfice (von Hippel, 1986) – au *self-manufacturer*, l'utilisateur/innovateur qui réagence les outils dont il

dispose pour les adapter à des besoins spécifiques (von Hippel, 2005). Dans la même lignée, en décrivant le processus d'innovation dans la phase d'amorçage ou *bootstrapping* – phase que, comme on verra, est dans les systèmes P2P un moment d'intéressement et enrôlement des usagers vital pour la survie même du dispositif – Thierry Bardini utilise la notion de co-évolution pour rendre compte de la manière dont les dispositifs et les usagers/concepteurs évoluent en même temps, s'adaptant mutuellement à travers d'interactions constantes (Bardini, 2000).

Pour une analyse de systèmes qui non seulement sont en train de faire leurs premiers pas, mais opèrent dans un environnement structurellement fragile et incertain comme les « marges du réseau », il convient enfin de noter les apports de la sociologie des usages à la conceptualisation de l'« ouverture » du dispositif (ouverture à l'initiative et l'activité d'une pluralité d'acteurs) en tant que dynamique structurante du dispositif lui-même, en particulier pour ce qui est des technologies de réseau. Le modèle d'innovation « *Permanently Beta* », que Gina Neff et David Stark ont développé à partir de l'observation des communautés de logiciel libre, est à cet égard particulièrement intéressant. En suivant comment les usagers se font de la place (et comment les développeurs leur font de la place) dans la mise au point du logiciel, une version *beta* (d'essai) après l'autre, les auteurs en arrivent à remettre en question la nécessité d'une stabilisation du dispositif technique (Neff & Stark, 2003). Pour en venir à une conclusion qui résonne avec les histoires de P2P racontées dans cette thèse : l'instabilité, le changement, le « perpétuellement provisoire » pourrait bien être la seule caractéristique durable – et structurante – des dispositifs en réseau et de leur développement.

## SECTION 2. LE PAIR-À-PAIR, UN MODÈLE DE RÉSEAU INFORMATIQUE AUX USAGES MULTIPLES

Si la discussion précédente autour de la littérature sur l'innovation a pu de temps en temps ramener l'attention du lecteur à l'objet plus spécifique de cette thèse, l'application d'un modèle décentralisé ou P2P aux services Internet, le moment est venu dans cette deuxième partie d'en faire le centre de mes propos. Cette section introduit une série de travaux centrés sur le P2P comme modèle de réseau informatique sous-jacent à des systèmes complexes. Il s'agit en premier lieu de travaux en sciences informatiques sur les réseaux distribués, parfois mobilisés par les acteurs eux-mêmes ; j'aborderai ensuite des travaux à vocation juridique qui visent à préciser et explorer les questions de sécurité et de *privacy* de l'information « en réseau »<sup>42</sup>. En essayant de montrer comment ces travaux ouvrent la voie à un traitement du P2P « alternatif » comme objet d'étude interdisciplinaire par excellence, mon but est de contribuer à éclairer le dense substrat technique qui est la nécessaire précondition de mes analyses, tout en montrant clairement que le P2P ne peut pas se résumer à ce seul substrat technique.

Je vais par ailleurs laisser délibérément de côté – en cohérence avec ma démarche qui souhaite agrandir la focale rivée sur le P2P comme outil de partage de fichiers – le vaste débat existant en littérature sur les batailles économiques et juridiques qui, à commencer par les « Napster wars » (Marshall, 2002), ont pris pour objet les pratiques et logiciels de partage de fichiers couverts par le droit d'auteur. Il en sera de même, sauf exceptions, pour le vaste ensemble de recherches, débats, controverses, chiffres et rapports décortiquant, dans une perspective d'économie des échanges P2P<sup>43</sup>, la « défaite » de l'industrie des contenus numériques face à la décennie de partage non autorisé de fichiers musicaux à laquelle le P2P aurait ouvert la porte, ainsi que les fondements des arguments qui établissent des relations de causalité immédiate entre le partage en ligne et les chutes dans les ventes de disques. Enfin, en accord avec une autre facette de ma démarche, je ne m'attarderai pas sur un ensemble de travaux que l'on pourrait qualifier de « sociologie des usages décentralisés en ligne ». Les auteurs de ces travaux s'intéressent aux nouvelles formes de coopération en ligne (Aguillon & Cardon, 2007), aux réseaux sociaux numériques (Boyd, 2004 ; Cardon, 2008), ou encore aux communautés en ligne (Gensollen, 2009 ; Auray, 2012) dans des contextes fort intéressants et novateurs, mais qui – hormis certains cas spécifiques qui intègrent

---

<sup>42</sup> Je trace un portrait schématique de la co-évolution des « générations » de réseaux P2P, des poursuites judiciaires dont ils ont fait l'objet, et des technologies de surveillance qui ont essayé de les contrer dans [Musiani, 2011].

<sup>43</sup> Maya Bacache-Beauvallet et Pauline Meinzel en font une excellente synthèse dans un document de travail présenté en janvier 2012 à un séminaire du projet ADAM (Bacache-Beauvallet & Meinzel, 2012). Elles y concluent notamment que, parmi les études centrées sur le partage de fichiers qui dominent actuellement le secteur, commencent à se frayer un chemin des éléments d'analyse tels que la réputation, la confiance et les types de relations établis entre les pairs. Cela permet (et surtout, permettra) d'appréhender le P2P en économie en faisant le lien avec des notions de grand intérêt pour la problématique de cette thèse, comme l'asymétrie de l'information et l'incertitude de la présence des ressources.

une réflexion explicite sur les formes de la solidarité technique sous-jacente au P2P (Beuscart, 2002 ; Dauphin & Dagiral, 2005) – ne se concentrent pas, contrairement à ce qui est mon intention, sur le lien entre les façons dont les applications prennent forme, notamment les évolutions de leur architecture technique, et leurs possibles influences sur la circulation des contenus, les droits ou les relations sociales.

Depuis dix ans, le P2P est devenu l'un des termes les plus largement discutés dans le domaine des technologies de l'information et de la communication (Schoder & Fischbach, 2003 ; Shirky et al., 2001). Il repose sur l'hypothèse que, dans un réseau d'égaux ou de pairs, à l'aide de systèmes de communication et d'échange appropriés, deux ou plusieurs individus sont en mesure de collaborer spontanément, sans nécessairement avoir besoin de coordination centrale (Schoder & Fischbach, 2003 ; Schollmeier, 2001). Après avoir présenté les principes caractéristiques des réseaux P2P, je proposerai un cadrage de leurs usages, ainsi qu'ils sont exposés dans nombre de travaux en sciences informatiques, à l'aide d'un modèle à trois niveaux ou « couches » (*layers*) basé sur deux contributions récentes (Oram, 2001 ; Schoder, Fischbach & Schmitt, 2005). Je souhaite mettre ainsi en lumière les dynamiques et caractéristiques, spécifiques à ces réseaux par rapport au modèle client/serveur, sur lesquelles travaillent les développeurs interviewés au cours de cette thèse. Sont en jeu la possibilité d'un passage à l'échelle plus souple, la coordination auto-organisée et décentralisée de ressources précédemment sous-utilisées ou limitées, la tolérance aux pannes/redondance, le renforcement de réseaux *ad-hoc* ; ainsi que les possibilités d'organisation et d'inclusion des utilisateurs qui peuvent plus difficilement être mis en œuvre en utilisant les approches client/serveur « traditionnels ».

#### PARTAGE, DÉCENTRALISATION, AUTONOMIE

Le partage des ressources et des services distribué(e)s, la décentralisation et l'autonomie sont généralement reconnus dans la littérature comme les trois principes caractéristiques des réseaux P2P.

En premier lieu, dans un réseau P2P chaque nœud peut fournir à la fois les fonctionnalités de client et de serveur, c'est-à-dire qu'il peut agir à la fois comme fournisseur et comme consommateur de services ou de ressources (comme peuvent être informations, photos, bande passante, capacité de stockage et capacité de calcul). Parfois, ces nœuds du réseau sont dénommés *servents*, mot composé dérivé des termes client et serveur. Deuxièmement, il n'y a pas d'autorité centrale de coordination pour l'organisation du réseau pendant sa phase de configuration ; une autorité n'est pas non plus présente dans l'utilisation successive des ressources et de la communication entre les pairs du réseau. Cela concerne, en particulier, le fait qu'aucun nœud n'exerce un contrôle central sur l'autre. À cet égard, la communication entre pairs se fait directement.

Une distinction est souvent faite entre les réseaux P2P purs et ceux qui sont dits hybrides ; on verra plus tard comment ces derniers en viennent à être la seule

alternative viable pour l'utilisation d'un modèle P2P dans la pratique<sup>44</sup>. En raison du fait que tous les composants du réseau ont des droits égaux et des fonctions équivalentes, les réseaux P2P « purs » représentent l'idéal-type pour la conception d'un système P2P. Dans ces structures, il n'existe aucune entité qui a une vision globale du réseau (Barkai, 2001 ; Yang & Garcia-Molina, 2001). Dans les réseaux P2P hybrides, les fonctions sélectionnées, telles que l'indexation ou l'authentification, sont attribuées à un sous-ensemble de nœuds qui, par conséquent, assume le rôle d'une entité de coordination. Ce type d'architecture de réseau combine les principes du P2P et du client/serveur (Minar, 2001 & 2002). En troisième lieu, chaque nœud dans un réseau P2P peut déterminer de façon autonome quand et dans quelle mesure il met ses ressources à la disposition d'autres entités.

Sur la base de ces trois principes ou caractéristiques, le P2P, loin d'être un développement récent qui aurait commencé avec Napster, peut être considéré comme une des plus anciennes architectures dans le monde des télécommunications (Oram, 2001). En ce sens, Usenet, avec ses groupes de discussion, et l'Internet à ses débuts, ou ARPANET, peuvent être classés en tant que réseaux P2P. En conséquence, il y a des auteurs qui soutiennent que le P2P est en train de ramener l'Internet à ses origines, à l'époque où chaque ordinateur avait des droits égaux dans le réseau (Minar & Hedlund, 2001). La baisse des coûts d'un côté, et la disponibilité croissante de l'autre de la capacité de calcul (ou « cycle processeur »), de la bande passante, de la capacité de stockage, le tout accompagné par la croissance de l'Internet, ont créé de nouveaux champs d'application pour les réseaux P2P. Dans un passé récent, ceci a entraîné une augmentation spectaculaire du nombre d'applications P2P, et des discussions controversées concernant les limites et les performances, ainsi que les implications économiques, sociales et juridiques de ces applications (Schoder, Fischbach & Teichmann, 2002 ; Smith, Clippinger & Konsynski, 2003).

#### QUESTIONS DE « COUCHES », QUESTIONS DE TERMINOLOGIE

De quoi parle-t-on quand on parle de P2P, ou plus précisément, de quel P2P parle-t-on selon les cas ? En m'appuyant sur deux contributions (Oram, 2001 ; Schoder, Fischbach & Schmitt, 2005) récentes, je propose ici un modèle à trois niveaux ou « couches », se composant de infrastructures P2P, applications P2P, et communautés/collectifs P2P (Tableau 1), qui peut aider à éclairer la terminologie à l'œuvre dans la littérature en sciences informatiques dans la théorie et la pratique du P2P.

Le premier niveau concerne les infrastructures P2P. Celles-ci sont positionnées au-dessus des réseaux de télécommunications existants, et agissent comme une base pour tous les niveaux. Les infrastructures P2P assurent les fonctions de communication, intégration et traduction entre les composants du réseau. Ils fournissent à chacun d'entre eux des services qui permettent de localiser et de communiquer avec leurs

---

<sup>44</sup> Le chapitre 3 explore tout particulièrement ce point.

pairs dans le réseau et d'identifier, utiliser et échanger des ressources, ainsi que de lancer des processus de sécurité comme l'authentification et l'autorisation. Le deuxième niveau concerne les applications P2P, qui utilisent les services des infrastructures P2P. Elles visent à permettre la communication et la collaboration des entités en absence de contrôle/supervision central/e. Le troisième niveau se concentre sur les phénomènes d'interaction sociale, en particulier, la formation de communautés/collectifs et leurs dynamiques internes. Plusieurs auteurs notent comment le terme *peer* (pair) se réfère pour les premiers deux niveaux essentiellement à des entités techniques, tandis que pour le troisième niveau il est interprété dans un sens non technique : le pair est alors une personne (voir notamment Schoder, Fischbach & Schmitt, 2005).

<p><b>Communautés P2P</b></p> <p>Communautés virtuelles qui se servent de plateformes d'interaction pour favoriser des activités de communication et collaboration entre personnes</p>
<p><b>Applications P2P</b></p> <p>Programmes qui utilisent les services des infrastructures P2P, visant à permettre la communication et la collaboration des entités en absence de contrôle central</p>
<p><b>Infrastructures P2P</b></p> <p>Mécanismes et techniques qui assurent les fonctions de communication, intégration et traduction entre les composants informatiques (les applications en particulier)</p>

Tableau 1. Modèle à trois niveaux précisant la terminologie relative au P2P. D'après [Schoder, Fischbach & Schmitt, 2005] et [Oram, 2001].

### *INFRASTRUCTURES P2P*

Le terme d'« infrastructures P2P » se réfère aux mécanismes et techniques qui assurent les fonctions de communication, intégration et traduction entre les composants informatiques en général, et les applications, en particulier. La fonction principale est la fourniture de l'interopérabilité, dans le but d'établir une infrastructure P2P intégrée et puissante. Cette infrastructure agit comme une plateforme de service P2P, avec des interfaces de programmation et des logiciels d'intermédiation standardisés, qui, en principe, peut être utilisée par n'importe quelle application (Schoder & Fischbach, 2003 ; Shirky et al., 2001 ; Smith et al., 2003).

Par rapport aux services que l'infrastructure P2P rend disponible pour les applications, la sécurité est devenue particulièrement importante (Barkai, 2001). Les améliorations

de la sécurité sont actuellement considérées comme le principal défi que doit être résolu pour que les réseaux P2P deviennent intéressants pour des usages « monétisables ». L'utilisation partagée des ressources a souvent lieu entre pairs qui ne se connaissent pas et, par conséquent, ne se font pas nécessairement mutuellement confiance. Dans de nombreux cas, l'utilisation d'applications P2P implique de donner à des tiers l'accès aux ressources d'un système interne, par exemple, afin de partager des fichiers ou des cycles processeur. L'ouverture d'un système d'information afin d'accorder l'accès à des tiers, ou de leur permettre de communiquer, peut avoir des effets secondaires critiques, qui impliquent fréquemment que des mécanismes de sécurité classiques, tels que les logiciels pare-feu, soient contournés. Un autre exemple est la communication via le logiciel de messagerie instantanée. Dans ce cas, la communication a souvent lieu sans l'utilisation de techniques de chiffrement ou d'encryptage, ce qui met en péril la confidentialité et la sécurité des échanges. Les principaux défis concernant les infrastructures P2P résident donc à cette heure dans le développement de techniques et méthodes d'authentification, autorisation, vérification de la disponibilité, intégrité et confidentialité des données (Bursell, 2005).

### *APPLICATIONS P2P*

Dans la littérature des premières années 2000, les applications P2P sont souvent classées selon les catégories de messagerie instantanée, partage de fichiers, grilles informatiques et collaboration (Schoder et Fischbach, 2003 ; Shirky et al., 2001). Cette forme de classification a nécessité des modifications au fil du temps, car elle ne parvient plus à établir des distinctions claires – en particulier parce que aujourd'hui, dans de nombreux cas, ces catégories ont fusionné et sont intégrées. Les applications P2P sont par ailleurs plus clairement distinguées selon l'aspect des ressources qui y sont distribuées, communiquées ou partagées ; aspect qui mieux se prête à apporter une compréhension des principes de base des réseaux P2P et de la façon dont ils fonctionnent (Taylor & Harrison, 2009). Les pages suivantes donnent donc une vue d'ensemble des approches possibles à la coordination de différents types de ressources dans les réseaux P2P. Ces ressources peuvent être de l'information, des fichiers, de la bande passante, des capacités de stockage ou de calcul.

#### *Partage de fichiers, la technique à la base d'un usage controversé*

Si le modèle de téléchargement de contenus numériques en streaming à partir de sites Web s'est imposé dans les années les plus récentes comme un concurrent sérieux du partage de fichiers en P2P<sup>45</sup>, celui-ci demeure probablement l'application la plus répandue de ce modèle. Jusqu'à 70% du trafic réseau sur l'Internet a été attribué à l'échange de fichiers dans un passé récent, notamment de fichiers musicaux : dans

---

<sup>45</sup> La mesure, la quantification et la modélisation de cette évolution sont par ailleurs problématiques, du fait que de plus en plus de réseaux P2P « cachent » intentionnellement leur trafic (Karagiannis et al., 2004 ; Musiani, 2011).

l'immédiat après-Napster, on estimait qu'il y avait plus d'un milliard de téléchargements de fichiers musicaux effectués chaque semaine (Stump, 2002). Il est caractéristique du partage de fichiers que les pairs qui ont téléchargé les fichiers dans le rôle de client les rendent par la suite accessibles à d'autres pairs, adoptant le rôle de serveur. Un problème central pour les réseaux P2P en général, et pour le partage de fichiers en particulier, est la localisation des ressources (*lookup problem* : Balakrishnan et al., 2003). Dans le contexte des systèmes de partage de fichiers, trois algorithmes différents ont été mis au point: le modèle *flooded request*, le modèle de dossier centralisé, et le modèle de routage (*document routing* : Milojevic et al., 2002). La meilleure façon d'illustrer ces trois modèles est d'utiliser en tant qu'exemples les applications qui ont constitué leurs principales implémentations au fil du temps : Gnutella, Napster, et Freenet.

Les réseaux P2P qui sont basés sur le protocole Gnutella fonctionnent sans autorité centrale de coordination. Tous les pairs ont des droits égaux au sein du réseau. Les demandes de recherche d'un fichier particulier sont acheminées par le réseau selon le modèle *flooded request*, ce qui signifie qu'une requête est transmise à un nombre prédéterminé de pairs. Si ceux-ci ne peuvent pas répondre à la demande, ils la transmettent à d'autres nœuds jusqu'à ce que un niveau de « profondeur » de recherche établi a priori ait été atteint, ou que le fichier demandé ait été localisé. Les résultats de recherche positifs sont ensuite envoyés à l'entité qui demande la ressource, qui peut alors télécharger le fichier souhaité directement à partir de l'entité qui est en train de l'offrir. Le protocole FastTrack, fonctionnant sur ce principe, est à la base d'usages très répandus : ce protocole optimise les requêtes par le biais d'une combinaison de super-nœuds centraux, qui forment un réseau décentralisé similaire à Gnutella.

À cause de son modèle sous-jacent, qui inclut un répertoire centralisé, le début de Napster (1999, 2000) peut être considéré comme un exemple presque parfait d'un système P2P hybride, dans lequel une partie de la fonctionnalité de l'infrastructure, dans ce cas le service d'indexation, est fournie de manière centralisée par une entité de coordination. Au moment même où un pair se connectait au réseau Napster, les fichiers dont il disposait étaient enregistrés par le serveur de l'application. Lorsqu'une requête était établie, le serveur de Napster fournissait une liste de pairs rendant le fichier disponible au téléchargement. L'utilisateur pouvait alors obtenir les fichiers directement par les pairs qui les offraient.

La recherche et le stockage de fichiers au sein du réseau Freenet (Clarke, 2003 ; Clarke et al., 2002) a lieu *via* le modèle de routage (Milojevic et al., 2002). Une différence significative par rapport aux modèles qui ont été introduits jusqu'à présent est que les fichiers ne sont pas stockés sur le disque dur des pairs qui les fournissent, mais sont intentionnellement conservés à d'autres endroits dans le réseau. Cela s'explique par le fait que Freenet a été développé dans le but de créer un réseau dans lequel des informations peuvent être stockées et accessibles de manière anonyme. Entre autres choses, cela exige que le propriétaire d'un nœud de réseau ne sache pas quels documents sont stockés sur son disque dur local. Pour cette raison, les fichiers et les pairs se voient attribués par le système des numéros d'identification uniques. Quand un fichier est créé, il est transmis, *via* les pairs voisins, au pair doté du numéro

d'identification le plus proche numériquement du numéro d'identification du fichier, et y est stocké. Les pairs qui participent à la transmission du fichier sauvegardent le numéro d'identification du fichier et notent aussi le pair voisin auquel ils l'ont transféré, dans une « table de routage » à utiliser pour des requêtes ultérieures.

La recherche de fichiers a lieu en transmettant les interrogations de recherche sur la base des informations contenues dans les tables de routage des pairs individuels. Différemment des réseaux de recherche qui fonctionnent selon le modèle *flooded request*, quand un fichier demandé est localisé dans le système, il est renvoyé au pair qui en fait la demande par le même chemin. Dans certaines applications, chaque nœud sur cette route stocke aussi une copie/réplique du fichier afin d'être en mesure de traiter plus rapidement les requêtes de recherche à venir. Dans ce processus, les pairs stockent des fichiers seulement jusqu'à une capacité maximale. Lorsque leur stockage est épuisé, les fichiers sont supprimés selon le principe du « moins récemment utilisé ». Il en résulte une présence progressivement et proportionnellement plus élevée de copies des fichiers plus populaires dans le réseau, alors que, au fil du temps, les fichiers qui sont demandés moins souvent sont supprimés (Milojicic et al., 2002).

Plusieurs études ont reconnu le modèle de routage comme adapté pour être utilisé dans des larges communautés. Le processus de recherche, cependant, est plus complexe que, par exemple, dans le modèle *flooded request*. En outre, il peut déterminer la formation d' « îles » – c'est-à-dire, un cloisonnement du réseau, dans lequel les différentes communautés ne sont plus connectées à l'ensemble du réseau (Clarke et al., 2002 ; Langley, 2001).

### *Echange et utilisation partagée d'information*

Les réseaux P2P sont aussi déployés pour l'échange et l'utilisation partagée d'informations de présence (*presence information*), l'indicateur d'état qui exprime la capacité et la volonté de communiquer par les partenaires potentiels de ladite communication, ou encore pour la gestion documentaire, et pour la collaboration.

Les informations de présence jouent un rôle très important dans les applications P2P. Elles sont décisives dans l'auto-organisation des réseaux P2P, car elles fournissent des informations sur les pairs et les ressources disponibles dans le réseau. Elles permettent aux pairs d'établir un contact direct avec les autres pairs et d'en savoir davantage sur les ressources qu'ils rendent disponibles. Un exemple largement diffusé des applications P2P qui utilise les informations de présence est constitué par les systèmes de messagerie instantanée. Ces systèmes offrent aux pairs la possibilité de transmettre des informations *via* le réseau, par exemple, d'indiquer s'ils sont disponibles pour communiquer (Hummel, 2002). L'utilisation d'informations de présence est intéressante pour l'utilisation partagée de la capacité de calcul (les applications peuvent reconnaître indépendamment quels pairs sont à leur disposition dans une grille informatique, et déterminer comment les tâches de calcul peuvent être réparties entre les cycles processeur non utilisés des pairs respectifs), et dans les scénarios liés à la disponibilité de l'information dans l'informatique ubiquitaire (les périphériques mobiles peuvent indépendamment reconnaître les pairs qui sont disponibles dans son

environnement, par exemple pour demander des services Web, de l'information, du stockage ou de la capacité de calcul).

Les systèmes de gestion documentaire (Document Management Systems, DMS) sont généralement organisés de manière centralisée, et permettent le stockage, la gestion et l'utilisation des données de manière partagée. Toutefois, il est possible d'accéder seulement aux données qui ont été placés dans le dépôt central du DMS. En conséquence, un effort supplémentaire est nécessaire pour créer un index centralisé des documents pertinents. On a noté qu'une grande partie des documents créés dans une entreprise sont répartis entre les ordinateurs de bureau, sans qu'un référentiel central ait en fait connaissance de leur existence (Crowcroft & Pratt, 2002). Dans ce cas, l'utilisation des réseaux P2P peut venir en aide. Par exemple, il est possible de façonner des réseaux qui créent un dépôt connecté à partir des données locales des différents pairs. L'indexation et la catégorisation des données sont effectuées par chaque pair, sur la base de critères choisis individuellement.

En plus de relier les sources de données distribuées, les applications P2P peuvent offrir des services pour l'agrégation de l'information et la formation de réseaux de connaissances P2P auto-organisés. OpenCola, un logiciel libre de gestion et acquisition d'information<sup>46</sup>, a été l'une des premières applications P2P à offrir à ses utilisateurs la possibilité de recueillir des informations distribuées à partir des domaines de la connaissance qui les intéressent (Schoder, Fischbach, et Schmitt, 2005 ; Leuf, 2002). À cet effet, les utilisateurs créaient des dossiers sur leur bureau, auxquels ils assignaient des mots-clés correspondant à leur domaine d'intérêt. OpenCola recherchait ensuite le réseau de connaissances de façon indépendante et continue parmi les pairs disponibles qui affichaient les mêmes domaines d'intérêt, sans que ces recherches soient dépendantes d'une gestion centralisée de l'information. Les documents provenant de pairs pertinents étaient analysés, suggérés à l'utilisateur le cas échéant, et automatiquement dupliqués dans le dossier de l'utilisateur. Si l'utilisateur refusait les suggestions, les critères de recherche se réajustaient en conséquence. L'utilisation d'OpenCola déterminait un réseautage spontané des utilisateurs ayant des intérêts similaires, sans nécessité d'un contrôle central.

Enfin, les logiciels de collaboration (*groupware*) de type P2P permettent la gestion des documents au niveau de groupes de travail fermés. En conséquence, les membres d'une équipe de travail peuvent communiquer de manière synchrone, gérer en commun des réunions en ligne, et d'éditer des documents partagés, de manière synchrone ou asynchrone. Dans le *groupware* de type client/serveur, une zone de travail pour la gestion des données centrales doit être mise en place et administrée sur le serveur pour chaque groupe de travail. Afin d'éviter cette tâche de gestion supplémentaire, les réseaux P2P peuvent être utilisés pour le travail collaboratif. Ces systèmes, dont le plus célèbre est Groove, offrent des fonctions similaires (messagerie instantanée, partage de fichiers, notification, co-navigation, conférences, et bases de données dotées de synchronisation en temps réel) à ceux des produits client/serveur largement utilisés, mais ne nécessitent pas de gestion centrale des données. Toutes les

---

<sup>46</sup> Le système a été achetée par la firme canadienne Open Text Corporation en 2003, mais son site est disponible en archive sur <http://web.archive.org/web/20020721214044/http://opencola.com/>

données créées sont stockées sur chaque pair et sont synchronisées automatiquement. Si les pairs ne peuvent pas se rejoindre directement, il y a l'option de synchronisation asynchrone via un serveur de relais. Les logiciels de collaboration en P2P donnent aux utilisateurs la possibilité de mettre en place des espaces communs qui fournissent un environnement de travail partagé pour les équipes virtuelles formées sur une base *ad hoc*, et permet aussi d'inviter autres utilisateurs à travailler dans ces équipes (Edwards, 2002).

### *Utilisation efficace de la bande passante*

Puisque la demande de capacité de transmission des réseaux est en augmentation constante, notamment en raison de l'augmentation en quantité et qualité des données multimédia de grand volume, une utilisation efficace de la bande passante à disposition d'un système dans son ensemble est de plus en plus importante.

Actuellement, la plupart des services Internet sont construits autour d'une approche centralisée : les fichiers sont conservés sur le serveur d'un fournisseur d'informations et transférés de là au client qui en fait requête<sup>47</sup>. Dans ce cas, un problème se pose lorsque des augmentations spontanées de la demande exercent une influence négative sur la disponibilité des fichiers, parce que des « embouteillages » et des files d'attente se développent. Les approches de type P2P sont susceptibles de gérer de façon plus équilibrée les augmentations soudaines de charges sur le réseau, en tirant parti des voies de transmission qui ne sont pas pleinement exploitées : ce qui s'explique tout particulièrement dans leur rôle de « facilitateurs » de l'utilisation partagée de la bande passante fournie par les fournisseurs d'information (Taylor & Harrison, 2009).

Les réseaux P2P hybrides, notamment, peuvent gérer de façon plus équilibrée les charges sur le réseau par rapport aux architectures client/serveur pures. Seules les demandes initiales pour les fichiers doivent être servies par un serveur central ; les requêtes ultérieures peuvent être automatiquement transmises à des pairs au sein du réseau, qui ont déjà reçu et reproduit ces fichiers. Ce concept est le plus souvent appliqué dans les domaines de streaming de contenus numériques, comme est le cas de Tribler (Pouwelse et al., 2006)<sup>48</sup>. Dans ce type d'outils, les utilisateurs peuvent s'abonner à des chaînes d'information ou à des fournisseurs de logiciel, par qui ils souhaitent obtenir des informations, ou des mises à jour de logiciels. Lorsque de nouvelles informations sont disponibles, les fournisseurs d'information respectifs les transmettent aux pairs qui y ont souscrit. Après avoir reçu l'information, chaque pair agit instantanément en tant que fournisseur et transmet l'information à d'autres pairs. Les domaines d'application dans lesquels ces dessins ou modèles peuvent être mis en œuvre sont la distribution de « didacticiels » (logiciels pour *eLearning*) dans un intranet, la distribution d'antivirus et mises à jour de configuration de pare-feu, et

---

<sup>47</sup> La montée en puissance du modèle client-serveur dans le cadre de l'Internet « commercial » est un des moments fondamentaux de la « mise à l'épreuve » du modèle P2P dans l'histoire de l'Internet, et sera au centre du prochain chapitre.

<sup>48</sup> Ce logiciel constitue le cas d'étude exposé dans le sixième chapitre de la thèse.

aussi la mise à jour de jeux vidéos « massivement multi-joueurs » de manière distribuée sur les terminaux des joueurs/pairs (Kulkarni, 2009).

L'utilisation de modèles P2P peut aussi amener à des véritables utilisations partagées de la bande passante, accélérant ainsi le téléchargement et le transport de gros fichiers qui sont demandés simultanément par des entités différentes. Généralement, ces fichiers sont divisés en blocs plus petits. Chaque bloc est ensuite téléchargé par les pairs qui le demandent. Dans le premier cas chaque pair ne reçoit qu'une partie de l'ensemble du fichier. Par la suite, les parties du fichier sont échangées par les pairs sans avoir besoin de nouvelles requêtes à la source originale. Finalement, les pairs reconstruisent l'ensemble des parties pour former une copie exacte du fichier original. La mise en œuvre la plus célèbre de ce principe est celle trouvée dans BitTorrent (Cohen, 2003).

### *Gestion distribuée d'un espace de stockage*

Les solutions centralisées pour le stockage de données peuvent avoir certains inconvénients, tels que l'utilisation inefficace de la capacité de stockage, charge supplémentaire sur le réseau, ou la nécessité de solutions de sauvegarde ultérieures. Toutefois, l'augmentation de la connectivité et la disponibilité accrue de la bande passante permettent des formes alternatives de gestion du stockage qui répondent à certains de ces problèmes. Avec les réseaux de stockage P2P, il est fait l'hypothèse que seule une portion de l'espace disque disponible sur chaque ordinateur qui rejoint le réseau est utilisée. Un réseau de stockage P2P est un ensemble d'ordinateurs, formé sur la base des réseaux existants, qui partagent tout le stockage disponible sur le réseau. Différents exemples de systèmes actifs aujourd'hui, parmi eux Wuala, présentent (ou présentaient<sup>49</sup>) des similitudes de base dans la façon dont ils sont construits et organisés (Hasan *et al.*, 2005). Afin de participer à un réseau de stockage P2P, chaque pair reçoit une information d'identification dans le réseau, comportant une « clé » publique et une clé privée. Avec l'aide d'une fonction de hachage (fonction qui, à partir d'une donnée fournie en entrée, calcule une *empreinte* servant à identifier rapidement, bien qu'incomplètement, la donnée initiale), la clé publique est utilisée pour créer un numéro d'identification univoque pour chaque pair. Afin d'accéder au stockage sur un autre ordinateur, le pair doit soit mettre à disposition une partie de son propre espace de stockage, soit payer une cotisation. En échange de sa contribution, chaque pair se voit attribuer un volume maximum de données qu'il peut ajouter au réseau.

Quand un fichier doit être stocké dans le réseau, il est identifié avec un numéro d'identification unique, créé avec une fonction de hachage à partir du nom ou le contenu du fichier respectifs, ainsi que la clé publique du propriétaire. Le stockage du fichier et sa recherche dans le réseau se déroulent suivant le modèle de routage de document. En outre, un certain nombre de répliques des fichiers, déterminé librement, est également stocké. Chaque pair récupère la version actualisée de la table de routage

---

<sup>49</sup> Voir le chapitre 5 ; le développement des aspects distribués du modèle de stockage Wuala a été arrêté à la fin de 2011, mais il sera repris dans le détail dans mon analyse.

qui est utilisé pour le stockage et la recherche. Les pairs vérifient la disponibilité de leurs voisins à intervalles réguliers, afin d'établir quels pairs ont quitté le réseau et éventuellement chercher une compensation des « sorties » par de nouvelles « entrées » de pairs disponibles à stocker. Les nouveaux pairs qui rejoignent le réseau sont inclus dans la table de routage de la même façon.

Pour coordonner les réseaux de stockage en P2P, les clés doivent être produites et distribuées aux pairs respectifs, et l'utilisation de l'espace de stockage doit être supervisée et coordonnée. En règle générale, ces tâches que l'on pourrait qualifier d'« administratives » sont traitées par un certain nombre de pairs particulièrement performants, ou qui se distinguent également par un degré élevé de disponibilité dans le réseau : les « super-pairs ». Afin de veiller à ce qu'un manque de disponibilité de la part d'un de ces pairs choisis ne porte pas atteinte à l'efficacité fonctionnelle de l'ensemble du réseau, les pairs sont coordonnés par des algorithmes résistants aux défaillances arbitraires, appelés « Byzantins » (Castro, 2001). Les demandes sont traitées par tous les pairs sélectionnés qui sont aussi disponibles. Chacun envoie un résultat à l'entité qui a émis la demande, laquelle attend jusqu'à ce qu'un certain nombre de résultats identiques aient été reçus de ces pairs avant d'accepter le résultat comme correct.

Par le biais de la réplique de fichiers et la distribution aléatoire de numéros d'identification aux pairs, le réseau de stockage P2P vérifie automatiquement que les différentes copies du même fichier sont stockées à différents endroits géographiques. Un certain niveau de consensus parmi les spécialistes semble exister sur le fait qu'aucune administration supplémentaire, ou autre solution de sauvegarde, n'est nécessaire pour assurer une protection contre un incident local ou une perte de données (Foster, 2002 ; Hasan *et al.*, 2005 ; Kubiawicz *et al.*, 2005). Cette procédure permet également de réduire l'importance d'un problème qui est caractéristique des réseaux P2P : le manque de certitude qu'un pair spécifique sera disponible dans le réseau à un moment donné dans le temps. Dans le cas des réseaux de stockage en P2P, cela pourrait entraîner des contextes où aucun des pairs stockant le fichier demandé n'est disponible dans le réseau. Augmenter le nombre de répliques stockées à différents emplacements géographiques peut, toutefois, accroître la probabilité qu'au moins un de ces pairs sera disponibles dans le réseau et en assurer la redondance (Williams *et al.*, 2007). Les faibles coûts de gestion, résultant du caractère auto-organisé des réseaux de stockage P2P, et le fait que d'autres solutions de sauvegarde sont rarement nécessaires sont aussi considérés comme des avantages de ces systèmes (Hasan *et al.*, 2005).

### *Répartition de la capacité de calcul*

La prise de conscience du fait que la puissance de calcul disponible des entités formant un réseau est souvent non utilisée a été une des premières motivations pour l'utilisation des applications P2P afin de réunir une plus grande de calcul, notamment à la suite de la montée en puissance des ordinateurs domestiques et l'amélioration de la qualité de leurs connexions Internet (Oram, 2001). En parallèle, la nécessité de calculs à hautes prestations dans des secteurs tels que l'informatique biomédicale, la

logistique ou le secteur financier, a sensiblement augmenté. En utilisant les applications P2P pour assembler la puissance des cycles processeurs qui resteraient autrement inutilisés, il est possible d'atteindre une puissance de calcul que même les supercalculateurs les plus coûteux ne peuvent guère fournir. Ceci est réalisé en créant un groupe d'ordinateurs indépendants mais en réseau, dans lequel tous les nœuds de ce réseau sont combinés en un seul ordinateur logique. Les approches de la diffusion coordonnée et l'utilisation partagée des ressources de calcul dans des organisations virtuelles et dynamiques se situent actuellement dans le cadre du calcul par grilles informatiques (*grid computing* : Baker *et al.*, 2002 ; Foster & Kesselman, 2004). Le terme « grille informatique » naît d'une analogie avec les réseaux électriques : la plus grande quantité possible de ressources, la puissance de calcul notamment, devrait être disponible pour l'utilisateur, idéalement sans restriction et sans être liée à un endroit géographique particulier – de façon similaire à la manière dont l'énergie électrique est tirée d'une prise de courant (Bal, Lohr & Reinefeld, 2002).

L'un des projets les plus cités dans le contexte du P2P, qui, cependant, n'est qu'une première approximation de l'objectif de l'informatique en grille, est SETI@home (*Search for Extra Terrestrial Intelligence at Home* : Anderson, 2001 ; Anderson *et al.*, 2002). SETI@home est un projet scientifique lancé par l'Université de Californie à Berkeley, dans le but de découvrir des signaux radio émis par des intelligences extraterrestres. A cet effet, un radiotélescope à Puerto Rico enregistre une portion du spectre électromagnétique de l'espace. Ces données sont envoyées au serveur central de SETI@home en Californie ; celui-ci, en capitalisant sur le fait que la plus grande partie des cycles processeur sur les ordinateurs des privés et des entreprises reste inutilisée et plutôt que de faire analyser les données par un seul supercalculateur coûteux, divise les données en unités plus petites et envoie ces unités à plusieurs millions d'ordinateurs mis à disposition par les bénévoles qui se sont inscrits pour participer à ce projet. Le client SETI effectue les calculs au cours des cycles processeur inactifs des ordinateurs des participants, puis renvoie les résultats au serveur central.

Dans la littérature dédiée, SETI@home est parfois considéré comme un des premiers exemples d'une application P2P, et plus précisément, l'exemple par excellence de grille informatique (Oram, 2001; Miller, 2001). Cette évaluation, cependant, est considérée comme partiellement inexacte par d'autres auteurs, qui objectent que le cœur de SETI@home demeure en fait un classique modèle client/serveur, en raison du fait qu'un serveur central coordonne les tâches des nœuds et leur envoie des paquets de travail. Les pairs traitent les tâches qui leur ont été assignées et renvoient les résultats. Dans ce système, il n'y a pas de communication entre les nœuds individuels ; il s'agit donc d'un exemple de calcul distribué et récentralisé, plutôt que décentralisé, et peut d'ailleurs contribuer à clarifier la différence entre distribution et décentralisation, et entre les types de système auxquels ces deux termes renvoient. Cela n'empêche pas de reconnaître les caractéristiques P2P que le projet comporte au delà de l'absence de communication directe (Milojicic *et al.*, 2002). Les nœuds forment une communauté virtuelle et rendent des ressources disponibles, sous forme de cycles processeur inutilisés (*idle*). Les pairs sont dans une large mesure autonomes, car ils déterminent si et quand le logiciel SETI@home est autorisé à effectuer des tâches de calcul relatives au programme (Anderson, 2001). La réalisation commune de

ces types de tâches de calcul distribué, cependant, n'est possible que si les étapes analytiques peuvent être séparées et divisées en paquets de données.

La vision des grilles informatiques s'est étendue bien au-delà des projets comme SETI@home. À un stade avancé de développement, il devrait être possible pour chaque nœud du réseau non seulement de mettre à disposition ses propres ressources, mais également de tirer parti des ressources disponibles dans l'ensemble du réseau P2P<sup>50</sup> (Foster, 2002).

### *COMMUNAUTÉS P2P*

Le terme « communauté virtuelle » a été introduit en premier dans les sciences informatiques par les pionniers de l'Internet J. C. R. Licklider et Robert Taylor. Dans un article de 1968, qui a fait école, les deux chercheurs américains font la prédiction que « dans la plupart des domaines, [les communautés virtuelles] se composeront de membres séparés géographiquement, parfois regroupés en petites grappes et parfois travaillant individuellement. Elles seront communautés non pas parce que elles se trouveront dans un endroit commun, mais parce que elles auront des intérêts communs. » (Licklider & Taylor, 1968). Aujourd'hui, de nombreuses variantes et extensions de la définition d'origine peuvent être trouvées dans la littérature en sciences informatiques, auxquelles se sont ajoutées les contributions des sciences économiques et sociales ; il semble pourtant exister un minimum de consensus entre elles pour identifier dans les intérêts communs, les normes et valeurs communs, ainsi qu'une plateforme d'interaction commune, les éléments qui constituent une communauté virtuelle. Des éléments ultérieurs qui contribuent à qualifier les communautés virtuelles sont les liens affectifs, la continuité temporelle, la consolidation des relations, et l'autodétermination. Sur cette base, les développeurs définissent généralement, et traitent sur le plan opérationnel, les communautés P2P comme des communautés virtuelles qui se servent de plates-formes de communication et d'interaction pour favoriser des activités de communication, collaboration et coordination de groupes de travail ou de personnes.

Un certain nombre d'intérêts communs et d'infrastructures communes sont identifiés dans la littérature par rapport au contexte des communautés P2P. Les utilisateurs de réseaux de partage de fichiers, par exemple, souhaitent échanger de la musique et donnent vie à une plate-forme d'interaction destinée à cet usage. Dans les réseaux du type [...] @home, les utilisateurs sont liés par des intérêts: par exemple, par la recherche de formes de vie extraterrestre (SETI@home) ou par la recherche d'un remède contre le sida (FightAIDS@home). On observe souvent que la mise en place de normes et de valeurs communes augmente en fonction de la disponibilité de

---

<sup>50</sup> Les premières implémentations utilisables dans l'industrie sont déjà annoncées par les grands acteurs sur le marché, par exemple le Globus Project (<http://www.globus.org/>, commencé en 2004), qui travaille sur un intergiciel standardisé pour le calcul en grille et a été accueilli favorablement dans la communauté d'experts. Le projet est soutenu par la Global Alliance, qui inclut des acteurs importants tels que IBM, Microsoft et Sun.

sanctions et de mécanismes de réputation (Akrich & Méadel, 2012). Les premiers clients Napster, par exemple, ont permis aux utilisateurs de refuser l'accès à leurs propres ressources à des personnes sélectionnées, qui n'avaient pas de ressources à offrir (ou seulement des ressources non souhaitées), afin que les « free-riders » puissent être exclus de la communauté (Shirky, 2001).

Il est particulièrement intéressant de voir comment la littérature en sciences informatiques esquisse, en « partenariat timide » avec l'économie ou la psychologie, une analyse de questions telles que l'attachement émotionnel, le sentiment d'appartenance, ou l'adhésion continue à une communauté existante, et s'attache à comprendre dans quelle mesure elles contribuent ou peuvent contribuer au fonctionnement du dispositif. Un dénominateur commun semble être l'auto-détermination – caractéristique inscrite dès le début dans l'un des principes des réseaux P2P, l'autonomie des nœuds ; mais d'autres aspects sont susceptibles d'influencer le développement, la motivation, la stabilisation et le contrôle des communautés P2P, comme la confiance et la réputation.

#### *Création de confiance, établissement d'une réputation*

Dans les structures de collaboration virtuelle, la confiance est une condition essentielle, car les pairs doivent souvent ouvrir leur système interne à l'accès par d'autres pairs. En conséquence, les systèmes P2P efficaces sont aussi ceux qui permettent que la confiance s'établisse entre les partenaires de la communication ou de l'échange. La réputation peut donc se révéler un élément très important pour la création de confiance dans les réseaux P2P (Lethin, 2001).

La réputation est généralement établie à l'intérieur du système au moyen de données agrégées sur le comportement antérieur d'un acteur. Dans les réseaux P2P, il est possible de profiter d'un large éventail de mécanismes de réputation. Des exemples de ces mécanismes, à la fois centralisés et décentralisés, ont déjà été discutés et, en partie, mis en œuvre, dans le domaine du commerce électronique (Wei, Xie & Zhou, 2008).

La sociologie computationnelle et quantitative apporte dans ce domaine, au moyen de simulations à large échelle des comportements des communautés P2P, des éclairages intéressants au sujet de phénomènes tels que la formation et la dissolution de groupes au fil du temps, la coopération entre pairs, la spécialisation de certains d'entre eux, les systèmes de réputation, et les comportements antagonistes occasionnels. Le travail de David Hales, informaticien et sociologue ayant longtemps travaillé avec et sur l'équipe Tribler (voir chapitre 6), et de ses collègues, est particulièrement intéressant dans son exploration des caractéristiques de groupes particuliers qu'il appelle les « tribus virtuelles » (*virtual tribes*) ; il considère que la conception d'un mécanisme durable de construction de confiance entre pairs est une précondition nécessaire au développement de systèmes robustes et résistants, car pour nombre de systèmes distribués, l'imposition d'un contrôle central « n'est tout simplement pas une option, et on ne peut pas se reposer sur les mains invisibles et orthodoxes du marché ou de la théorie des jeux » (Hales, 2006; voir aussi Marcozzi & Hales, 2008; Hales, Arteconi,

Marcozzi & Chao, 2008).

### *Responsabilisation des pairs et « free-riding »*

Le succès des communautés de partage de fichiers mis en place par Napster au début, et puis par d'autres applications similaires, est dû à une dynamique pour le moins insolite. En effet, la maximisation individuelle de l'utilisation dans les communautés P2P conduit – en ce qui concerne la communauté dans son ensemble – à des résultats collectivement souhaitables (Shirky, 2001). Cela est dû au fait que quand un fichier est téléchargé, une réplique du fichier est ajoutée à la base de données de la communauté de partage. Notamment grâce à la faculté de ces réseaux de fragmenter les fichiers, si un seul utilisateur possède un fichier qui, d'un seul coup, est demandé par un grand nombre d'autres utilisateurs, ces derniers vont commencer par en télécharger une partie, puis une autre, jusqu'à reconstruire le fichier complet. Toutes les requêtes ultérieures sur ce fichier seront partagées entre tous les utilisateurs en possédant au moins une partie.

Ces dynamiques peuvent être endommagées par les *free-riders* (que l'on peut traduire par « resquilleurs », « clandestins », ou encore « parasites »), soit quand ceux-ci refusent l'accès au fichier téléchargé, soit quand ils le déplacent immédiatement après le téléchargement, de sorte que la base de données collective n'augmente pas. Les pairs pratiquant le *free-riding* utilisent les ressources disponibles dans le réseau P2P, mais ne fournissent en contrepartie aucune ressource (Adar & Huberman, 2000). Pour la plupart des systèmes P2P en général, et pour les systèmes de partage de fichiers en particulier, cela peut créer un problème important – la capture privée de biens non rivaux (Bauwens, 2005b ; Gensollen, 2009) – et empêcher un réseau de développer son plein potentiel. Le *free-riding* réduit la disponibilité de l'information ainsi que le niveau de performance du réseau. Pour les développeurs de réseaux P2P, une solution possible pour résoudre ce problème est la « responsabilisation » des pairs. Celle-ci est mise en œuvre à travers l'identification et l'affectation des ressources utilisées ou fournies par chacun des pairs, et par l'implémentation d'incitations (*incentives*) négatifs (débits) ou positifs (crédits, sous la forme de rémunérations monétaires ou de droits d'utilisateur. Cependant, la force exécutoire de ces encouragements, leur acceptation de la part des usagers, la confidentialité des données utilisateur – et en conséquence, l'applicabilité et fonctionnalité de ces mesures – continuent à poser problème.

### *Modèles d'affaires et chaînes de valeur*

Les communautés P2P peuvent atteindre des échelles qui aboutissent à des changements de la configuration des chaînes de valeur d'un secteur, ainsi qu'à la tendance à transférer le contrôle de chaque étape de création de valeur des acteurs centraux du marché aux consommateurs et, dans certains cas, à des nouveaux intermédiaires. L'industrie de la musique fournit un exemple de ces dynamiques (Hummel & Lechner, 2001, 2002 ; Shirky, 2001).

Les communautés virtuelles ont également été prises en compte en termes de stratégies commerciales pour atteindre des objectifs économiques : des modèles d'affaires ou *business models* (Armstrong & Hagel, 1996 ; Hagel & Armstrong, 1997). On met ici l'accent sur la mesure dans laquelle les communautés virtuelles peuvent effectivement être utilisées par des personnes ayant des intérêts commerciaux, et on cherche à déterminer si et comment des incitations monétaires et non monétaires motivent les membres potentiels à participer.

#### DROIT(S) EN RÉSEAU : SÉCURITÉ, *PRIVACY* ET SURVEILLANCE

La deuxième partie de cette section se propose d'introduire des travaux qui visent à préciser et explorer les questions de sécurité et de *privacy* de l'information « en réseau », en puisant dans les sciences juridiques et de manière ponctuelle, dans la sociologie du droit. Parmi la pléthore de travaux existants à ce sujet, je vais concentrer mon attention sur un sous-ensemble qui présente un lien étroit, mais peu souvent établi, explicité ou exploré, avec la littérature en sciences informatiques qui a informé mes propos jusqu'ici. Il s'agit de travaux qui s'attachent à analyser les différentes conceptions de la *privacy* qui sont envisagées et s'inscrivent dans les dispositifs d'information et de communication en réseau, en mobilisant leurs différentes couches, de l'architecture à l'interface et plus souvent, les deux. La discussion de ces travaux contribuera à éclairer et cadrer ces aspects de ma démarche qui prêtent une attention spéciale au droit et aux droits, dans la poursuite de trois objectifs. En premier lieu, il s'agit de mener à bien l'analyse de la « légalité » des services pris en considération, dont l'architecture évolue constamment et dont les instruments de régulation écrite, comme les politiques de confidentialité établies par les fournisseurs de services Internet, peinent souvent à suivre les rythmes trop rapides et les directions imprévisibles de changement. En deuxième lieu, on s'efforcera de fournir les outils nécessaires au dépassement, dans le cas du P2P, d'une analyse des rapports entre le droit et la technologie qui se focalise trop souvent sur un seul aspect : le fait que les technologies émergentes posent des défis aux régimes légaux existants, et créent un besoin de réforme de ces mêmes régimes légaux (Elkin-Koren, 2006 ; Delmas-Marty, 2012). En troisième lieu, il s'agit de suggérer que les objets et ressources inscrits et produits avec le P2P peuvent eux-mêmes être conçus et traités en tant que systèmes de définition et de protection des droits de l'utilisateur des services Internet.

Depuis la création du Web il y a vingt ans, un certain nombre de voix se sont élevées pour définir l'Internet comme l'« espace public » le plus large du monde d'aujourd'hui, en raison des millions de personnes qui chaque jour échangent des messages, produisent et reçoivent des connaissances, renforcent leur participations politique et sociale, jouent, vendent et achètent sur l'Internet (Di Maggio et al., 2001). En soulignant qu'Internet est actuellement l'infrastructure qui rend possible l'interconnexion à l'échelle mondiale, on soutient également qu'il doit être traité, au niveau du droit, en tant que bien commun (Massit-Folléa, 2008). À côté de ces arguments, on discute des méthodes pour préserver activement l'espace public créé

par le « réseau des réseaux », face aux défis posés par les moyens de gestion actuels des services basés sur Internet, qui posent à la fois des besoins et des risques inédits pour les usagers du réseau, dans leur double capacité de consommateurs (Johnson, Crawford, & Palfrey, 2004) et de citoyens (Rodotà, 1997 ; Solove, 2006).

On souligne par exemple que les plus récentes évolutions des pratiques criminelles en ligne, et même les luttes pour la préservation du droit d'auteur, ont été à la base de la création de nouveaux instruments de contrôle des contenus et de techniques de filtrage, qui, à leur tour, affectent négativement la confidentialité des communications personnelles (Diffie & Landau, 2007). Qu'ils se concentrent sur des dynamiques de privatisation ou de fragmentation du réseau, ou encore, sur les arènes de conflit où la liberté est souvent présentée comme l'ennemi de la sécurité, plusieurs travaux se concentrent sur les conditions techno-juridiques sous-tendant la « garantie du respect des libertés et des droits pour tous les utilisateurs de l'Internet », cherchant à articuler une définition de cette expression aussi séduisante que vague (Rodotà, 2006).

D'autres travaux suivent les pratiques de surveillance de la vie quotidienne de chaque internaute via les traces laissées sur le « *world wide web of surveillance* » (Lyon, 1988: 37-47), quand il demande ou fournit des biens et des services, lors de ses recherches d'informations, lorsqu'il établit des relations sociales en ligne impliquant de la création et du partage et lorsqu'il se déplace dans l'espace réel ou virtuel. Certains se concentrent tout particulièrement sur les manières dont les contours de la représentation sociale en ligne sont progressivement dessinés par les informations laissées par chaque transaction, dispersées dans une variété de bases de données et de réseaux, tandis que les notions de public et de privé voient leurs frontières reconfigurées et se confondent dans les traces plurielles laissées sur le réseau (Solove, 2006). Dans cette lignée, des auteurs intéressés par le concept de résistance ont souligné les « aspects numériques, algorithmiques et rhizomatiques<sup>51</sup> des nouvelles pratiques de surveillance » (Martin, van Brakel & Bernhard, 2009 : 215). On souligne que les conceptualisations traditionnelles de la surveillance, en tant que relation exclusive entre le surveillant et son objet, ne tiennent pas suffisamment en compte les « assemblages de surveillance » (et de réponse à la surveillance) qui sont en train d'émerger dans les médias en réseau, et qui transforment les buts et les hiérarchies de l'activité de surveillance en même temps qu'ils reconfigurent la notion de *privacy* (Haggerty & Erickson, 2000).

Le droit à la vie privée, ou *privacy*, relatif aux expressions et à l'explicitation en ligne des informations personnelles concernant un individu, a par ailleurs fait son entrée dans plusieurs textes constitutionnels parmi les plus récents, en compagnie d'autres droits – tels que la sauvegarde de l'environnement, l'objection de conscience, l'accès à l'information, le droit à communiquer, la protection des consommateurs et la participation aux décisions administratives. Si les débats sur la vie privée précèdent, et de loin, l'ère du numérique, le croisement entre les systèmes de protection de ce droit et la multiplicité des environnements en ligne le rend désormais, selon certains auteurs, un « *droit de troisième génération* », expression des nouvelles exigences qui se dégagent des progrès de la science et de la technique (Morbidelli, Pegoraro,

---

<sup>51</sup> Au sens de Deleuze et Guattari (1972 ; 1980).

Reposo, & Volpi, 2004: 44-45 ; Evans, 2007).

Ces données personnelles ont été définies par la jurisprudence européenne plus récente avec un vocabulaire intentionnellement inclusif et général : la directive européenne sur la protection des données<sup>52</sup> les interprète comme « toute information concernant une personne physique identifiée ou identifiable [...], c'est-à-dire] qui peut être identifiée, directement ou indirectement, notamment par référence à un numéro d'identification ou à un ou plusieurs éléments spécifiques, propres à son identité physique, physiologique, psychique, économique, culturelle ou sociale », tandis que leur traitement inclut « la collecte, l'enregistrement, l'organisation, la conservation, l'adaptation ou la modification, l'extraction, la consultation, l'utilisation, la communication par transmission, diffusion ou toute autre forme de mise à disposition, le rapprochement ou l'interconnexion, ainsi que le verrouillage, l'effacement ou la destruction<sup>53</sup>. » La définition inclut donc toute information relative à un individu identifiable ; elle prend aussi en compte la possibilité qu'une information ou un ensemble d'informations pourraient ne pas être suffisamment révélateurs en elles-mêmes, mais pourraient le devenir si associées ou agrégées à d'autres informations qui pourraient être détenues par une même entité. La définition aborde donc aussi la question des différentes manières d'identifier un individu : si le nom propre est un identificateur classique, la personne peut arriver à être identifiable grâce à la combinaison et à l'agrégation de nombre d'autres informations, variant du pseudonyme aux préférences d'achat, de visionnage, d'écoute, de recherche<sup>54</sup>.

---

<sup>52</sup> Formellement la Directive 95/46/CE du Parlement européen et du Conseil, du 24 octobre 1995, relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données (<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:FR:HTML>)

<sup>53</sup> Article 2(a) et (b).

<sup>54</sup> Une des problématisations les plus intéressantes du droit de l'utilisateur de connaître d'abord, et de contrôler ensuite, quelles données le concernant sont stockées et circulent dans quels endroits matériels et virtuels, est proposée avec le concept d'*habeas data*, en lien avec les discussions autour d'une charte des droits de l'Internet (Musiani, à paraître 2012). En tant qu'évolution du concept de *habeas corpus*, point de départ historique des libertés personnelles fondamentales, le concept d'*habeas data* a été d'abord introduit dans plusieurs systèmes légaux latino-américains (Rozo-Acuna, 2002), et ensuite proposé à l'« exportation » au niveau international afin de mettre en relief l'exigence de sauvegarde des données personnelles des usagers de l'Internet, et de réponse aux changements dans la technique (Rodotà, 2006b). Un des auteurs pionniers sur le sujet soulignait en 1988 que, aujourd'hui, le droit d'un État moderne doit prévoir un instrument qui garantisse à ses citoyens leur droit d'utiliser et de disposer librement de leur données personnelles, de la même façon qu'ils utilisent et disposent librement de leur corps » (Frosini, 1998).

Les problématiques concernant l'élaboration du concept d'*habeas data* couvrent les actions des gouvernements, mais aussi des entreprises proposant des services basés sur Internet (Ganascia, 2009). En ce qui concerne les gouvernements, on se demande s'ils peuvent réellement revendiquer le droit d'intercepter les communications électroniques quotidiennes pour des raisons de sécurité ou pour surveiller des activités criminelles ; pour les entreprises, en particulier celles qui se trouvent en position de rassembler des quantités importantes de données personnelles (les grands portails de recherche sur l'Internet et les outils de réseautage social), les questions d'enjeux commerciaux directs (Boyd, 2008) et de coopération avec les gouvernements (autoritaires ou non) dans la mise à disposition de données personnelles des utilisateurs à des services, institutions ou applications tierces sont directement posés. Enfin, une question largement débattue dans le cadre de l'*habeas data* est de savoir

Si le concept de *privacy* peut être appliqué, selon le contexte, à la capacité d'un individu ou d'un groupe de protéger ou de révéler sélectivement tout type d'information personnelle confidentielle, des traits physiques à l'orientation politique, les amples discussions théoriques autour de la notion concernent sa définition même, et ce qui doit en être exclu ou inclu. Cette section retrace brièvement la partie de ce débat qui concerne la *privacy* de l'information (*information* ou *informational privacy*), la relation entre la protection et la dissémination de données confidentielles, identifiables et reconnaissables, au moyen de technologies de recherche, stockage et réseautage.

Au cœur des travaux du philosophe de la technologie Luciano Floridi sont les manières dont la commodification des TIC, et l'émergence d'une société de l'information globale, sont en train de poser des défis au droit à la *privacy* informationnelle. Pour lui, la manière la plus appropriée de comprendre l'impact croissant que les TIC sont en train d'avoir sur la *privacy* informationnelle est en termes d'une révolution ontologique, qui a changé la nature du monde de l'information d'analogique à numérique. En comparant notamment trois interprétations de la *privacy* informationnelle, réductionniste, basée sur la propriété et basée sur l'ontologie, Floridi arrive à la conclusion que la *privacy* de l'information devrait être considérée comme un droit fondamental et donc que, comme c'est le cas pour d'autres droits fondamentaux, elle devrait être poursuivie par défaut (Floridi, 2005).

En s'attaquant à la question fondamentale de ce que doit être inclus dans la définition de *privacy* et comment, Herman Tavani souligne – en rapport, plus spécifiquement, à la problématique des données informationnelles – que la définition de ce périmètre est « en rapport étroit avec la question suivante : doit-elle être comprise comme un droit à part entière, ou plutôt conçue en termes des intérêts portés par les individus ? » (Tavani, 2008 : 133). Dans le cas spécifique des réseaux informatiques en P2P, l'auteur note qu'il pourrait être possible d'appliquer le même principe de « publicité » utilisé dans la technologie des *cookies*<sup>55</sup> pour déterminer si l'information personnelle accessible à la technologie devrait être déclarée une « situation normativement privée » ; c'est-à-dire, quand toute l'information nécessaire à la compréhension de la technologie est dévoilée à l'utilisateur, celui-ci a le pouvoir de prendre une décision informée sur le choix de l'accepter ou de la rejeter (Tavani, 2008 : 160).

---

si les usages commerciaux de l'Internet ne seraient pas désormais en train de prendre le dessus sur les autres types d'utilisation. Le danger serait alors que la *network society* s'identifie toujours plus à sa dimension de marché et d'échanges monétaires, dans laquelle seraient seulement reconnus comme droits ceux liés aux transactions de biens et de services (Castells, 2000).

<sup>55</sup> Un cookie (ou témoin de connexion) est une suite d'informations envoyée par un serveur HTTP à un client HTTP, que ce dernier retourne lors de chaque interrogation du même serveur HTTP sous certaines conditions. Un cookie peut être utilisé pour une authentification, une session, ainsi que pour stocker une information spécifique sur l'utilisateur, comme les préférences d'un site ou le contenu d'un panier d'achat électronique.

De façon intéressante, en explorant l'interaction entre technologie et *privacy* à l'ère numérique, Daniel Solove souligne qu'une définition de *privacy* informationnelle, et des manières de la protéger, ne peut pas se passer d'une analyse fine de la manière dont différentes architectures de réseau influencent les pratiques de recueil et d'utilisation de l'information : « le droit de la vie privée ne doit pas seulement se concentrer sur les remèdes et les sanctions pour les individus lésés, mais sur la construction d'une architecture capable de gérer les flux de données en croissance constante qui caractérisent la société de l'information » (Solove, 2004 : 10-11). De manière semblable, Colin Bennett et Charles Raab soutiennent que les sciences politiques se doivent de prêter une attention particulière à la « régulation de l'information personnelle comme importante 'monnaie d'échange' à la fois de l'organisation capitaliste contemporaine et de l'Etat moderne » ; plutôt que pondérer la difficulté inhérente à la stabilisation d'une définition générale de *privacy*, il est souhaitable d'adopter comme point de départ les « suppositions douteuses sur l'individualisme et la distinction entre la société civile et l'Etat comme entités, » et de questionner ces suppositions afin de faire ressortir la *privacy* comme objet « politique » (Bennett & Raab, 2003 : 6).

De l'autre côté de la « *privacy equation* », des auteurs ont aussi considéré si et comment, dans des contextes spécifiques, la question de l'équilibre entre *privacy* individuelle et responsabilité sociale dans la construction de la sécurité et de la protection collectives devrait être beaucoup plus mise à l'épreuve, et en débat, qu'elle ne l'est couramment. Dans son ouvrage controversé de 1999, *The Limits of Privacy*, Amitai Etzioni a notamment fait l'argument que la *privacy* est en effet trop protégée et privilégiée, au détriment du bien commun et de la durabilité du lien social. En citant, entre autres, l'exemple des cyber-terroristes qui échangent bien plus aisément grâce à la montée en puissance des technologies d'encryptage, Etzioni soutient que la sécurité collective est à ce jour « systématiquement négligée à cause d'une déférence excessive attribuée à la vie privée » : le droit à la vie privée devrait être défini et considéré en tant que « droit individuel, à rééquilibrer à l'aune des préoccupations pour le bien commun – ou en tant que bien parmi des biens, sans privilégier a priori aucun d'entre eux » (Etzioni, 1999, p. 4).

Dans le contexte de la protection des données personnelles dans les activités de partage et réseautage en ligne, le concept de *privacy* est aujourd'hui en train d'acquiescer encore des nouvelles facettes, à cause des reconfigurations de la relation entre propriété et données personnelles, et de la perméabilité croissante des frontières du numérique qui ouvrent des nouvelles possibilités de surveillance de l'utilisateur/consommateur des services Internet (Katyal, 2005). Ces deux notions – le rapport entre propriété et exposition de données confidentielles, et la perméabilité des frontières des interactions dans l'espace numérique – permettent à Sonia Katyal de conceptualiser la *privacy* informationnelle comme la capacité d'une personne à contrôler quelles informations personnelles la concernant sont révélées sur l'Internet, et de gérer leur accès par d'autres personnes (en particulier « recherches ou appropriations dépassant la limite raisonnable », Katyal, *ibid.* : 302) ; capacité qui est constamment mise à l'épreuve par l'écosystème de technologies d'encryptage, de préservation de l'anonymat et de surveillance que constituent l'Internet d'aujourd'hui.

Très récemment, grâce notamment aux travaux de la psychologue et juriste<sup>56</sup> Ann Cavoukian, une conceptualisation de la *privacy* comme principe implanté (*embedded*) dans la technologie est en train de prendre pied, sous l'étiquette de « Privacy by Design » (Cavoukian, 2006 ; 2009 ; 2010 ; Schaar, 2010). Le PbD, pouvant se traduire en français par l'expression « la prise en compte de la vie privée dès la conception », est problématisé comme le principe techno-juridique selon lequel toute technologie exploitant les données personnelles doit intégrer la protection de la vie privée dès sa conception, et s'y conformer tout au long de son cycle de vie. C'est une forme de prévention du risque d'exploitation abusive de ces données qui se met en place, en intégrant un dispositif technique de protection juridique dès la conception des solutions informatiques destinées aux services Internet. Si le concept et l'ontologie même du PbD sont fortement débattus, des objets, des marchés, des réalités économiques commencent à se construire autour de ce concept, en entraînant l'intérêt et le suivi de la part d'instances de régulation nationales, européennes et internationales<sup>57</sup>.

En se concentrant, dans une perspective juridique, sur la relation entre *privacy* informationnelle et protection du droit d'auteur, Sonia Katyal arrive à une conclusion qui dépasse le cadre des *copyright wars* pour éclairer les manières dont la *privacy* s'inscrit dans les objets et architectures techniques (2004 ; 2005 ; 2009). Nombre de stratégies d'application du droit d'auteur, toujours plus omniprésentes et invasives, ont vu le jour au cours des dernières années, et partagent le fait de baser leur mise en œuvre et leur capacité de contrôle sur des mécanismes de surveillance « privée ». Si dans le passé, les législateurs et les spécialistes ont concentré leur attention sur d'autres méthodes de surveillance plus visibles, concernant le marketing, l'emploi ou encore la sécurité nationale, le phénomène de la « *piracy surveillance* » – les systèmes extra-judiciaires de monitoring et exécution qui identifient ou découragent les infractions de la part des consommateurs, est « théorisé de manière incomplète, dépourvu de frontières techniques, et en puissance, incontrôlable juridiquement » (Katyal, 2005 : 227).

Comme le soulignent Frances Grodzinsky et Herman Tavani, la surveillance et les manières d'échapper à la surveillance sont actuellement au centre des débats au

---

<sup>56</sup> Et actuelle Information and Privacy Commissioner de l'état de l'Ontario, Canada (<http://www.ipc.on.ca/>).

<sup>57</sup> Le PbD est désormais proposé outre-Atlantique comme un principe obligatoire à intégrer dans toutes les TIC et les technologies de sécurité comme la vidéosurveillance, basées sur la collecte, l'analyse et l'échange des données personnelles, surtout grâce au travail d'Ann Cavoukian en tant que commissaire à la vie privée et l'information de l'Ontario. Au niveau de l'Union Européenne, la Commissaire européenne chargée de la justice, le Contrôleur Européen de la Protection des Données (CEPD) et le G29 ont récemment exprimé le souhait que ce principe soit intégré à la législation européenne, en tant que principe concernant toute institution ou organisation, publique ou privée, pour qui les données personnelles constituent une importante ressource fonctionnelle et stratégique. En France, l'Agence Nationale de la Recherche a récemment organisé un atelier sur les enjeux du PbD (« Privacy by Design (PbD). Mettre la technologie au service de la vie privée : Enjeux, limites et perspectives », 23 mars 2012, Paris), qui a souligné que le PbD est actuellement l'un des chantiers les plus intéressants pour expérimenter l'articulation entre normes et standards techniques dans le domaine des TIC (voir mon compte-rendu de l'atelier, <http://adam.hypotheses.org/1230>).

croisement entre *privacy*, propriété et expression sur les réseaux parce que, quand la charge d'y identifier les infractions a été placée sur les détenteurs de droits d'auteur, cela a légitimé en quelque sorte la montée en puissance « d'une industrie entièrement nouvelle où les propriétaires de contenus ont le droit de parcourir l'Internet à la recherche de potentiels contrefacteurs » (Grodzinsky & Tavani, 2005 : 247). Cette « nouvelle » surveillance – tout en altérant la notion de propriété, de bouclier défenseur à arme offensive – contribue à faciliter et à rendre pervasifs les enregistrements des activités des consommateurs, y compris celles qui ne constituent pas une violation ; à imposer des standards d'usage et d'expression ; à réduire ou à proscrire des activités considérées comme inacceptables. La dialectique entre le droit à la vie privée et l'ensemble d'activités en réseau qui sont regroupées sous l'étiquette de « piratage » est donc importante non seulement en tant que vitrine de pratiques de surveillance sous-explorées, mais aussi en tant que démonstration du besoin de résoudre les tensions entre vie privée et présence active sur les réseaux de façon qui reflète finement, et protège, la relation entre les nouvelles technologies et les libertés personnelles (Katyal, 2004).

C'est dans cette perspective que s'inscrivent les travaux de Niva Elkin-Koren (2002 ; 2006 ; 2012) ; selon la juriste israélienne, la relation entre loi et technique se focalise trop souvent sur un aspect en particulier, les défis que les technologies émergentes posent aux régimes légaux existants, créant dès lors un besoin de réforme de ces mêmes régimes légaux ; les mesures juridiques qui concernent la technique à la fois comme cible de régulation et instrument d'exécution devraient, par ailleurs, considérer que la loi ne répond pas simplement aux nouvelles technologies, mais contribue à les façonner et en influence le design (Elkin-Koren, 2006 : 15). Les systèmes de surveillance technique et juridique qui n'implémentent pas cet aspect vont probablement influencer l'innovation technique de manière à empêcher ou limiter ses bénéfiques socio-économiques potentiels (*ibid.*, 21), ou sacrifier les principes fondamentaux de la *privacy* informationnelle au bénéfice d'une capacité de contrôle dépourvue de limites (Katyal, 2004; 2005).

Si cette conception de la surveillance n'est pas entièrement attribuable au développement des technologies en pair-à-pair, ou encore moins à l'explosion du « piratage numérique » à la suite des usages massifs de systèmes P2P pour le partage de fichiers, elle est particulièrement bien illustrée dans ce domaine en tant que représentant de la « nature paradoxale de l'Internet, [qui offre] une capacité d'expression apparemment illimitée [...] en même temps qu'un déploiement important de capacités de surveillance panoptique » (Katyal, 2005 : 228). Les usagers des réseaux P2P partagent et mettent à disposition les uns des autres des ressources de calcul et de connexion telles que la bande passante, l'espace de stockage de leur disque dur, et la capacité de calcul du processeur : en bref, les ordinateurs des usagers composant le système constituent l'infrastructure de ce système. Les défis techniques qui sont au centre de la littérature discutée dans la première partie de cette section entraînent donc un certain nombre de questions liées à la *privacy*.

En premier lieu, le flux de données qui provient d'un pair peut être compromis par les autres pairs du réseau qui sont en train de contribuer au processus même de la transmission des données, car l'architecture P2P implique que les pairs aient un accès

direct aux paquets de données. En deuxième lieu, l'utilisateur nécessite une connaissance technique du logiciel qu'il est en train d'utiliser et doit avoir une idée des types de matériaux et d'informations partagés (ou de ceux qu'il ne souhaite pas partager), car « il est bien possible de se retrouver à partager le disque dur dans son entièreté, y compris des informations confidentielles telles que la boîte de messagerie électronique ou des documents privés » (Suvanto, 2005). Plusieurs études notent que les utilisateurs, pour la plupart, ignorent ou sont seulement partiellement conscients des implications de l'utilisation d'une application P2P vis-à-vis de la *privacy* ; cela est lié en partie à la grande facilité d'usage de plusieurs applications P2P de partage de fichiers, qui rend la tâche plus aisée pour des attaquants potentiels (Li, 2007). Enfin, les informations envoyées sur le réseau par le logiciel client P2P à des fins de connexion entre pairs soulèvent également des questions relatives à la protection de la vie privée ; ces informations incluent l'adresse IP des utilisateurs, ce qui, dans certains des premiers (et très populaires) systèmes P2P de partage de fichiers, était utilisé ouvertement et exposé directement, posant des questions de manque d'anonymat et ses répercussions juridiques (Suvanto, 2005).

Afin de contrer ces limites et s'opposer à leur exploitation par une entité quelconque qui aurait l'intention de compromettre la stabilité et l'intégrité du réseau – et faciliter ainsi l'adaptation des technologies P2P aux services Internet – les développements de recherche les plus récents dans le domaine du P2P se sont concentrés notamment sur deux aspects : l'encryptage des contenus et le renforcement de l'anonymat. D'un côté, l'encryptage du trafic P2P permet non seulement de sécuriser les données de manière plus sûre, mais les flux de connexion P2P eux-mêmes sont moins facilement identifiables en tant que tels, rendant le trafic de données plus résistant aux attaques et aux blocages. L'anonymisation des pairs permet par ailleurs de protéger l'identité des nœuds/usagers sur le réseau, en particulier lorsqu'elle est combinée avec des techniques d'encryptage :

Si l'anonymat absolu ne peut vraiment exister sur un réseau, un P2P anonymisé en fournit suffisamment pour qu'il soit extrêmement difficile de trouver la source ou la destination d'un flux de données. Il le fait en faisant de tous les pairs d'un réseau des expéditeurs et des récepteurs universels, en rendant virtuellement impossible de déterminer si un pair est en train de recevoir un paquet de données ou est simplement en train de le transmettre [...] l'utilisation de l'encryptage conjointement au P2P anonyme permet probablement l'expérience d'usage P2P la plus sûre d'aujourd'hui (Li, 2007).

En conclusion, la *privacy* de l'information et des données personnelles en réseau peut être définie et conceptualisée – ainsi qu'assurée ou entravée – au moyen de différentes modalités de traitement et de gestion technique des données, et selon différentes priorités dérivant de la concurrence entre droits, des évolutions des usages, et de la mise en œuvre de technologies de surveillance et contre-surveillance. Des questions telles que l'agrégation, la circulation et le stockage de l'information et des contenus (Preibusch et al., 2007), la pertinence, et la réactivité à la demande, de modèles de réseau distribués et décentralisés pour la distribution à coût limité de contenus numériques (Wood, 2010), leur stabilité et redondance dans la préservation des données (Hales, 2006), et le rôle des intermédiaires sans les connexions entre

usagers (Elkin-Koren, 2006), sont de plus en plus parties intégrantes de discours et pratiques techno-juridiques autour de la confidentialité, la *privacy* et la sécurité. Ces discours et pratiques, comme on le verra au fil des prochains chapitres, se construisent dès les premières phases de la démarche de développement et d'innovation des services Internet au moyen du P2P ; notamment dans les choix qui déterminent la forme, l'organisation et le modèle de réseau constituant l'architecture de ces services.

### SECTION 3. DE L' « INTÉRÊT DES TUYAUX » : ÉTUDIER LES ARCHITECTURES

« Le P2P, c'est des tuyaux. La plupart des gens ne s'intéressent pas aux tuyaux, » soulignait il y a quelques années Dan Bricklin – le développeur de la première feuille de calcul, *VisiCalc* – dans un des premiers ouvrages portant sur le potentiel « perturbateur » des technologies P2P (Bricklin, 2001 : 59). La « plupart des gens » auxquels Bricklin se réfère dans cette citation est constituée bien sûr par les utilisateurs des premières et populaires applications de partage de fichiers en P2P qui, comme Napster, connaissaient leur heure de gloire à l'aube du XXI<sup>e</sup> siècle.

Dan Bricklin a très probablement eu raison dans son évaluation des raisons du succès auprès du grand public des applications P2P pionnières, destinées à servir le partage massif de contenus numériques : celui-ci doit son succès à la capacité de ces systèmes à trouver et obtenir rapidement un contenu spécifique qui suscite l'intérêt de l'utilisateur, plutôt qu'à l'architecture en pair-à-pair qui les sous-tend. Pourtant, cette remarque a joué un rôle très important dans la définition du regard que j'ai porté sur des objets variés, dont le point commun – et la spécificité par rapport à leurs homologues centralisés plus célèbres – était la logique P2P qui sous-tendait leur modèle technique. La qualité de mon analyse allait en grande partie dépendre de ma capacité à m'« intéresser aux tuyaux », ou, comme Susan Leigh Star l'a très bien formulé, à « faire remonter à la surface le travail invisible » (Star, 1999 : 385) sous-tendant les pratiques, les usages et les échanges en réseau – comme partie intégrante et incontournable des « processus de constitution, d'organisation et de changement de [...] la société en réseau » (Castells, 2000 : 693).

Cette dernière section est donc l'occasion de rassembler et discuter un ensemble de travaux réunis autour d'un intérêt pour l'architecture du réseau Internet comme dispositif politique, juridique et économique – une « propriété relationnelle » (Star & Ruhleder, 1996 : 113), la trame même des comportements et des interactions des usagers en réseau. Il s'agit de travaux qui ont profondément marqué mon approche consistant à rechercher dans les « couches inférieures » d'un réseau certaines des raisons à la base de choix de développement, de réappropriation par les usagers, de formes d'organisation et de régulation. Ces travaux ouvrent la voie à une analyse des trajectoires d'innovation de l'Internet qui contribue à éclairer comment des formes spécifiques de distribution et décentralisation de l'architecture (ou leur manque) influencent procédures, pratiques et usages. Ils constituent le fondement de la démarche qui, au cours de cette thèse, essaiera de montrer les influences réciproques entre le développement de l'architecture et de l'organisation technique des services P2P alternatifs – leurs points de passage obligés, lieux d'échange et de stockage, croisements des flux – et la définition de leurs usagers, en tant que consommateurs, partageurs, producteurs et « gestionnaires » d'informations personnelles et contenus numériques.

L'architecture d'un réseau ou d'une application peut être définie comme la structure technique qui le ou la sous-tend (van Schewick, 2010), conçue selon une « matrice de concepts » (Agre, 2003) : c'est-à-dire, son agencement logique et structurel, fait d'équipements de transmission, protocoles de communication, éléments d'infrastructure, et connectivité entre ses composantes ou nœuds<sup>58</sup>. Comme le souligne Barbara van Schewick, le cloisonnement des disciplines a pu dans le passé entraîner une conception des architectures techniques en tant qu'artefacts « pertinents seulement pour les ingénieurs », et à laisser donc exclusivement à la compétence de ces derniers (2010 : 2).

Cependant, à l'intérieur de la « galaxie » des *Internet studies*, se trouve aujourd'hui un nombre croissant d'initiatives interdisciplinaires, ayant pour pierres angulaires des travaux comme ceux de Yochai Benkler sur le partage en ligne en tant que véritable paradigme de production économique (Benkler, 2004) et ceux de Lawrence Lessig, déjà évoqués, sur le « *code as law* » (Lessig, 2002). Ces travaux partagent un intérêt pour ces aspects des artefacts socio-techniques qui, bien que « transparents » pour l'utilisateur par volonté de leurs créateurs, représentent une part importante et intégrante des pratiques et des usages en réseau. En particulier, les *software studies*, les *critical code studies* et les *cyberinfrastructure studies* ont récemment relevé le défi de l'interdisciplinarité, en puisant dans les approches de la sociologie des techniques et des STS pour explorer les qualités sociales et politiques des infrastructures. Des auteurs travaillant au croisement de l'informatique, de la sociologie, de l'économie et du droit explorent également des approches méthodologiques novatrices des architectures, en travaillant à intégrer dans une même perspective le développement et l'évolution de celles-ci avec les pratiques et les usages qui ont lieu « à leur surface ». Le reste de cette section explore ces travaux de manière plus détaillée.

#### *POUR UNE INTÉGRATION DES ARCHITECTURES ET DES PRATIQUES*

Une référence certaine pour toute tentative d'intégration des liens entre architectures et usages dans l'analyse des technologies de réseau est constituée par les travaux que Susan Leigh Star, Geoffrey Bowker et collègues ont mené au cours des derniers quinze ans. Les infrastructures informationnelles sont explorées dans ces travaux en tant que systèmes socio-techniques complexes, informés non seulement par des éléments matériels et virtuels invisibles à l'utilisateur final, mais aussi par des facteurs tels que l'organisation sociale et le partage des connaissances – et les informant à leur tour (Star & Ruhleder, 1996; Neumann & Star, 1996; Star, 1999; Star, 2002; Star & Bowker, 2006).

---

<sup>58</sup> Le *IEEE Standard for Architectural Description of Software-Intensive Systems* (IEEE P1471/D5.3) définit l'architecture [technique] comme « l'organisation à la base d'un système, représentée par ses composantes, les relations qu'elles entretiennent les unes avec les autres et à l'environnement extérieur, et les principes qui guident son design et son évolution » (Bredemayer & Malan, 2001).

Dans son « appel à l'étude de choses ennuyeuses » (Star, 1999), Susan Leigh Star transmet de manière efficace l'idée que les choix dans le design d'une architecture de réseau, les spécifications techniques qui y sont incorporées, les standards et les séquences de développement ne sont pas moins importants pour l'étude des systèmes d'information puisqu'ils sont « des mécanismes cachés, sous-tendant ces procédures qui sont plus familières aux chercheurs en sciences sociales (Star, 1999 : 337). Ainsi qu'elle l'écrit dans un article sur l'ethnographie des infrastructures, qui a fait école par la suite :

Il est nécessaire de creuser un peu afin de déterrer les tensions<sup>59</sup> inhérentes à la création et la conception d'un système, et redonner leur qualités narratives à celles qui semblent n'être que des listes mortes [...] La plupart des études ethnographiques des systèmes d'information comportent implicitement l'étude des infrastructures. [...] Par ailleurs, il est plus facile de rester dans les compétences traditionnelles de notre champ d'études : le discours, la communauté, l'identité, les dynamiques de groupe, comme elles sont médiées aujourd'hui par les technologies de l'information. [...] Mais étudiez un système d'information et négligez ses standards, ses fils, son contexte, et vous allez manquer des aspects tout aussi essentiels d'esthétique, de justice, et de changement (Star, 1999 : 337-339).

Cette approche « relationnelle » entraîne d'importants changements dans les méthodes, puisque la portée du travail de terrain s'élargit jusqu'à inclure les arènes où les formes de l'infrastructure et de son architecture sont observées, déconstruites, reconstruites, et où des décisions politiques – de nom ou de fait – sont prises quant aux codes, normes techniques, bricolages, reconfigurations (Star & Bowker, 2006 : 151-152), où le chercheur s'attache à une combinaison d'« analyses historique et littéraire, d'instruments traditionnels comme des entretiens et des observations, analyse de système, et *usability studies* » (Star, 1999 : 382).

Des courants d'étude émergents comme les *software studies*, les *critical code studies* et les *cyberinfrastructure studies* (Manovich, 2001; Fuller, 2008; Marino, 2006; Ribes & Lee, 2010) puisent dans cette approche dans la poursuite de leur objectif, celui d'équilibrer « l'utilisation de termes critiques comme 'virtualité' [...] et l'engagement dans une recherche documentaire méticuleuse afin de récupérer et stabiliser les traces matérielles des nouveaux médias » (Kirschenbaum, 2003). C'est donc la matérialité du logiciel, du code, des flux de données – des éléments qui, dans l'expérience du développeur et de l'utilisateur de l'Internet, sont appelés « virtuels » – qui est réaffirmée, et la relation entre ces éléments, ou niveaux, qui est explorée :

La signification [du système] dérive du fonctionnement du code, mais n'est pas limitée aux procédures que le code met en œuvre. Avec les *critical code studies*, les praticiens peuvent s'attaquer aux systèmes humains et informatiques complexes du niveau de l'ordinateur au niveau de la société dans laquelle ces code objects circulent et exercent leur influence (Marino,

---

<sup>59</sup> Star utilise le mot *drama*, à comprendre dans ce contexte comme « un état, une situation ou une série d'événements impliquant des conflits de forces ou de positions ».

2006).

*ARCHITECTURES : POLITIQUES, SOCIALES, LÉGALES*

A la littérature à vocation STS, explorant les qualités sociales et politiques des « tuyaux » des systèmes d'information, sont nécessairement complétés un ensemble de travaux qui, pour la plupart, se revendiquent explicitement de l'interdisciplinarité, et se réunissent autour d'un intérêt pour l'architecture du réseau Internet comme dispositif politique, juridique et économique.

Philip Agre, spécialiste des systèmes d'information et pionnier de l'Internet, a exploré au cours de sa carrière éclectique la relation entre l'architecture technique et les institutions, en particulier la différence entre « l'architecture comme politique » et « l'architecture comme un substitut de la politique » (Agre, 2003). Il fait l'argument que les technologies nous parviennent souvent « enveloppées d'histoires qui concernent la politique » ; si ces histoires peuvent ne pas nous expliquer les raisons des développeurs, elles sont utiles pour rendre compte de l'« énergie » qui rend une technologie sociale de par sa propre nature (*ibid.*, 39). En définissant les architectures comme les matrices de concepts (par exemple la distinction entre un client et un serveur) inscrits dans la technologie, et les institutions comme les matrices de concepts qui organisent le langage, les règles, les professions et autres catégories sociales dans des secteurs spécifiques, Agre suggère que l'histoire de l'ingénierie du calcul et des réseaux distribués, d'un côté, et l'histoire politique du changement institutionnel au moyen d'architectures décentralisées, de l'autre, ne sont pas naturellement reliées. Elles se reconfigurent et évoluent constamment, et pour que ces reconfigurations et évolutions partagent une direction commune, elles ont besoin de « travail » :

Des institutions décentralisées n'impliquent pas nécessairement des architectures décentralisées, ou vice-versa. La motivation à la base des architectures décentralisées ne sert pas *a priori* l'objectif politique de décentraliser la société. Cependant, les architectures et les institutions co-évoluent inévitablement, et dans la mesure où elles peuvent être conçues, elles devraient l'être ensemble [...] Certaines directions d'innovation des technologies de l'information et de la communication ouvrent en effet de nouvelles possibilités de changement institutionnel. Mais pour explorer ces possibilités, les développeurs auront besoin de meilleures idées au sujet des institutions (Agre, 2003 : 42).

Au sein d'un projet à large échelle qui recherche, dans le corpus de *Requests for Comments* (RfC) de l'Internet Engineering Task Force, des indications sur les manières dont les concepteurs techniques de l'Internet comprenaient et s'engageaient dans des questions politiques et juridiques, Sandra Braman a très récemment exploré comment la problématique centrale du design technique de l'Internet était de construire le « réseau des réseaux » sur des architectures qui non seulement toléraient, mais facilitaient le changement et la modification. En examinant les manières dont le

changement et la stabilité mêmes étaient conceptualisés par les designers de l'Internet, Braman conclut que l'étude de la « conception pour l'instabilité » (*design for instability*) inscrite dans son architecture est une partie intégrante du social et du légal qui se construit non seulement de par et avec l'Internet, mais aussi dans ses relations avec les autres TIC (Braman, 2011).

Au croisement de l'économie, du droit et de l'informatique, Barbara van Schewick a récemment exploré les manières dont l'architecture de l'Internet, et de ses applications, est pertinente pour l'économie. Son travail examine comment les changements, surtout les choix de design, dans l'architecture Internet (qu'elle définit de manière opérationnelle comme la « structure technique sous-tendant » le réseau des réseaux) ont une incidence sur l'environnement économique de l'innovation ; elle évalue l'impact de ces changements à partir d'une perspective politique et juridique (2010 : 2). Van Schewick souhaite contribuer à combler une lacune dans l'étude et la compréhension des décisions des innovateurs des « couches inférieures » de l'Internet ; elle soutient qu'après de nombreuses années de recherche sur les processus d'innovation, on a maintenant une compréhension fine de la manière dont ceux-ci influencent et sont influencés par des changements dans le droit, les normes, les prix. Mais l'impact réciproque de l'innovation et de l'architecture reste sous-exploré, probablement à cause des qualités « intrinsèquement techniques » de cette dernière :

Si l'architecture d'une maison en décrit sa structure interne de base, il en est de même pour l'architecture d'un système complexe – ses composantes, ce qu'elles font, et comment elles interagissent pour créer les fonctionnalités du système. Que cette structure technique puisse avoir des conséquences économiques est une contribution relativement récente. En général, on pense aux architectures comme artefacts techniques d'exclusive pertinence des ingénieurs. Mais comprendre comment l'architecture de l'Internet se rapporte à l'innovation nous force à penser, plus généralement, à comment les architectures affectent l'innovation (van Schewick, 2010 : 4).

Elle conclut que, traditionnellement, les décideurs politiques se sont servis du droit pour entraîner les effets économiques souhaités ; l'architecture constitue *de facto* une manière alternative d'influencer les systèmes économiques, et en tant que telle, elle devient un autre instrument dont les acteurs de l'économie Internet, des développeurs aux politiques en passant par les usagers, peuvent se servir pour faire avancer leurs intérêts (ibid., 389).

#### QUELLES ARCHITECTURES POUR LE FUTUR INTERNET ?

En conceptualisant les architectures de réseau comme politiques, sociales et légales, certains auteurs s'attachent plus spécifiquement à examiner les manières dont certaines d'entre elles se rapportent ou peuvent se rapporter à l'innovation dans le domaine des services Internet. Les trajectoires d'innovation actuelles du réseau des réseaux rendraient de plus en plus évident que ses évolutions (et in-volutions)

dépendront probablement, à moyen et à long terme, de la topologie et des modèles organisationnels et techniques des applications basées sur Internet, et de l'infrastructure qui les sous-tend (Aigrain, 2011).

Le développement de services basés sur des architectures (au moins partiellement) décentralisées est reconnu depuis des années comme l'un des axes de transformation potentiellement plus importants pour l'Internet : le concept de distribution et décentralisation est en quelque sorte inscrit et informé dans les débuts de l'Internet, notamment dans l'organisation et la circulation des flux de données, mais sa topologie actuelle intègre ce principe structurant de manière limitée (Minar & Hedlund, 2001 ; Berners-Lee, 2010). Les limites de cet urbanisme de l'Internet, « classique » et prédominant depuis les débuts de sa phase commerciale et son appropriation par le grand public, ressortent et se proposent à l'attention des spécialistes de façon régulière, notamment à la lueur du succès retentissant qu'ont connu certains services comme les médias sociaux. Alors que chaque internaute est devenu, au moins potentiellement, un consommateur mais aussi un distributeur et un producteur de contenus numériques, cet urbanisme conduit à centraliser des masses considérables de données dans certaines régions de l'Internet, alors même qu'elles ont vocation à être rediffusées aussi vite dans de multiples places d'un réseau désormais pleinement globalisé (Moglen, 2010).

#### *LES ARCHITECTURES ET LA « VALEUR SOCIALE » DU MODÈLE INTERNET*

Le mode actuel d'organisation des services Internet et de la structure du réseau qui rend possible leur fonctionnement, avec ses points de passage obligés, ses carrefours plus ou moins contraints, ses « routes à péage », soulève nombre de questions en termes à la fois d'une utilisation optimisée des ressources de stockage, et de la fluidité, rapidité et efficacité des échanges électroniques. S'y ajoutent des interrogations sur la sécurité des échanges et la stabilité du réseau : une série de dysfonctionnements et de pannes aux conséquences globales, qu'on a pu introduire brièvement dans le premier chapitre, a récemment attiré l'attention de plusieurs commentateurs techniques sur des questions de sécurité et de protection des données qui sont inhérentes à la structure actuelle de l'Internet (Bortzmeyer, 2008).

Ces questions influencent largement l'équilibre des pouvoirs entre développeurs, usagers et opérateurs de réseau, jusqu'à rejoindre des questions de neutralité de l'Internet : c'est le cas, par exemple, des « bridages » imposés par les opérateurs à certains usages spécifiques, ou de l'optimisation du réseau pour certains types de trafic ou protocoles au détriment d'autres (Schafer, Le Crosnier & Musiani, 2011 : 66-68). Comme le souligne Barbara van Schewick, en influençant les appropriations de l'Internet par les usagers – de manière qui crée plus ou moins d'ouverture, d'échange, de diffusion – les changements dans l'architecture du réseau des réseaux influencent sa « valeur sociale » :

Mais la valeur sociale des architectures [...] dépasse cela [les questions de neutralité technique]. L'Internet a le potentiel d'augmenter les libertés

individuelles, fournir une plateforme pour une participation démocratique plus poussée, proposer plus de réflexivité et de culture critique [...] mais la capacité de l'Internet de réaliser ce potentiel est strictement liée à des caractéristiques telles que la liberté de choix de l'utilisateur, la non-discrimination, la non-optimisation (van Schewick, 2010 : 387).

Ces caractéristiques peuvent être atteintes de manières différentes en concevant de façons différentes l'architecture technique qui les sous-tend (Moglen, 2010). Recourir à des architectures décentralisées et à des formes d'organisation distribuées du réseau est donc une manière alternative d'aborder certaines questions critiques de gestion du réseau, dans une perspective d'efficacité, de sécurité et de « développement numérique durable » au moyen d'une meilleure gestion des ressources, et de la maximisation de la valeur de l'Internet pour la société<sup>60</sup>.

### *L'ARCHITECTURE P2P ENTRE PASSÉ ET FUTUR*

Le principe de décentralisation a gouverné, depuis les origines de l'Internet, la circulation des transmissions et des communications sur le réseau des réseaux (Aigrain, 2011). Cependant, l'introduction du World Wide Web en 1990 a largement conduit à la diffusion d'architectures client-serveur – les services basés sur Internet les plus répandus, qu'il s'agisse de réseaux sociaux, d'instruments de communication ou de services de stockage de contenus numériques, sont basés sur des modèles techniques et économiques dans lesquels les utilisateurs demandent de l'information, des données, des services à de puissantes « fermes » de serveurs qui stockent de l'information et gèrent le trafic du réseau (van Schewick, 2010 : 70). Même si le trafic sur l'Internet fonctionne sur le principe de la distribution généralisée, il a donc désormais pris la forme d'une concentration autour des serveurs qui donnent accès aux contenus. Pourtant, ce mode d'organisation des services et de la structure du réseau n'est pas le seul, ni peut-être le plus efficace possible et des alternatives sont recherchées (Aigrain, 2010, 2011; Moglen, 2010).

Comme on verra plus dans le détail dans le prochain chapitre, le modèle P2P n'est pas, en soi, novateur dans l'histoire de l'Internet. En fait, l'Internet des origines avait été conçu, essentiellement, en tant que système P2P, avant que le réseau ne commence à être peuplé par un nombre d'utilisateurs finaux toujours en croissance, et ne devienne l'instrument au moyen duquel des millions de clients/consommateurs communiquent avec un ensemble de serveurs « relativement privilégié » (Minar & Hedlund, 2001 : 4).

Au fur et à mesure que d'autres évolutions de l'Internet se mettaient en place –

---

<sup>60</sup> Cette idée est explorée ultérieurement par le philosophe Michel Bauwens (2005a ; 2005b) qui, en proposant une vision du modèle P2P qui est basée sur sa conception en tant que technologie de réseau informatique, mais se propose de la dépasser pour préconiser que le P2P soit le fondement d'une « théorie générale » des interactions humaines collaboratives et directes – un phénomène émergent, pervasive et social qui pourrait être en train d'influencer profondément les manières dont la société et la civilisation humaine même est organisée.

l'augmentation de la quantité et de la qualité de bande passante à disposition de chaque utilisateur ; la puissance croissante des terminaux domestiques ; la diversification des « activités numériques » au delà de la navigation du Web et de l'envoi de messages – les conditions se formaient pour un autre changement, voire un revirement, basé sur « la connexion réciproque et directe de machines domestiques, formant des groupes et collaborant pour devenir des moteurs de recherche, des outils de calcul et des systèmes de fichiers créés par les utilisateurs » (Minar & Hedlund, 2001 : 3). C'est donc en relevant les moments et points matériels où l'architecture de l'Internet a été soumise à des tensions que les développeurs d'applications en viennent à se retourner avec profit vers l'Internet d'il y a quinze ou vingt ans, en considérant les manières plus efficaces de résoudre un problème (Figueiredo *et al.*, 2008).

La problématisation de l'architecture de l'Internet, et du modèle P2P en particulier, dont on a fourni les appuis théoriques au cours de cette section, n'est pas sans écueils. S'il s'agit d'éviter une interprétation réductionniste de l' « effet P2P » comme synonyme de pratiques douteuses sur les réseaux, il convient également d'éviter les pièges que le P2P, un modèle *a priori* doté de fortes connotations d'égalité et de décentralisation, peut proposer. Comme l'a souligné Philip Agre, il est particulièrement facile dans le cas du P2P de juxtaposer l'architecture aux histoires des institutions, des individus et des groupes, en assumant que l'un détermine l'autre – mais il s'agirait d'un raccourci trompeur :

Dans le cas des technologies P2P, l'histoire officielle des ingénieurs est que l'effort de calcul doit être distribué pour refléter la nature du problème. Mais l'histoire de cette ingénierie n'explique pas à elle seule les sentiments forts que le modèle P2P évoque souvent. Ces sentiments dérivent d'une histoire politique qui est souvent démentie avec véhémence par les technologues, mais répandue dans la culture : le P2P 'tient la promesse' de décentralisation faite par l'Internet. En minimisant le rôle des éléments de calcul centralisé, selon l'histoire officielle, les systèmes P2P seront immunes à la censure, au monopole, à la régulation, et à d'autres formes d'exécution d'une autorité centralisée. Cette juxtaposition de l'ingénierie et de la politique est assez répandue, et pour une raison évidente : les artefacts construits, comme l'Internet, sont inscrits (*embedded*) dans la société de façons compliquées. [...] Le cas du P2P [est utile] pour analyser le façonnage réciproque entre ingénierie et politique – ou, comme je me plais à le dire, entre architectures et institutions. [Mais pour ce faire, il faut comprendre que] l'architecture est du politique, tout en n'assumant pas que l'architecture soit un substitut du politique (Agre, 2003 : 39-42).

La clé pour une analyse et une compréhension fine des systèmes P2P semble donc résider dans une approche qui étudie et explore les architectures en tant que trame d'interactions, de choix, de controverses, de négociations, et examine comment ceux-ci façonnent, en retour, les conceptualisations et reconceptualisations de ces systèmes. Il s'agit d'arriver à « lire ces couches invisibles de contrôle et d'accès. Et pour comprendre comment celles-ci opèrent, il est nécessaire de 'déconstruire' les moments routiniers, voir ennuyeux [de la démarche d'innovation], pour en faire

ressortir les narratives qu'ils contiennent et les scènes qui, dans les coulisses [...], sont partie intégrante de la culture matérielle des sciences de l'information » (Star, 2002 : 110).

## EN GUISE DE SYNTHÈSE

En puisant dans la grande variété de travaux qui font aujourd'hui part du champ des études de l'Internet – sans pour autant prétendre à une impossible exhaustivité – ce chapitre a souhaité mettre en relief les perspectives, les approches et les outils qui guident l'analyse des matériaux empiriques présentés dans les chapitres suivants. Il a montré l'intérêt de suivre le développement d'innovations « en réseau », comme les services Internet en P2P, à l'aide d'outils conceptuels et méthodologiques qui dérivent des croisements entre les *Internet studies*, la sociologie de l'innovation technique et la sociologie des usages des TIC. La première section a tout particulièrement souligné l'attention de ces champs d'étude à la contextualité et à la pratique des processus d'innovation, au suivi de l'innovateur comme « ingénieur hétérogène », et à la capacité des utilisateurs à co-construire les dispositifs, leurs usages, et leurs représentations. Le chapitre a également cherché à établir des liens entre des travaux en informatique et en droit sur les réseaux décentralisés, avec l'idée d'éclairer le substrat technique qui est nécessaire précondition à mes analyses – tout en montrant clairement que le P2P ne peut pas se résumer à ce seul substrat technique. Enfin, le chapitre a discuté un ensemble de travaux réunis autour d'un intérêt commun pour l'architecture du réseau Internet comme dispositif politique, juridique et économique, pour ouvrir la voie à une analyse qui montre les influences réciproques entre le développement de l'architecture et de l'organisation technique des services P2P alternatifs, et la responsabilité, le degré de contrôle, les pratiques de leurs usagers.

C'est la perspective que je souhaite adopter pendant que, au fil des chapitres suivants, je proposerai une réélaboration de l'histoire du P2P et suivrai trois démarches d'innovation socio-techniques, basées sur le choix d'un modèle P2P et destinées à fournir une « alternative » à des services Internet bien connus : un moteur de recherche, un espace de stockage en ligne, un logiciel de streaming vidéo.

## CHAPITRE 3

### P2P « DÉCENTRALISÉ » : S'ÉLOIGNER DU SERVEUR POUR LE REPENSER

#### INTRODUCTION

Internet est une ressource partagée – un réseau coopératif construit sur des millions d'*hôtes* répandus dans le monde. A ce jour, un nombre toujours croissant d'applications utilise le réseau, envoie des paquets, et consomme de la bande passante. Dès 1994, le public commence à rejoindre en masse la communauté d'ordinateurs qui composent l'Internet, posant problème pour la durabilité de la plus basique des ressources : la bande passante. De plus, la confiance croissante dont les utilisateurs investissent le « réseau des réseaux », pour des usages et des applications à la fois fondamentaux et critiques, entraîne de nouveaux critères de sécurité, conduisant à la mise en place de pare-feu qui divisent fortement l'Internet en régions s'auto-alimentant et se suffisant à elles-mêmes.

Le système Internet passe à l'échelle de façon bien plus importante qu'il n'était prévu dans les intentions originelles de ses développeurs, et les paquets d'information se frayent un chemin entre des fournisseurs d'accès au réseau de plus en plus encombrés. La modélisation du trafic par des lois statistiques classiques, qui suffisaient jusqu'aux années 1990, présente désormais de nombreuses difficultés additionnelles dues au « caractère décentralisé, stochastique [...], non linéaire [...], complexe » (Baccelli, 2005 : 2) d'un réseau très étendu et impliquant des interactions entre et avec de nombreux intermédiaires<sup>61</sup>. Les raisons de cette complexification sont à chercher dans l'utilisation de nouveaux protocoles de routage et de contrôle des transmissions, qui ont donné lieu à l'extension potentiellement indéfinie et indéfinissable du réseau (Baccelli, 2005 : 3), tout en apportant des variations – et des tensions – aux « protocoles coopératifs » TCP et IP (Minar & Hedlund, 2005 : 13).

C'est ainsi qu'au cours de ces dernières années, un grand nombre de modèles sont proposés afin de prendre en compte les nouvelles complexités du réseau, tels que la multi-fractalité du trafic global ou l'équité du partage d'un même canal de

---

<sup>61</sup> On analyse plus dans le détail les questions de modélisation de l'Internet et leur « mise en controverse » dans [Musiani & Schafer, 2011].

communication entre plusieurs flux de données (Baccelli, 2005) : en particulier, autour de l'année 2000, certains développeurs commencent – dans l'effort de répondre à certains points d'interrogation posés par la massification des usages – à revenir sur des chemins déjà parcourus. Aux côtés d'un modèle de réseau ayant survécu à la croissance importante des cinq années précédentes, apparaissent des modèles qui tiennent d'un passé un peu moins récent.

En particulier, au sein d'un large mouvement appelé « peer-to-peer », initié par l'application de partage de fichiers musicaux Napster<sup>62</sup>, les millions d'utilisateurs connectés à l'Internet commencent à utiliser leurs ordinateurs personnels et domestiques, toujours plus puissants, pour faire autre chose que naviguer sur le Web et échanger des courriels électroniques. Les machines, dans les maisons et dans les bureaux, (re)commencent à se connecter les unes aux autres directement, formant des groupes et des clusters, pour prendre la forme de moteurs de recherche, machines de calcul et systèmes de fichiers créés par des utilisateurs – et composés par eux.

Nombre de questions et réflexions qui tiennent à la fois du juridique, de l'économique, du social et du technique sont soulevées par ces reconfigurations. Plusieurs fournisseurs d'accès au réseau, ayant programmé leurs systèmes sur la base de l'idée que les utilisateurs passaient le plus clair de leur temps en ligne à télécharger données et informations depuis des serveurs centraux, soulèvent des objections de nature économique aux modèles P2P. Certains commencent par empêcher l'accès aux services P2P sous prétexte que ceux-ci constituent une violation des conditions d'usage et qu'ils consomment, voire gaspillent, trop de bande passante<sup>63</sup>, tandis que derrière ces revendications se cachent souvent des débats sur la nature licite ou illicite de certains types d'échanges et de contenus partagés. Par ailleurs, de nombreuses entreprises du secteur IT tentent d'identifier les problèmes spécifiques qui se posent quand l'architecture de l'Internet est durement mise à l'épreuve par les scénarios de « concentration » actuels – et, en cherchant les meilleures façons de contourner ces problèmes et les solutions les plus efficaces, ils se retournent, de façon peut-être surprenante, vers l'Internet d'il y a quinze ou vingt ans.

Le premier Internet a été créé, foncièrement, en tant que système P2P. Au fil du temps, il est de plus en plus construit selon un modèle client-serveur, avec des millions de clients/consommateurs communiquant avec un ensemble relativement réduit de serveurs. Nombre d'applications construites sur le modèle P2P, que l'on a vu naître au cours de la dernière décennie, utilisent à nouveau l'Internet de la façon dont il avait été conçu : comme un moyen de communication entre machines qui partagent des ressources l'une avec l'autre, comme des égaux. Ce modèle de réseau ne doit pas sa « nouveauté » à son concept, donc : l'innovation réside dans son passage à l'échelle et ses implémentations particulières. Les concepteurs

---

<sup>62</sup> Comme on verra plus tard, le P2P est loin d'être né avec Napster, mais c'est avec ce logiciel que ce modèle de réseau informatique se fait connaître du grand public.

<sup>63</sup> Voir par exemple la controverse qui a suivi, en février 2011, la décision de Telecom Italia de poser des limitations au trafic P2P en heure de pointe (<http://www.webnews.it/2011/02/21/telecom-italia-minaccia-il-p2p/>). Voir aussi [Elkin-Koren, 2006].

d'applications courantes cherchent à apprendre des systèmes Internet distribués ou avec d'importantes composantes distribuées, comme Usenet ou le Domain Name System ; dans d'autres cas, ils cherchent à inverser, modifier ou reconfigurer les changements auxquels l'Internet a fait face dans sa phase de commercialisation et d'ouverture au grand public, afin d'adapter ou réadapter le réseau à de nouvelles applications P2P. Comme souligne à ce propos un développeur :

En tout cas, les leçons qu'on apprend de ces systèmes sont instructives, et pourraient nous aider – nous, les concepteurs d'applications – à empêcher la mort de l'Internet.<sup>64</sup>

La section qui suit présente une réélaboration de l'histoire du P2P. A partir de cette histoire, on souhaite mettre en lumière les aspects dont les développeurs d'aujourd'hui s'inspirent dans la démarche de « décentralisation » de leurs applications. On discutera ensuite des implications qu'ont ces démarches pour les définitions et reconfigurations à la fois du concept de décentralisation, et de l'« objet » serveur qui en découle.

---

<sup>64</sup> Entretien avec Gwendal Simon, enseignant-chercheur à Télécom Bretagne, 19 mars 2009.

## SECTION 1. INTERNET DÉCENTRALISÉ, ENTRE COOPÉRATIONS ET TENSIONS

L'Internet tel qu'il a été conçu à l'origine – dans la deuxième partie des années 1960 – était un système P2P. Le but de l'ARPANET originel était de partager des ressources de calcul sur le territoire américain. Le défi principal auquel cet effort a dû faire face a été l'intégration de différents types de réseaux existants, mais aussi des futures technologies, dans un réseau commun, qui permettrait à chaque *hôte* d'être un participant égal dans le jeu. Les quatre premiers hôtes sur ARPANET – University of California at Los Angeles, Stanford Research Institute, University of California at Santa Barbara et University of Utah – étaient déjà des sites de calcul indépendants, avec un statut égal, et ARPANET les connectait l'un à l'autre en tant que tels.

L'Internet des débuts était aussi beaucoup plus ouvert et libre que le réseau d'aujourd'hui. Les pare-feu étaient inconnus jusqu'aux années 1980, et généralement, toute machine connectée à l'Internet pouvait envoyer des paquets à une autre de son choix, alors que le réseau d'aujourd'hui est beaucoup plus cloisonné. La toile était l'« aire de jeux » de chercheurs guidés par le principe de coopération, et qui n'avaient que très peu de besoin de protection. Les protocoles et les systèmes étaient suffisamment obscurs – et suffisamment spécialisés – pour rendre les failles de sécurité rares et, pour la plupart, négligeables.

Les premières « *killer applications* »<sup>65</sup> de l'Internet, FTP et Telnet (proposés en mai 1983)<sup>66</sup>, étaient quant à elles des applications *client/serveur*. Pour échanger des communications et pour échanger des fichiers, respectivement, les clients Telnet et les clients FTP devaient se connecter à des serveurs. Mais si l'application en elle-même était client/serveur, les modèles d'usage dans leur ensemble étaient symétriques : chaque hôte sur le Net pouvait se connecter au moyen de FTP ou Telnet à tout autre hôte, et dans ces débuts des mini-ordinateurs et des unités centrales, les serveurs se comportaient souvent dans l'économie du réseau en tant que clients.

Cette symétrie fondamentale était la caractéristique qui rendait l'Internet si radicalement différent des systèmes d'interconnexion précédents, tels que la téléphonie – et à son tour, elle a donné lieu à la naissance d'une variété de systèmes plus complexes, comme Usenet et le DNS, qui se servaient de modèles de communication peer-to-peer de façon intéressante pour les développeurs d'aujourd'hui.

---

<sup>65</sup> L'étiquette « *killer application* » désigne un programme informatique si attrayant qu'il justifie à lui seul, pour de nombreux consommateurs, l'achat ou l'adoption d'un type particulier d'ordinateur, de système d'exploitation ou d'autre dispositif. En certains cas, il finit par « invalider » les concurrents existants, les contraignant à s'aligner sur ses spécificités ou à disparaître.

<sup>66</sup> Au moyen de la Request for Comments (RFC) 854 de l'IETF [<http://tools.ietf.org/html/rfc854>].

Le système de forums Usenet, qui a fait sa première apparition en 1979, implémente un modèle de contrôle décentralisé qui, à certains égards, est l'aïeul d'applications P2P beaucoup plus récentes. Le modèle de gouvernance « anarchiste » et l'organisation sociale de Usenet ont été étudiés en détail au fil des années (Hauben & Hauben, 1996 ; Pfaffenberger, 1996 ; Latzko-Toth, 2010 ; Paloque-Bergès, 2012)<sup>67</sup>. Non moins intéressantes, en termes d'influence du dispositif sur les innovations plus récentes, sont l'architecture et les procédures de fonctionnement de ce système: il s'agit en effet d'un système qui, sans contrôle central, copie des fichiers entre ordinateurs. Chaque site participant au réseau sauvegarde ses copies des messages postés dans les forums, et se synchronise périodiquement avec ses pairs.

Le système Usenet était à l'origine basé sur une installation appelée *Unix-to-Unix copy protocol* – UUCP, un mécanisme par lequel une machine Unix pouvait automatiquement en « appeler » une autre, échanger des fichiers avec elle, et puis se déconnecter. Ce mécanisme permettait à Unix d'échanger des fichiers, des emails, des mises à jour de système, ou d'autres types de messages. Usenet utilisait ce protocole pour échanger des messages rentrant dans plusieurs sujets, de façon à ce que les étudiants de *University of North Carolina* et *Duke University* puissent chacun « poster » des messages sur un sujet donné et lire des messages écrits par d'autres sur le même sujet. Usenet s'est diffusé à partir de ces deux hôtes originaux, jusqu'à connecter des centaines de milliers de sites. Au fur et à mesure que le réseau grandissait, croissaient aussi le nombre et la structuration des sujets sur lesquels un message pouvait être posté. Aujourd'hui, Usenet utilise un protocole de type TCP/IP, connu en tant que *Network News Transfer Protocol* (NNTP), introduit en 1985 en tant que solution plus flexible et adaptée à des transferts Internet informels que l'UUCP. NNTP permet à deux machines sur le réseau Usenet de découvrir de nouveaux groupes de discussion, ou groupes de nouvelles (*newsgroups*) de façon efficace, et d'échanger de nouveaux messages à l'intérieur de chacun d'entre eux ; depuis le boom de l'Internet des années 90, la presque totalité de la distribution Usenet est sur protocole NNTP.

Le modèle de base de Usenet donne une grande capacité de contrôle local, et une administration du réseau relativement simple (Latzko-Toth, 2010). Un nouvel usager Usenet se connecte au reste du monde en installant une connexion d'échange de nouvelles avec au moins un autre serveur de nouvelles sur le réseau Usenet. Usenet est donc organisé autour du principe des *newsgroups*, qui rassemblent chacun des articles (contributions) sur un sujet précis. Les sujets des *newsgroups* sont organisés selon une hiérarchie ; une fois connectés à un serveur fournissant un service Usenet, les utilisateurs peuvent choisir les groupes mis à disposition par ce serveur auxquels ils souhaitent s'« abonner ». Pour chaque groupe auquel il est abonné, l'utilisateur peut alors voir tous les nouveaux articles mis à disposition sur ce groupe (le serveur envoie une notice à l'utilisateur pour lui signaler un nouvel article, qu'il lui transmet)

---

<sup>67</sup> Je me suis aussi servi de Usenet comme exemple d'une gouvernance de réseau « anarchiste » dans mon mémoire de Master (Musiani, 2007).

et tous les articles reçus par le serveur depuis un certain temps. Lorsqu'un utilisateur, à son tour, envoie un article sur un serveur Usenet, celui-ci le diffuse à tous les autres serveurs avec lesquels il a conclu des accords d'échange d'articles (littéralement, il l'« alimente »), et ainsi de suite. Chaque serveur conserve une copie de cet article, et peut ensuite le mettre à disposition des utilisateurs ayant accès à ce serveur.

Aujourd'hui, le volume du trafic Usenet est énorme, et assurer la charge de l'ensemble des *newsgroups* ou des messages n'est certes pas à la portée de tout serveur. Les fournisseurs de service Usenet peuvent donc contrôler la taille de l'installation en spécifiant quels *newsgroups* vont être pris en charge par un serveur en particulier. De surcroît, l'administrateur peut fixer une durée de validité des articles par groupe, ou par hiérarchie, de façon à ce que les articles dans un *newsgroup* soient maintenus durant une période donnée mais pas plus. Ces contrôles permettent à chaque fournisseur de service Usenet de rejoindre le réseau, et d'y définir leur participation, en accord avec leurs propres critères. Plusieurs fournisseurs décident, par exemple, de ne pas accepter des *newsgroups* qui contiennent et transmettent des données de nature sexuellement explicite, ou illégales sous certains aspects. Il s'agit là d'une différence marquée par rapport à Freenet, qui, par choix de design, ne permet pas à un usager de voir quels types de données sa machine est en train de recevoir et de transmettre.

Usenet a influencé certains des exemples les plus réussis de structures de contrôle décentralisées sur le Net. Il n'y a pas d'autorité centrale qui gouverne le système de transmission des nouvelles, mais l'ajout de nouveaux *newsgroups* à la hiérarchie des sujets est contrôlé par un processus rigoureux, qui utilise le groupe Usenet *news.admin* pour proposer et discuter la création de nouveaux groupes (voir Pfaffenberger, 1996 et Auray, 2007). Une fois qu'un nouveau groupe est proposé, et discuté pendant une période de temps prédéterminée, toute personne dotée d'une adresse email peut soumettre un avis et demander un vote pour ou contre sa proposition. S'il obtient gain de cause, un nouveau message de groupe est envoyé et se propage à travers le réseau Usenet.

Une forme « institutionnalisée » d'anarchie, *alt.\**, est également prévue, et peut subvertir le processus *news.admin* de manière codifiée. Un *newsgroup* de type *alt.\** peut être ajouté à tout moment par tout usager, mais les sites préférant éviter *a priori* les situations problématiques qui peuvent en résulter peuvent choisir de ne pas avoir l'option. A l'intérieur de Usenet, chaque hôte participant peut prévoir ses propres politiques locales, mais le réseau dans son ensemble fonctionne grâce à la coopération et la bonne volonté de la communauté : le succès du modèle « démocratique » de ce réseau pionnier réside avec toute probabilité dans son mélange de régulation par l'organisation (procédurale) et de distribution de l'architecture (Auray, 2007).

Beaucoup des systèmes P2P émergeant au cours des dix dernières années n'ont pas encore atteint pleinement le contrôle décentralisé,<sup>68</sup> mais considèrent qu'y travailler

---

<sup>68</sup> Et de leur côté, les systèmes à la Freenet évitent délibérément de déléguer aux administrateurs locaux le contrôle sur le contenu de leurs machines car ce contrôle affaiblirait de façon importante les finalités de connexion entre politiques dissidents du système (Langley, 2001).

est une des questions primordiales à résoudre. Dans la plupart des cas, la question intéressante à aborder est : combien de contrôle (ou quel degré de contrôle ?) faut-il donner aux administrateurs locaux des applications ?

Pour certains, les systèmes modernes de P2P feraient bien de s'inspirer du protocole utilisé par Usenet. Les co-fondateurs de *Popular Power*, une start-up proposant une solution de calcul partagé, remarquent à ce propos :

NNTP, en tant que protocole, contient un certain nombre d'optimisations (...) par exemple, les messages de nouvelles maintiennent un titre qui retrace leur parcours de transmission d'un serveur de nouvelles à l'autre, appelé *path header*. Si le serveur A reçoit une requête du serveur B, et la copie du message de A liste B dans son parcours, A n'essayera pas de retransmettre ce message à B. Puisque le but de la transmission de type NNTP est d'être sûr que chaque serveur de nouvelles sur Usenet peut recevoir un article s'il le veut ; le *path header* évite une inondation de messages répétés. Gnutella, par exemple, n'a pas utilisé un tel système en transmettant les requêtes de recherche, aussi un seul nœud Gnutella peut recevoir plusieurs fois la même requête (Minar & Hedlund, 2001).

La nature ouverte et décentralisée d'Usenet a, par ailleurs, des inconvénients aussi bien que des avantages. Usenet a représenté un des plus grands succès du premier Internet, dans la mesure où il a survécu depuis 1979 et qu'il continue d'héberger des communautés d'experts florissantes, ayant largement dépassé ses débuts modestes. Pourtant, sous plusieurs aspects, l'éthos de confiance et la nature décentralisée du protocole en ont parfois réduit l'utilité et l'ont relégué au rang de canal de communication trop bruyant. En particulier, comme on le verra, Usenet a été une victime privilégiée du spam dans les premières phases de développement de l'Internet commercial. Cependant,

son système de contrôle décentralisé, les méthodes que ce réseau déploie pour éviter la saturation du réseau, et quelques autres caractéristiques en font un excellent objet de leçon pour nous les concepteurs des nouveaux systèmes P2P<sup>69</sup>.

#### *DOMAIN NAME SYSTEM, L'HYBRIDE SOUPLE*

Le Domain Name System (DNS, système de noms de domaine) est un autre exemple de système qui combine un type d'interconnexion pair à pair avec un modèle hiérarchique de propriété de l'information. Pour le développeur d'outils de type P2P, ce qui paraît le plus remarquable dans le DNS est sa capacité de passer à l'échelle, des quelques milliers de hôtes qu'il était destiné à héberger en 1983 jusqu'aux centaines de millions qui forment désormais l'Internet. Les leçons qui dérivent du

---

<sup>69</sup> Entretien avec Olivier Auber, chercheur/entrepreneur de l'Internet et fondateur du laboratoire Anoptique, 16 février 2009.

DNS sont directement applicables aux applications contemporaines en P2P, surtout celles destinées au partage de données.

Le DNS a été établi, à l'origine, en tant que solution à un problème de partage de fichiers. A l'aube de l'Internet, la façon d'associer une étiquette plus facilement lisible à une adresse IP était de se servir d'un seul fichier, *hosts.txt*, qui était copié à tout le réseau Internet périodiquement. Au fur et à mesure que la toile se peuplait de milliers d'hôtes, et que la gestion de ce fichier devenait impossible, le DNS a été développé, pour distribuer le partage des données sur l'ensemble de l'Internet, à ce moment encore une structure égalitaire.

L'espace de nommage du DNS est, par ailleurs, « naturellement » hiérarchique : une hiérarchie intégrée au système à son niveau le plus profond. Il s'agit d'un mécanisme de délégation de responsabilité dans la gestion des différentes parties de la base de données DNS. Chaque domaine a une autorité, le serveur qui enregistre les hôtes pour cette base de données. Quand un hôte sur l'Internet veut connaître l'adresse d'un nom en particulier, il le demande au serveur de nommage le plus proche. Si ce serveur ne le connaît pas, il délègue la requête à l'autorité pour l'espace de nommage. Cette requête peut, à son tour, être déléguée à une autorité supérieure, et le processus peut se répéter jusqu'aux serveurs racine pour l'Internet dans son ensemble. Pendant que la réponse est renvoyée vers le demandeur, le résultat est mis en mémoire (cache) aux différents stades du parcours, pour que la requête suivante puisse être traitée plus efficacement. Les serveurs de nommage opèrent soit en tant que clients, soit en tant que serveurs.

Dans son ensemble, le DNS fonctionne correctement aujourd'hui, d'une façon presque surprenante au vu de son passage à l'échelle très importante (presque 10 000 fois sa taille originelle).<sup>70</sup> Plusieurs éléments fondamentaux de son design sont susceptibles d'être répliqués dans les systèmes distribués contemporains :

Premier élément : les hôtes peuvent opérer soit comme clients, soit comme serveurs, en propageant si besoin les requêtes. Ces hôtes facilitent aussi le passage à l'échelle en mettant les réponses en cache. Deuxième élément : la manière de propager les requêtes de données à travers le réseau. Tout serveur DNS peut en interroger un autre, mais normalement, il y a un parcours standard dans la chaîne d'autorité. La charge est distribuée sur l'ensemble du réseau DNS, pour que tout serveur individuel ne doive servir que les nécessités de ses clients et l'espace de nommage qu'il gère individuellement<sup>71</sup>.

---

<sup>70</sup> Par ailleurs, l'« ordonnancement » d'un système qui n'avait pas prévu, à l'origine, d'ordre politique a rendu nécessaire la mise en place d'un système de régulation des disputes, *Uniform Dispute Resolution Policy (UDRP)*, organisé sur le modèle du droit des marques, après l'invention du système lui-même. Selon Cécile Méadel et Meryem Marzouki, l'imposition de l'ordre économique-politique lui s'est mis en place avec l'UDRP a « redoublé » les aspects hiérarchiques inscrits dans l'architecture (Méadel & Marzouki, 2005). Voir [Hancock, 2001] ; [Mueller, 2001, 2002] ; je fais une analyse procédurale de l'UDRP dans [Musiani, 2009, 47-49].

<sup>71</sup> Entretien avec Finn Brunton, développeur et enseignant à la New York University, 5 septembre 2010.

Le succès du DNS, en tant qu'hybride doté d'une capacité d'évolution souple, a donc été lié à l'introduction d'une dynamique distribuée dans un système conçu de manière totalement hiérarchique.

Ainsi, depuis ses premiers stades de développement, l'Internet a été construit en se basant sur des modèles de communication en pair à pair. Cela amène les développeurs à regarder vers ces premières expériences pour leur design de nouveaux systèmes P2P, puisque les questions et problèmes à affronter dans les nouvelles applications ont beaucoup en commun avec les problèmes abordés par Usenet et le DNS il y a vingt ans. Il s'agit, comme on l'a vu, de dispositifs avec des spécifications distribuées, articulées avec des formats de régulation à la fois technologiques et politiques – d'un côté, Usenet, une gestion interne de conflits avec un lourd appareillage de consultation, et de l'autre, un alignement sur le droit des marques.

## SECTION 2. LA DIFFUSION MONDIALE D'INTERNET ET SES CONSÉQUENCES SUR L'ARCHITECTURE DU RÉSEAU

En 1994, l'explosion commerciale de l'Internet en change radicalement la forme, transformant le réseau des réseaux, tranquille utopie d'intellectuels passionnés, en agité médium de masse. Des millions de nouveaux utilisateurs font leur apparition sur le Net, une vague représentant des nouveaux profils d'internautes : les *ordinary people* intéressés par la facilité avec laquelle on peut entrer en contact, visiter des pages d'information, et acheter en ligne, plutôt que par les secrets et les détails de structuration de réseaux complexes. La transformation de l'Internet en un phénomène de culture de masse entraîne bientôt des changements d'une portée considérable dans l'architecture du réseau, un impact qui n'est pas sans conséquences sur la création d'applications P2P dans l'Internet d'aujourd'hui. Ces changements s'explicitent notamment dans la façon dont le réseau est utilisé, dans la rupture du principe de coopération sur le réseau, le déploiement de pare-feu, et l'augmentation de liens de réseau asymétriques tels que l'ADSL et les câbles Ethernet.

### LE PASSAGE AU MODÈLE CLIENT/SERVEUR

Le modèle de réseau sous-tendant les applications destinées à l'utilisateur final – non seulement en termes de consommation de bande passante, mais aussi en termes de contact, relation et communication entre machines – change radicalement avec l'avènement de l'Internet commercial et la connexion de millions d'utilisateurs « domestiques » des années 1990. Les protocoles de modem deviennent toujours plus populaires<sup>72</sup>, les applications sont de plus en plus programmées pour des modems et des lignes analogiques à basse vitesse, et les entreprises commencent à doter leurs réseaux de pare-feu et de *Network Address Translation*<sup>73</sup>. Beaucoup de ces changements reflètent les priorités d'usage communes à ce moment historique particulier ; la plupart concernent le téléchargement de données plutôt que la publication d'informations ou leur téléchargement vers l'amont.

Le navigateur Web, et plusieurs autres applications qui naissent dans les premières phases de la commercialisation de l'Internet, se basent sur un simple protocole client/serveur : le client amorce une connexion à un serveur connu, en télécharge des

---

<sup>72</sup> Par exemple SLIP (*Serial Line Internet Protocol*, ou *protocole Internet de liaison en série*, protocole de liaison Internet simple n'effectuant ni contrôle d'adresse, ni contrôle d'erreur, permettant l'envoi d'une trame composée uniquement des données à envoyer suivies d'un caractère de fin de transmission) et PPP (*Point to Point Protocol*, ou *protocole point à point*, il autorise un adressage des machines et est doté d'un contrôle d'erreur). Voir Internet Engineering Task Force, RFC 1661, Point to Point Protocol, <http://abcdrfc.free.fr/rfc-vf/rfc1661.html>.

<sup>73</sup> Un routeur fait du *Network Address Translation* (NAT) (« traduction d'adresse réseau ») lorsqu'il fait correspondre les adresses IP internes non uniques et souvent non routables d'un intranet à un ensemble d'adresses externes uniques et routables. Ce mécanisme permet notamment de faire correspondre une seule adresse externe publique visible sur Internet à un ensemble d'adresses d'un réseau privé.

données, et se déconnecte. Quand l'utilisateur a terminé ce téléchargement et que les données sont collectées, le processus se répète. Ce modèle s'est avéré une façon simple et directe de mettre en place nombre d'applications qui impliquaient un service à l'utilisateur, de la navigation du Web, au visionnage des vidéos en *streaming*. De leur côté, les développeurs commencent à y mettre des chariots de courses, des transactions de stocks, des jeux interactifs, et beaucoup d'autres « biens ». Les machines qui hébergent un client web ne nécessitent pas, elles, une adresse reconnue ni permanente. Elles n'ont pas besoin d'une connexion permanente à l'Internet, ni de gérer les besoins de plusieurs usagers. Elles doivent juste « savoir comment poser une question, et comment écouter et comprendre la réponse »<sup>74</sup>.

Cependant, toutes les applications utilisées au niveau domestique ne répondent pas à ce modèle. La messagerie électronique, par exemple, repose beaucoup plus sur une communication à double sens entre un serveur et un client email ; par ailleurs, dans ce cas, le client est souvent en train de « parler » à un serveur sur le réseau local (soit le serveur du FAI, soit le serveur d'une entreprise). Les systèmes de *chat* tel que l'Instant Messenger de AOL, largement diffusés, ainsi que les systèmes Usenet, ont des propriétés « locales » semblables. En conséquence, les instructions de configuration typiques d'un FAI pour le *chat*, l'email ou le partage de nouvelles sont détaillées, mais elles n'ont pas besoin d'être comprises par l'utilisateur pour les usages communs. Mises à part quelques exceptions qui non seulement exigeaient une configuration manuelle de la part de l'utilisateur, mais « méritaient aussi son effort »<sup>75</sup>, le modèle de téléchargement à sens unique est à ce stade plus simple, bien que souvent moins transparent, et fonctionne sans que l'utilisateur ait besoin de s'investir dans le processus de configuration ; le modèle à double sens est utilisé moins fréquemment.

Les premières visions du Web traçaient un portrait de grand « égalisateur de communications » – un système permettant à tout usager de publier son point de vue au lieu d'être purement et simplement un consommateur de médias. L'explosion commerciale de l'Internet dirige rapidement une grande majorité de son trafic vers le paradigme « *downstream* » propre à la télévision et aux médias traditionnels ; même si la fabrication de contenus *ad hoc*, liée à la spécificité de chaque requête et de chaque contribution des utilisateurs, reste une première évolution par rapport à ce modèle – une évolution destinée à s'inscrire dans la durée.

## LA RUPTURE DU PARADIGME DE COOPÉRATION

Le premier Internet a été dessiné en accord avec des principes de coopération et avec la « bonne ingénierie » qui donnent la priorité aux questions d'efficacité et d'optimisation technique du dispositif. Toute personne travaillant sur l'Internet avait un même but : construire un réseau fiable, efficace et puissant. L'entrée de l'Internet dans sa phase commerciale change les équilibres, les avantages et les structures ; elle

---

<sup>74</sup> Entretien avec G. Simon, 19 mars 2009, cité.

<sup>75</sup> Entretien avec O. Auber, 16 février 2009, cité.

provoque nombre de points de stress et de faiblesse. Ceux-ci rendraient plus évidente la susceptibilité à la « tragédie des communs » (Ostrom, 1999), le dilemme découlant de situations dans lesquelles plusieurs individus, agissant de manière indépendante et consultant rationnellement leur propre intérêt, finissent par épuiser une ressource partagée limitée, même quand il est clair que ce n'est l'intérêt à long terme de personne que cela se produise. Ce phénomène s'est actualisé de plusieurs façons, notamment dans l'augmentation importante de spam circulant sur le réseau et dans les défis rencontrés dans la construction de protocoles de réseau efficaces, pouvant gérer les ressources nécessaires à son fonctionnement dans une perspective de bien commun.

#### *DES AGENTS NON COOPÉRATIFS : LE SPAM*

Les spam, ou messages électroniques commerciaux non sollicités, sont aujourd'hui un phénomène omniprésent sur l'Internet. Au cours de la phase pré-commerciale du réseau, la publicité non sollicitée était généralement reçue avec surprise et indignation. La « fin de l'innocence » (Minar & Hedlund, 2001) a lieu le 12 avril 1994, le jour où les avocats Laurence Canter et Martha Siegel postent le premier spam commercial de masse sur Usenet, le « *Green Card Spam* » : il s'agit d'une publicité postée individuellement sur tous les groupes de discussion Usenet, qui atteint le monde entier avec un message publicisant les services de la firme de Canter et Siegel<sup>76</sup>. À l'époque, ce genre d'action est sans précédent et engendre une forte désapprobation. Non seulement la plupart du public potentiel touché n'est pas intéressé par le service, mais beaucoup de gens estiment que Canter et Siegel ont, *de facto*, volé les ressources de Usenet. Les deux avocats ne sont pas ceux qui paient pour la transmission de la publicité : les coûts sont, plutôt, pris en charge par Usenet dans son ensemble.

Dans l'Internet contemporain, le spam n'apparaît pas comme surprenant ; Usenet y a largement cédé, et les FAI fournissent maintenant des services de filtrage de spam pour le courrier électronique de leurs utilisateurs, à la fois pour aider ceux-ci et pour défendre leurs propres infrastructures. La messagerie électronique et Usenet comptaient sur la coopération des individus, et sur leur volonté individuelle et collective de ne pas inonder les ressources communes avec des courriers publicitaires indésirables : c'est ce principe de coopération qui tombe en panne. A ce jour, l'Internet ne s'est pas doté d'une technologie vraiment efficace pour empêcher le spam.

Dans ce cas, le manque d'attribution de responsabilité dans l'architecture de l'Internet devient un problème, et pose des nouvelles questions. Puisque n'importe quel hôte peut se connecter à tout autre hôte, et que les connexions sont presque anonymes, les usagers peuvent insérer du spam dans le réseau à tout moment. Commence alors une sorte de « course aux armements »<sup>77</sup> pour essayer de responsabiliser les utilisateurs :

---

<sup>76</sup> Voir [Everett-Church, 1999], pour plus de détails.

<sup>77</sup> Entretien avec un cadre d'une entreprise de logiciels de sécurité, 2 octobre 2009, Open World

la fermeture des relais ouverts pour l'envoi de messages, le suivi des sources de spam sur Usenet, les représailles contre les spammeurs. Mais la bataille est perdue depuis le début : aujourd'hui, le spam est un phénomène avec lequel les utilisateurs (et les dispositifs de communication eux-mêmes) ont, *de facto*, appris à vivre.

Les concepteurs de P2P voient une leçon importante dans la naissance du spam : sans l'attribution de responsabilité au sein du réseau, il est difficile d'implémenter les règles de la responsabilité sociale.

Tout comme Usenet et le courrier électronique, les systèmes P2P actuels courent le risque d'être envahis par des publicités non sollicitées. Il est difficile de concevoir un système dans lequel tout usage socialement inapproprié peut être complètement prévenu. Des technologies de responsabilité et responsabilisation, telles que les systèmes de réputation ou l'identification chiffrée, peuvent être des outils précieux pour aider à gérer un réseau P2P<sup>78</sup>.

Il est important, renchérit un autre développeur, d'implémenter et renforcer ces capacités depuis le début<sup>79</sup> : des propositions qu'il y a eu jusqu'ici pour mettre à jour Usenet et la messagerie électronique, aucune n'a été largement diffusée.

#### *DES PROTOCOLES COOPÉRATIFS : TCP ET LE « BEST EFFORT »*

Un principe fondamental de conception de l'Internet est le *best effort*, le « meilleur effort » de livraison de paquets : cela signifie que la composition et la structure du réseau Internet ne garantit pas qu'un paquet va rejoindre sa destination, mais tout simplement que le réseau « fera de son mieux » pour que cela arrive.

Les protocoles de niveau supérieur, tel TCP, créent des connexions fiables en détectant quand un paquet est perdu et en le renvoyant. Une des raisons principales pour laquelle les paquets ne sont pas livrés sur l'Internet est l'encombrement : si un routeur du réseau est débordé, il commence à supprimer des paquets au hasard. TCP y remédie en modifiant la vitesse à laquelle les données sont envoyées. Lorsque le réseau est encombré, chaque connexion TCP ralentit de façon indépendante, cherchant à trouver le taux optimal d'échange tout en ne perdant pas trop de paquets. Par ailleurs, non seulement les connexions TCP individuelles optimisent leur utilisation de bande passante, mais le protocole est également conçu pour permettre à l'Internet dans son ensemble de fonctionner efficacement. Le comportement collectif de nombre de connexions TCP individuelles qui ralentissent de façon indépendante entraîne une diminution de la saturation au niveau du routeur ; l'algorithme TCP est donc, en fait, une façon pour les différents pairs de gérer une ressource partagée, sans

---

Forum, Paris.

<sup>78</sup> Entretien avec O. Auber, 16 février 2009, cité.

<sup>79</sup> Entretien avec Alban Galland, développeur et doctorant à l'ENS Cachan, 15 octobre 2009.

coordinateur central.

Le problème est que l'efficacité du protocole TCP à l'échelle de l'Internet dans son ensemble requiert, fondamentalement, de la coopération : à chaque utilisateur du réseau, il est demandé de jouer avec les mêmes règles. Comme explique un développeur :

Les performances d'une connexion individuelle au moyen de TCP sont inversement proportionnelles à la racine carrée du taux de perte de paquets – ladite « *TCP rate equation* », loi fondamentale pour une bonne gouvernance de l'Internet. Les protocoles qui suivent cette loi sont connus comme les protocoles « *TCP-friendly* ». Il est possible de concevoir d'autres protocoles qui ne suivent pas cette équation, et qui vont rudement essayer de consommer plus de bande passante qu'ils ne devraient. Ces protocoles peuvent causer des ravages sur le Net.<sup>80</sup>

Ces « ravages » seraient dus non seulement à l'utilisation d'une quantité trop importante et non équitable de ressources, mais aussi à l'endommagement d'une ressource commune pour tous – un problème de réseautage qui a été cité comme exemple classique de la tragédie des communs (Ostrom, 1999), à laquelle l'Internet d'aujourd'hui est tout à fait vulnérable.

Le problème est à la fois théorique et très pratique : ces protocoles ont été élaborés au cours de ces dernières années par des entreprises aux exigences commerciales, tandis qu'en parallèle, d'autres composantes du secteur IT, le monde de l'*open source* en tout premier lieu, ont exprimé l'inquiétude croissante que les protocoles hostiles commencent à nuire à l'Internet.

Un des premiers exemples est la fonctionnalité ajoutée par Netscape à son navigateur qui donne la possibilité de télécharger plusieurs fichiers en même temps. En 1996, les ingénieurs de Netscape découvrent que si on télécharge des images intégrées en parallèle, plutôt qu'une seule à la fois, la page entière se télécharge plus rapidement, ce que les utilisateurs apprécient. Mais une question se pose alors : cette utilisation de bande passante est-elle équitable? Non seulement on force le serveur à envoyer plusieurs images simultanément, mais on multiplie les canaux TCP, et on contourne les algorithmes TCP contre la saturation.

Une controverse s'installe lorsque Netscape introduit cette fonction. Elle s'apaisera une fois que, Netscape ayant rendu publiques les caractéristiques de son navigateur, on découvre dans la pratique que la stratégie de téléchargement en parallèle ne porte pas indûment préjudice à l'Internet. Aujourd'hui, cette technique est la norme pour tous les navigateurs et demeure incontestée, mais le même type de questionnements s'est posé pour les programmes visant à accélérer le téléchargement simultané des différents morceaux d'un même fichier, menaçant à nouveau de bouleverser la gestion délicate du réseau Internet afin d'en éviter la saturation.

La diffusion du spam et le fonctionnement de l'algorithme TCP partagent une même

---

<sup>80</sup> Entretien avec un cadre d'une entreprise de logiciels de sécurité, 2 octobre 2009.

caractéristique : ils démontrent la fragilité du bon fonctionnement durable de l'Internet, et la nécessité de coopération que celui-ci implique. Dans le cas du protocole TCP, le système a résisté, et le réseau a été préservé. Dans le cas de spam, cependant,

« ...la bataille a été perdue et un comportement asocial et non coopératif est avec nous pour toujours. La leçon pour les concepteurs de systèmes est donc d'examiner la question de la 'politesse' entre usagers en amont<sup>81</sup>. »

Une alternative se pose donc aux développeurs : ou bien il est nécessaire de concevoir des systèmes qui n'ont pas besoin de la coopération comme principe fondateur de leur bon fonctionnement, ou bien, plus vraisemblablement, il faut créer des incitations à la coopération, en récompensant les comportements responsables ou en implémentant des mécanismes de vérification, afin que les initiatives déséquilibrées ou nuisibles puissent être punies.

#### LA FIN DU RÉSEAU OUVERT

En même temps que la nature coopérative de l'Internet se voit menacée, les administrateurs de réseau mettent en place diverses mesures de gestion, qui finissent par se traduire dans un Internet beaucoup moins ouvert. Notamment, la symétrie du réseau – le principe selon lequel si un hôte peut accéder au réseau, tout le monde sur le réseau peut atteindre ce hôte – commence à être érodée au milieu des années 1990 par trois tendances controversées : le déploiement toujours plus fréquent de pare-feu, l'augmentation des adresses IP dynamiques, et la popularité de la *Network Address Translation* (ou « traduction d'adresse réseau », NAT).

L'Internet devient un système plus mûr, et avec sa maturation vient la nécessité de sécuriser le réseau, pour protéger chaque hôte individuel d'un accès illimité par les autres. Les utilisateurs moyens ne pouvant pas gérer de manière autonome les risques de sécurité pour leurs machines qui résultent d'une conception symétrique des réseaux, leurs gestionnaires se tournent vers les pare-feux en tant qu'outils permettant de contrôler l'accès à ces machines. Les pare-feux se trouvent au point de contact entre le réseau interne et l'Internet à l'extérieur. Ils filtrent les paquets, et choisissent quel trafic laisser passer et à qui refuser l'accès. Le pare-feu met à l'épreuve le modèle fondamental de l'Internet : certaines parties du réseau ne peuvent plus parler de façon totalement libre à d'autres parties. Ainsi, un outil de sécurité très utile, les pare-feux, pose-t-il un premier obstacle sérieux aux modèles de communication en peer-to-peer.

Le fonctionnement typique du pare-feu est de permettre à quiconque est situé à l'intérieur d'un réseau d'amorcer et établir une connexion avec tout autre internaute, mais d'empêcher des hôtes au hasard sur l'Internet d'initier des connexions avec des hôtes sur le réseau interne – « une porte à sens unique par laquelle on peut sortir, mais

---

<sup>81</sup> Entretien avec un cadre d'une entreprise de logiciels de sécurité, 2 octobre 2009.

pas entrer. »<sup>82</sup> En outre, les connexions en sortie peuvent être limitées exclusivement à certaines applications, comme FTP et le Web, et bloquer le trafic d'autres portes au niveau du pare-feu. Un hôte protégé de cette façon ne peut plus facilement fonctionner comme un serveur : il ne peut être qu'un client.

Permettre à un hôte sur l'Internet de n'être qu'un client, et pas un serveur, est un thème transversal à beaucoup de changements de l'Internet après son explosion commerciale. Avec l'augmentation du nombre d'utilisateurs dotés d'une connexion à l'Internet par modem, la pratique de donner à chaque hôte une adresse IP fixe devient impraticable, le nombre d'adresses IP n'étant plus suffisant. L'affectation dynamique des adresses IP prend pied jusqu'à devenir la norme pour de nombreux hôtes sur l'Internet, où l'adresse d'un ordinateur particulier peut maintenant changer même une fois par jour. Les fournisseurs de bande passante, de leur côté, trouvent les adresses IP dynamiques utiles pour le déploiement de services toujours disponibles. Le résultat final est que de nombreux hôtes sur l'Internet, se déplaçant constamment, ne sont pas facilement accessibles. Par rapport à l'Internet des débuts où les hôtes étaient statiques, les choses se compliquent : les applications P2P servant des buts tels que la messagerie instantanée ou le partage de fichiers, pour contourner ce problème, doivent désormais construire des répertoires dynamiques de hôtes.

Cette tendance s'accroît ultérieurement quand on commence à ne plus attribuer une adresse Internet publique et valide à un hôte, mais à utiliser la NAT pour cacher cette adresse derrière un pare-feu. Par rapport au type de trafic peer-to-peer, la NAT allie les problèmes des pare-feu et ceux des adresses IP dynamiques : la véritable adresse de l'hôte n'est pas seulement instable, elle n'est plus accessible. Toute communication doit parler un langage relativement simple, que le routeur NAT puisse comprendre, ce qui entraîne une grande perte de flexibilité dans les communications entre applications. Par exemple,

de nombreux jeux coopératifs sur l'Internet ont eu des problèmes avec la NAT (...) chaque joueur veut être en mesure de contacter tous les autres joueurs, mais les paquets ne peuvent pas passer par le routeur NAT. Le résultat est que un serveur central doit être établi pour agir en tant que routeur de messages au niveau de l'application, en servant la même fonction que le protocole TCP/IP servait avant<sup>83</sup>.

Les pare-feux, les IP dynamiques, et la NAT sont nés d'un besoin de sécuriser l'architecture Internet au fur et à mesure que celle-ci évoluait et se peuplait d'utilisateurs, et ils ont sans aucune espèce de doute contribué de manière déterminante à amener des millions d'ordinateurs clients sur l'Internet, de façon rapide et souple. Ces mêmes technologies ont, par ailleurs, profondément reconfiguré les équilibres de l'infrastructure Internet dans son ensemble, en « reléguant » la plupart des ordinateurs au seul statut de clients. Architecture que les nouvelles applications P2P mettent, une fois de plus, en discussion.

---

<sup>82</sup> Entretien avec G. Simon, cité.

<sup>83</sup> Entretien téléphonique avec un développeur de jeux en ligne, 16 décembre 2009.

Une autre tendance de l'Internet des années 1990 qui pose problème aux applications P2P est la montée en puissance de connexions asymétriques au réseau, telles que l'ADSL ou les modems câblés. Afin de maximiser l'efficacité du câblage disponible, les fournisseurs de services à large bande choisissent de fournir une bande passante asymétrique. L'installation domestique typique d'une ligne ADSL ou d'un modem câblé est conçue de façon à ce qu'elle offre trois à huit fois plus de bande passante en téléchargeant des données depuis l'Internet qu'en envoyant des données vers le réseau, favorisant ainsi des usages de type client plutôt que de type serveur.

La raison pour laquelle ce déséquilibre entre téléchargements vers l'amont et vers l'aval est largement tolérée par le public est liée au rôle de « *killer application* » de l'Internet que le Web a occupé pendant ces années : la plupart des usagers du Web n'ont que très rarement besoin d'être plus qu'un client. Même les usagers qui publient leurs propres pages Web ne le font pas, généralement, depuis leur connexion domestique à bande large, mais depuis des serveurs dédiés appartenant à des tiers, tels IT GeoCities ou Exodus. Si dans les premiers jours d'existence du Web la situation était moins claire – la possibilité que chaque usager se dote d'un serveur web personnel n'était pas écartée *a priori* – il devint bientôt clair que le Web commercial comportait en soi des éléments d'asymétrie – beaucoup de clients pour peu de serveurs – et que la plupart des usagers étaient, dans ce contexte, bien servis par l'asymétrie de la bande passante.

Les applications P2P mettent à nouveau et radicalement en débat l'approche selon laquelle les usagers finaux exécutent presque exclusivement des opérations de téléchargement vers l'aval, et non vers l'amont. Déjà les premières applications de partage de fichiers, telles Napster ou la première version de Gnutella, renversent ce modèle d'utilisation de bande passante : les terminaux servent souvent beaucoup plus de fichiers qu'ils n'en téléchargent. Le canal de téléchargement vers l'amont ne peut souvent pas satisfaire la demande, ce qui, à cause des détails de fonctionnement du protocole TCP, entraîne également des répercussions sur les bonnes performances du canal de téléchargement vers l'aval.

L'architecture du réseau Internet, au début des années 2000, reflète cette asymétrie : les connexions ADSL, au moyen de modems câblés (et plus généralement tout réseau géré par un FAI) sont conçues pour que le flux des bits se dirige vers l'utilisateur (Taylor & Harrison, 2001 : 6). Cette assomption d'asymétrie est remise en discussion par la technologie P2P ; par ailleurs, il y a très peu de solution « directe » qu'une application P2P à elle seule puisse mettre en œuvre pour la contourner.

### SECTION 3. ENTRE ANCIENS MODÈLES ET INNOVATION : LA « RENAISSANCE » DU P2P

Le début des années 2000 voit l'avènement de nouveaux dispositifs qui se revendiquent du P2P : ceux-ci profitent à la fois des leçons fournies par les premiers modèles de l'Internet, et de la mise en place de caractéristiques novatrices. Ils permettent de nuancer et de séparer les concepts de création de l'information et de publication ; ils ouvrent la voie à des nouvelles organisations du trafic Internet, en soumettant les pare-feu à des efforts jusque là inédits ; ils expérimentent avec la conception d'applications décentralisées, ce qui pose à la fois des défis nouveaux et ouvre des nouvelles possibilités.

#### CRÉATION ET PUBLICATION DE L'INFORMATION

Une des promesses de l'Internet est depuis le début que ses usagers deviennent capables de publier les informations dont ils sont les créateurs, par exemple en utilisant des sites web personnels pour révéler au public leurs points de vue et leurs intérêts. Les pratiques d'auto-publication deviennent certes plus populaires avec l'avènement de l'Internet commercial ; pourtant les usagers passent, pour l'essentiel la plupart de leur temps en ligne à lire (et télécharger) de l'information, et moins à en publier (Minar & Hedlund, 2001 : 15-16). Comme on l'a vu, les fournisseurs de services et d'accès Internet construisent leur offre en s'appuyant sur cette asymétrie.

Les applications P2P qui jouissent d'un succès considérable au début des années 2000 sont porteurs d'une réalité intermédiaire entre le supposé cas idéal des origines, « tout le monde publie », et l'apparente réalité de l'Internet commercial, « tout le monde consomme ». Ces applications, Napster en particulier, rendent très facile la publication de données dont on n'est pas l'auteur, tandis que les machines des utilisateurs sont utilisées comme répéteurs pour la retransmission des données qu'elles reçoivent. Le design du réseau qui repose sur l'existence d'un nombre limité d'auteurs, et l'asymétrie de bande passante comme principe d'optimisation, se voient profondément remis en discussion par ce développement. Un nombre important de réseaux, en premier lieu les systèmes des campus et des petites entreprises, se trouvent surchargés.

Les systèmes P2P se voient rapidement bannis de ces réseaux, parce qu'ils changent *de facto* la direction et l'importance des flux de données, et ainsi faisant, ils prouvent la fausseté d'une assomption qui a guidé le développement de l'Internet des années 1990. Les serveurs ne sont pas tous placés sous le contrôle des éditeurs de contenus ; les éditeurs et les auteurs des contenus ne correspondent pas toujours à un même emplacement sur le réseau. Les usagers n'ont pas besoin de créer du contenu pour vouloir et pouvoir le publier (Dauphin & Dagiral, 2005) ; la faisabilité de la publication de masse par les « lecteurs » est démontrée par la grande échelle que certains de ces systèmes sont capables d'atteindre.

## LES DÉFIS POUR LA « SÉCURISATION » DES RÉSEAUX

Les outils de type P2P posent aussi des nouvelles questions et problèmes pour les outils de sécurisation du premier Internet commercial. En particulier, ils soumettent à des nouveaux efforts le « port 80 »<sup>84</sup>, que le trafic http utilise lorsque les utilisateurs naviguent sur le Web (Minar & Hedlund, 2001). Sauf exception, les pare-feux filtrent le trafic en se basant sur sa direction (en entrée ou en sortie) et son port de destination. Comme le Web est l'application principale de nombreux utilisateurs Internet, presque tous les pare-feu autorisent les connexions en sortie sur le port 80, même si leur politique est par ailleurs très restrictive.

Dans les premières années de l'Internet, le numéro de port indiquait habituellement quel type d'application était en train d'utiliser le réseau : « le pare-feu pouvait compter sur le fait que le port 80 est uniquement destiné au trafic Web »<sup>85</sup>. Mais précisément parce que de nombreux pare-feu autorisent les connexions au port 80, les auteurs d'autres types d'applications commencent à acheminer le trafic de ces applications à travers ce port. Streaming audio, messagerie instantanée sont envoyés au moyen du port 80, et la plupart des applications P2P destinées au partage de fichiers commencent à faire de même, contournant ainsi les politiques de sécurisation des réseaux : les pare-feu sont programmés de façon à ne pas se rendre compte du fait qu'ils sont en train de permettre l'accès à un trafic que l'administrateur du réseau désire précisément bloquer.

Cela pose deux points de tension. Tout d'abord, puisque le numéro de port a été contourné, il n'y a pas de moyen efficace pour un pare-feu d'identifier les applications qui le traversent. Une nouvelle génération de pare-feux commence donc à voir le jour, qui permet d'analyser le trafic réel qui le traverse pour voir s'il s'agit d'un flux http légitime ; cette dynamique encourage cependant, à son tour, les concepteurs d'applications à masquer le trafic de celles-ci en http, donnant ainsi lieu à une « course aux armements qui ne profite à personne »<sup>86</sup>. Un deuxième point ouvert est que, même si une application a une raison légitime de passer par le pare-feu, il n'y a pas de moyen simple pour l'application de demander l'autorisation de le faire. Le pare-feu, en tant que mesure de sécurisation du réseau, est dépassé : tant qu'un pare-feu permet à un type de trafic quelconque de s'y glisser, les applications P2P trouvent un moyen de le faire.

---

<sup>84</sup> Un port logiciel (identifié au moyen d'un numéro) permet, sur un ordinateur donné, de distinguer différents « interlocuteurs », c'est-à-dire des programmes informatiques qui, selon les cas, écoutent ou émettent des informations sur ces ports.

<sup>85</sup> Entretien avec G. Simon, 19 mars 2009, cité.

<sup>86</sup> Id.

## QUELLE DÉCENTRALISATION ?

Ainsi qu'on l'a introduit dans le chapitre précédent, les systèmes P2P ne sont pas synonymes de décentralisation, bien qu'ils y soient étroitement liés. Dans un système totalement décentralisé, non seulement tous les hôtes sont des partenaires égaux, mais il n'y a pas d'hôtes avec des rôles particuliers, de facilitateur ou d'administrateur. Dans la pratique, la construction de systèmes entièrement décentralisés, ainsi que leur préservation dans le temps, présentent nombre de défis techniques ; de nombreuses applications P2P adoptent donc des approches hybrides pour résoudre ces problèmes. Les exemples qu'on a présentés au cours de ce chapitre et du précédent en sont des exemples : le DNS est décentralisé dans la conception du protocole qui le gouverne, mais un principe organisationnel hiérarchique y est profondément inscrit. Il existe de nombreux autres exemples de systèmes qui sont en première instance décentralisés, mais dont l'organisation est semi-centralisée au niveau de l'application, tels que Usenet et plusieurs outils de messagerie instantanée ainsi que de partage de fichiers.

Usenet, en particulier, est un exemple éclairant de l'évolution d'un système décentralisé. La propagation des nouvelles par Usenet est symétrique ; il s'agit d'hôtes qui partagent du trafic. Par ailleurs, puisque les coûts pour maintenir une *news feed* complète sont élevés, il existe dans la pratique une ossature de nœuds « primaires » qui se chargent de tout le trafic et le servent à un grand nombre de *leaf nodes*, ou « nœuds/feuilles »<sup>87</sup> dont le rôle est surtout de recevoir des articles. Co-existent donc dans Usenet une tendance à la propagation du trafic de façon hiérarchique, même si les protocoles sous-jacents ne l'exigent pas. Cette forme de centralisation « molle » s'avère viable et économique pour plusieurs des nouveaux systèmes P2P, avec des systèmes de transmission de données à coût élevé.

Beaucoup d'autres applications P2P aux débuts des années 2000 présentent une facette décentralisée, tout en s'appuyant sur un facilitateur central pour coordonner les opérations :

Aux yeux de l'utilisateur d'un système de messagerie instantanée, l'application dont il se sert semble de type peer-to-peer, envoyant les données directement à l'ami qui est leur destinataire. Mais tous les principaux systèmes de messagerie instantanée sous-tendent un serveur qui facilite la communication entre les nœuds.<sup>88</sup>

Les développeurs introduisent l'« objet serveur » dans leur projets pour différents buts : celui-ci peut gérer l'association entre le nom de l'utilisateur et son adresse IP courante, stocker temporairement les messages si l'utilisateur est déconnecté, et

---

<sup>87</sup> En informatique, un arbre est une structure de données qui émule une arborescence hiérarchique avec un ensemble de nœuds reliés. Un nœud est une structure qui peut contenir une valeur, une condition, ou représenter une structure de données distincte (qui peut être un arbre elle-même). Chaque nœud dans un arbre a zéro ou plusieurs nœuds-fils (*child nodes*). Un nœud qui a un descendant est appelé son nœud parent, tandis que les nœuds qui n'ont pas de descendants sont appelés **nœuds feuilles**. Ils sont également appelés nœuds terminaux.

<sup>88</sup> Entretien avec O. Auber, 16 février 2009, cité.

acheminer les messages aux utilisateurs qui se protègent d'un pare-feu ; ou encore, dans le cas des systèmes qui permettent une communication directe entre clients lorsque cela est possible, il est présent comme solution de repli. Dans la plupart des cas où ces approches hybrides sont déployées, il s'agit de parvenir, avec difficulté, à un compromis avec la forme du réseau dans son ensemble : une approche entièrement décentralisée de messagerie instantanée ne fonctionnerait pas sur l'Internet d'aujourd'hui. Pourtant, permettre la communication de client à client, lorsque cela est possible, a des avantages en termes de passage à l'échelle.

Napster, le pionnier des applications de partage de fichiers en P2P, est bien sûr un autre exemple d'un système hybride, ce qui a entraîné sa perte. D'un point de vue technique, le partage de fichiers dans Napster était décentralisé : un client téléchargeait un fichier Napster directement de la machine d'un autre client de Napster. Mais le répertoire des fichiers était centralisé, les serveurs de Napster répondant à des requêtes de recherche et agissant en tant qu'intermédiaires des connexions entre les clients. A nouveau d'un point de vue technique, cette approche hybride a été très puissante en termes de passage à l'échelle : le répertoire restait efficace face à l'augmentation de la quantité de matériel mis à disposition sur le réseau, il utilisait peu de bande passante.

Au moment même où Napster connaît son heure de gloire, il paraît bientôt évident que les expérimentations avec des modèles décentralisés ne sont pas sans détracteurs. Dans la pratique, certaines applications peuvent mieux travailler avec un design entièrement centralisé sans utiliser de technologie P2P :

Prenons par exemple une recherche sur une grande base de données relativement statiques. Les actuels moteurs de recherche Web sont capables de servir jusqu'à un milliard de pages à partir d'un seul et même endroit. Les algorithmes de recherche, à ce jour, ont été optimisés à fond pour un fonctionnement centralisé, il semble y avoir peu d'avantages à la diffusion des opérations de recherche sur un réseau peer-to-peer. Il faut aussi faire attention aux « nouveaux » échecs liés au réseau engendrés par la décentralisation : manque de fiabilité, synchronisation des données incorrectes...<sup>89</sup>

Une leçon se profile donc pour les développeurs en P2P des générations suivantes : la nécessité d'équilibrer la puissance des modèles peer-to-peer et les complications, voire les limites, des systèmes décentralisés.

---

<sup>89</sup> Entretien avec S. Abiteboul, professeur à l'ENS Cachan, 15 octobre 2009.

## SECTION 4. AUX MARGES DU RÉSEAU

Dans un scénario où dominent les connections asymétriques et la concentration des ressources, le P2P se profile donc comme la technologie capable de fournir à un système la capacité de servir des ressources avec une grande disponibilité à un coût inférieur, tout en maximisant l'utilisation des ressources de la part de chaque pair connecté au réseau. Les solutions client-serveur tirent leur force du fait que le fournisseur ajoute, dans les cas où le nombre de clients augmente, des ressources additionnelles – bande passante, équipement, accès physique. Le P2P peut obtenir un niveau de robustesse semblable en partageant les demandes de ressources sur l'ensemble du réseau. Cependant, dans la pratique, le P2P « pur » ou intégral – celui qui utilise une structure de réseau complètement décentralisée – n'est presque jamais utilisé, au profit de solutions hybrides ou de compromis. Des choix architecturaux extrêmes, dans une direction ou dans l'autre, ne se révèlent que rarement la meilleure option quand il s'agit de construire un système « utilisable » (*usable*, Taylor & Harrison, 2009 : 197) : le développement des « nouvelles » applications P2P est une histoire de compromis et d'épreuves.

La définition de la décentralisation en train de se construire dans ces projets et dispositifs amène, en effet, à repenser ce qu'est un serveur, alors même que la démarche des développeurs vise souvent à s'en détacher le plus possible, afin d'augmenter la spécificité des dispositifs par rapport à leur rivaux centralisés plus célèbres et répandus.

« Bien que la décentralisation ne soit pas une condition nécessaire à un système P2P, elle est une condition souhaitée » (Taylor & Harrison, 2009 : 125) : cette phrase suffit à résumer les questions soulevées à la fin de la section précédente. En particulier, lors de la phase de conception d'un système P2P, la plupart des développeurs se trouvent face à un compromis : mêler suffisamment de décentralisation et de centralisation pour que le système soit à la fois résistant aux ruptures et capable de passer à l'échelle en cas d'augmentation substantielle du nombre d'utilisateurs. Plutôt que la présence ou l'absence d'un ou plusieurs serveurs, ce qui qualifie le P2P comme une technologie « décentralisée » est le fait que la responsabilité du réseau se voit poussée, de manière plus ou moins centrale mais toujours importante, aux marges ou à la périphérie du réseau, au lieu d'être maintenue dans un point central et concentré.

### ENTRE DÉCENTRALISATION ET PASSAGE À L'ÉCHELLE, DANS UN ENVIRONNEMENT FRAGILE

Plusieurs tentatives visant à extraire les points communs aux différents systèmes étiquetés sous le terme « peer-to-peer », et à donner une définition inclusive, ont été mises en œuvre dès la naissance de ces systèmes à l'aube du nouveau millénaire. En 2000, Clay Shirky tente la définition suivante :

*P2P is a class of applications that takes advantage of resources e.g. storage,*

*cycles, content, human presence, available at the edges of the Internet. Because accessing these decentralized resources means operating in an environment of unstable connectivity and unpredictable IP addresses, P2P nodes must operate outside the DNS system and have significant or total autonomy from central servers. (Shirky, 2000).*

Il s'agit donc de systèmes d'ordinateurs ou de dispositifs « aux marges du réseau », ceux qui opèrent dans des environnements transitoires et souvent fragiles, qui se connectent et déconnectent fréquemment du réseau, sont souvent cachés derrière un pare-feu, utilisent des protocoles de transport et systèmes d'exploitation souvent différents avec des systèmes formés par une quantité de terminaux souvent énorme et plusieurs millions de pairs interconnectés.

La deuxième partie de la définition de Shirky en limite par ailleurs le degré de globalité, en excluant du P2P des systèmes de calcul distribué ou réparti, tels que SETI@home, et ce à cause du contrôle central inhérent à ces systèmes. Cette définition, en se concentrant sur l'environnement des dispositifs et des ressources, marque aussi un changement par rapport aux définitions précédentes qui s'étaient focalisées exclusivement sur le statut de « servants » (serveur et client à la fois) des nœuds d'un système P2P, et sur le fait que, pour être de « vrais » systèmes P2P, ceux-ci devaient être complètement décentralisés :

*A peer can act as both a client and a server (they call these servants, i.e. server and client in Gnutella). (...) The network is completely decentralized and has no point of control. Peers in a Gnutella network are typically connected to three or four other nodes and to search the network a query is broadcast throughout the network. (Shirky, 2000).*

Bien que dans les systèmes P2P les plus modernes, la définition de « peer » réponde souvent à celle employée dans Gnutella, les réseaux P2P ne doivent pas être nécessairement complètement décentralisés. Les implémentations plus modernes de Gnutella, par exemple, utilisent une approche à la fois centralisée et décentralisée, afin de faciliter le passage à l'échelle du réseau et augmenter l'efficacité des recherches. Ces réseaux sont implémentés en utilisant des « super-pairs » qui conservent en cache les informations d'emplacement des données ; les pairs peuvent ainsi obtenir une réponse à leurs recherches en ne recherchant que dans une petite fraction de l'ensemble du réseau.

## PARTAGER DES RESSOURCES À LA PÉRIPHÉRIE DE L'INTERNET

Ainsi, la définition de Shirky semble ici plus appropriée pour décrire une nouvelle « classe » d'applications, dont le dénominateur commun est le fait d'opérer dans un environnement en reconfiguration constante. D'autres auteurs et praticiens le suivent, en constatant qu'une telle définition décrit les systèmes P2P à large échelle actuellement présents sur le marché d'une façon beaucoup plus précise que l'idéal-type du « vrai » P2P (Foster & Iamnitich, 2003 ; Ripeanu, 2001 ; Adar & Huberman, 2000 ; Gunther, 2002).

La définition donnée par Oram (2004), postulant que le P2P est

*« any networking technology where crucial responsibility lies at the endpoints »*,

renforce aussi l'idée que ce qui caractérise les réseaux P2P est la responsabilité des nœuds qui se trouvent à la périphérie. La périphérie est donc constituée par des nœuds qui ne font pas partie de l'infrastructure de l'Internet ; c'est-à-dire qu'ils ne sont pas une partie intégrante des réseaux qui composent l'Internet, ou le DNS. Ces nœuds, qui ne font pas partie du tissu de base de l'Internet, forment par ailleurs la majorité des nœuds présents sur le « réseau des réseaux ». Oram remarque aussi :

*« (N)ow the problems of peer-to-peer are the problems we all face. Peer-to-peer exposed the weaknesses that exist in the current implementation of the Internet. (...) In fact, definitional inadequacies aside, peer-to-peer isn't really a set of technologies as much as it is a set of problems. »* (Oram, 2004).

Cette intuition, qui est par ailleurs implicite dans la deuxième partie de la définition de Shirky, est importante car elle contribue à faire émerger l'ensemble de facteurs et dynamiques sociotechniques qui concernent le développement d'une variété d'applications en P2P. Des questions telles que « Comment peut-on exploiter la grande quantité de ressources, telles que la puissance de calcul ou les données, disponibles en régime de quasi-gratuité sur l'Internet ? », « Comment peut-on connecter deux personnes, ou deux machines, et comment les deux connexions peuvent-elles s'entremêler ? », « Comment la liberté de circulation des données, proxy de la liberté de communication et de parole, peut-elle être assurée dans l'âge de l'information ? » ont amené aux technologies P2P. Des communautés très différentes se sont appropriées ces questions avec des préoccupations variées, ce qui rend le P2P difficile à décrire. De surcroît, tout en essayant de répondre à ces questions, elles ont créé des bases d'utilisateurs qui ont largement dépassé les attentes des développeurs, pour prendre des directions et des configurations inattendues, ce qui a influencé ensuite les développements successifs de façon très importante, comme on pourra observer au cours des chapitres suivants.

Il est donc possible de définir, pour notre argument, le P2P en tant qu'ensemble de technologies déployées et développées pour déplacer le stockage, le calcul, la responsabilité ou le pouvoir aux marges du réseau Internet, et ce pour répondre à une large variété de préoccupations différentes. En relevant ce défi, les développeurs de plusieurs de ces systèmes se sont focalisés sur les manières d'entretenir la stabilité du réseau, malgré la nature inconstante et imprévisible des nœuds aux marges. Cela s'est souvent traduit par l'utilisation d'architectures décentralisées afin de gérer une éventuelle défaillance des nœuds, et leur mobilité. Plusieurs de ces architectures décentralisées se sont révélées être, à leur tour, robustes et capables de passer à l'échelle de façon efficace : comme a noté le développeur Paul James, « la meilleure façon de passer à l'échelle est d'utiliser l'ordinateur de quelqu'un d'autre<sup>90</sup> » (James, 2011).

---

<sup>90</sup> « You can't scale better than by utilising someone else's computer ».

## OPTIMISER SANS STABILISER

Comment organiser des réseaux décentralisés de façon à obtenir une performance optimale ? Tel est sans doute le but déclaré de tout ingénieur des réseaux P2P, mais les moyens d'y parvenir semblent beaucoup moins évidents. La performance des réseaux P2P est « non-déterministe »<sup>91</sup>, note un développeur, puisque le réseau est le plus souvent construit de façon *ad hoc* et, en raison de la nature temporaire de la disponibilité de ses pairs, il est impossible de stabiliser le réseau et ses frontières pour une performance optimale. Cela n'oblige pas, pour autant, à renoncer à la performance : les ingénieurs s'engagent donc dans un processus de négociation d'algorithmes qui puissent tirer le meilleur parti de la situation à tout moment donné.

La performance n'est pas définie une fois pour toute au moyen d'une mesure valide et absolue, mais elle prend forme à la suite d'une série de mesures empiriques temporairement valides :

On se pose ces questions sans cesse, et ce n'est pas parce que la performance ne peut pas être déterminée une fois pour toute que la question est moins délicate. Au contraire, comparés à d'autres types de réseaux, les réseaux P2P sont bien les plus sensibles à des questions de performance, puisqu'on ne peut pas obtenir une réponse stable<sup>92</sup>.

Les types de questions abordées sont, par exemple : combien de temps va prendre la recherche d'un fichier ; quelle quantité de bande passante cette recherche va-t-elle consommer ; combien de passages faudra-t-il à un paquet pour rejoindre un pair qui se trouve à la marge opposée du réseau ; le réseau est-il capable de résister à la disparition d'un pair – ou, de façon plus contre-intuitive, à l'ajout d'un de ces pairs ; au cas où le réseau verrait sa taille augmenter de quelques centaines, de plusieurs milliers voire de millions de pairs, le passage à l'échelle peut-il se passer sans problèmes ? La recherche d'optimisation d'un réseau cesse, donc, d'être synonyme d'obtention progressive de la stabilité, pour acquérir une signification d'exploration et de recherche d'une série d'équilibres transitoires.

C'est, par ailleurs, de cette négociation constante qu'un système P2P est susceptible de tirer sa force en cas de défaillance. Les réseaux décentralisés opèrent, comme nous l'avons déjà rappelé, dans des environnements où les pairs sont constamment en train de se connecter et se déconnecter, et les ressources qu'ils hébergent peuvent devenir indisponibles pour différentes raisons, liées tout autant aux motivations des individus, à une participation discontinue ou à des contraintes techniques – y compris les défaillances du matériel, les coupures électriques et les attaques délibérées. Dans ces cas, susceptibles de cibler autant les terminaux domestiques que des serveurs dédiés, des algorithmes pouvant gérer la restructuration continue du cœur du réseau sont

---

<sup>91</sup> Entretien téléphonique avec Ian Taylor, développeur et auteur d'un manuel sur les technologies P2P pour étudiants en informatique, 23 septembre 2010.

<sup>92</sup> Entretien avec Fabrice Le Fessant, développeur P2P et chercheur à l'INRIA, 4 avril 2009.

souvent plus efficaces que les pare-feu. Comme le souligne un développeur, les différentes conceptions de la sécurité ont souvent un rôle très important dans les choix du modèle de réseau sous-tendant un service : « Les systèmes P2P doivent traiter les défaillances en tant que norme, pas en tant qu'exception<sup>93</sup> », et ils doivent chercher à assurer la redondance du système au moyen d'une constante négociation de l'incertitude.

---

<sup>93</sup> « *P2P systems need to treat failures as normal occurrences, not freak exceptions* ». Entretien avec I. Taylor, septembre 2010.

## EN GUISE DE CONCLUSION : LE BAR SICILIEN ET LES SYSTÈMES DISTRIBUÉS

*During the writing of our book, I decided to get away and took a three-week writing trip to Sicily. Interestingly (and unsurprisingly) one night I was in a pub, and there, they adopted a rather bizarre serving scheme... A customer had to walk to a server by a till at the centre of this long bar, order his drink, then walk back to his barstool and pass this receipt on to the bartender who would bring him his drink. I thought this was very strange, as they turned an inherently decentralized service into a centralized one! Why, I thought to myself, could I not simply pay the bartender directly?*

*Anyway, after a few drinks the bar got rather packed, and I soon realised that this scheme in fact was working rather well. This kind of disturbed me, because there I was, pondering and outlining the pros and cons of decentralized systems in my head for the book, and these Sicilians have completely circumvented the idea! So, perhaps this is all wrong... Sometime later, though, I was saved by the realization of two facts, namely: Italians, in general, don't drink so much, and I predicted that this could never work, for example, in Britain: but most importantly, I then realised that they were not running a centralised system at all. They were, in fact, running a brokered system: they centralized discovery, and therefore payment, but the service of pouring drinks was still decentralised: hence, the scalability.*

*Furthermore, I did notice that during the night every time the cashier wandered off for a cigarette, not only was there a non-trivial accumulating number of patrons but also, the bartenders finished serving their current order and then stood around doing absolutely nothing; that is, the whole system came to a halt. Hence, I concluded that this pub in Agrigento had, in fact, the same taxonomy as centralized, Napster-like P2P systems: centralized discovery and therefore with (sic) intermittent availability, with no redundancy. However, due to its decentralized communication, services that have been discovered, and paid for here, could still be concluded as normal.*

Si j'ai ici retranscrit la totalité de cette anecdote que m'a racontée Ian Taylor,<sup>94</sup> développeur et co-auteur d'un populaire manuel sur les technologies P2P pour étudiants en informatique, ce n'est pas pour réaffirmer les qualités de buveur des Italiens par rapport à celles des Britanniques (bien que la chose mérite, sans doute, ultérieure attention par les experts). En guise de conclusion de ce chapitre, l'histoire racontée par Ian résume bien l'exploration continue qu'est la mise en opération d'un système en pair à pair : les principaux défis, épreuves et critiques auxquels se trouvent confrontés ses développeurs, face au choix d'introduire un niveau de décentralisation plus ou moins élevé dans un réseau, qu'il doive servir à un moteur de recherche, un partage de fichiers, un réseau social, ou encore des hybrides entre ces

---

<sup>94</sup> Entretien avec Ian Taylor, septembre 2010. Il raconte d'ailleurs cette même anecdote dans [Taylor & Harrison, 2009].

applications.

Comme l'histoire du premier Internet et l'histoire en train de se faire du « nouveau » P2P viennent de le montrer, une architecture P2P est le résultat d'un compromis compliqué entre décentralisation, présence d'intermédiaires, passage à l'échelle, organisation autonome des terminaux/utilisateurs, et résistance aux pannes. La décentralisation peut opérer à plusieurs niveaux, ou bien s'adapter à une couche de l'application et pas à d'autres ; la recherche d'un bon point d'équilibre peut, en fin de compte, ramener à un modèle où l'intermédiation est partiellement présente, tout comme dans l'organisation des serveurs et du système de paiement du bar sicilien.

Un modèle technique distribué peut signifier un avantage économique pour l'entreprise à ses premiers pas, ou pour un utilisateur qui, en permettant à ses pairs de tirer profit de ses propres ressources matérielles, voit sa qualité de service améliorée – ou encore les deux à la fois. En même temps, le serveur ou le point central du réseau, classiquement défini comme « ordinateur hébergeant un ensemble de logiciels, dont le rôle est de répondre de manière automatique à des demandes envoyées par des clients via le réseau », se voit saisi, mis à l'épreuve, réapproprié, transformé. Comme on verra plus dans le détail au cours des prochains chapitres, en puisant dans nos cas d'étude, il assumera tour à tour les rôles de sauvegarde principale de la sûreté des données, de garant de la stabilité du dispositif pendant sa phase d'amorçage, ou encore de « délégué de tâches » centralisé pour un réseau d'utilisateurs/nœuds qui conduit ensuite ses opérations de manière distribuée au sein du dispositif – comme le caissier l'est pour les barmans dans le pub d'Agrirento.

Les dynamiques de développement d'une architecture définie comme « décentralisée » ou « P2P » montrent la validité de l'intuition d'Andy Oram quand il qualifiait, en 2004, l'architecture P2P comme « ensemble de problèmes aussi bien qu'ensemble de technologies ». Plutôt que la présence ou l'absence d'un ou plusieurs serveurs au sein du système, ce qui qualifie le P2P comme une technologie « décentralisée » est le fait que la responsabilité du réseau – en termes de ressources fournies, de gestion de données, de durabilité du système – se voit poussée, de façon plus ou moins centrale mais toujours fondamentale, aux marges ou à la périphérie de celui-ci.

Les chapitres qui suivent – l'analyse de trois cas d'étude, un moteur de recherche en P2P, une application pour le stockage distribué de données, un projet de streaming vidéo sur architecture P2P – exploreront les implications de ce choix de topologie de réseau, notamment vis-à-vis du rôle de l'utilisateur dans les dispositifs analysés, et des dynamiques de sécurité et de protection des données personnelles.

## CHAPITRE 4

### FAROO, LE MOTEUR DE RECHERCHE « RÉPARTI SUR LES UTILISATEURS »

#### INTRODUCTION

*« Copy the entire Internet to one system? Strange idea ! »*

*Faroo White Paper, 2007*

Une idée bien étrange – et pas forcément très efficace – que celle qui consiste à essayer de copier toute l’information circulant sur l’Internet sur un seul, gigantesque système. C’est ainsi que Georg Schneider, créateur de la start-up Faroo, commence à présenter à ses investisseurs et clients pionniers de la valeur ajoutée de son idée – le développement d’un moteur de recherche « totalement » décentralisé. Il a utilisé à peu de choses près ces mêmes mots, tirés d’un White Paper partiellement publié en ligne<sup>95</sup>, lors de notre premier entretien en mai 2010.

Un billet de blog de début 2010 nous introduit de façon intéressante dans le principe de fonctionnement de base, et la spécificité, d’un moteur de recherche P2P :

*Peers across the Internet join into a network via local software. You can enter a search subject and get nearly real-time results from across the world. The end result is that you find what’s relevant based on what people are talking about, instead of any optimized algorithm. The technology works just like any other peer-to-peer network. Each computer (or peer) hosts a bit of information. When requested, that information is transferred to whomever asks for it<sup>96</sup>.*

On peut déjà noter dans cette définition une première opposition implicite à l’apologie de l’automatique versus la médiation humaine, identifiée à Google (Cotte & Després-Lonnet, 2004) : le géant californien de la recherche d’information en ligne

---

<sup>95</sup> Les premières pages du White Paper sont disponibles en ligne à l’adresse <http://www.faroo.com/hp/p2p/whitepaper.html>. J’ai obtenu le document complet à la suite de ma prise de contact avec Georg Schneider.

<sup>96</sup> Entrée blog de Brad McCarthy (2010), “A look at Faroo’s P2P Search. Will this one make it?”, <http://thenextweb.com/apps/2010/05/25/p2p-search-is-here-but-who-will-use-it/>.

sera par la suite identifié plus clairement par l'équipe Faroo comme l'entité à laquelle se confronter – le monopole qui, au cours des années plus récentes, a en partie effacé la diversité de moteurs, de publics et de types de requêtes encore présente à l'aube du nouveau millénaire (Assadi & Beaudouin, 2002). Assez simplement esquissés mais tout à fait présents sont aussi les éléments centraux de la définition d'un moteur de recherche P2P, que nous retrouverons de façon bien plus approfondie dans la brève, mais dense, histoire de Faroo. Tout d'abord, le fait que le système soit P2P implique que les utilisateurs installent un logiciel sur leur machine, un client, qui leur sert à accéder au système. L'ordinateur devient dès lors part d'un réseau de recherche composé d'utilisateurs : le client P2P installé indexe les sites web visités, et partage avec les autres usagers – si on lui en fait demande – ces informations de recherche et de découverte d'information, sans qu'elles soient pour autant maintenues dans une base de données unique, ou un index central.<sup>97</sup>

En deuxième lieu, les résultats de recherche sont censés être plus proches de « ce dont les gens parlent », et représenter les affinités et les préférences communes des usagers du système. Cela est dû au fait que le système s'intéresse directement aux sites web que les usagers eux-mêmes visitent, plutôt qu'à des informations dont l'ordre et la pertinence seraient prédéterminés par des « algorithmes optimisés » selon des critères qui ne sont que partiellement transparents (comme c'est le cas pour Google). Il convient cependant de noter qu'il s'agit là du postulat de Faroo et de son diagnostic sur Google : PageRank, l'algorithme à la base du fonctionnement du « géant » de la recherche en ligne, surtout après le « *social turn* » qui a caractérisé les choix stratégiques de l'entreprise de Mountain View en 2011, est censé, lui aussi, rendre compte de « ce dont les gens parlent »<sup>98</sup>.

L'idée d'un moteur de recherche P2P, donc,

*« takes the idea of a centralized search engine like Google and spreads it out to the users. With this service installed, your computer becomes a micronode on the [...] search network »<sup>99</sup>.*

Avec l'installation du service, chaque ordinateur devient un nœud-usager d'un système distribué et décentralisé, « réparti sur les utilisateurs », qui, en émettant des requêtes de recherche, partage également des informations sur lui-même, et sur « ce à quoi il tient ».

Officiellement établie en tant que firme en 2008 (mais les débuts du développement du moteur datent de 2006), Faroo est une start-up d'origine allemande, créée par Georg Schneider et sa femme Susanne, un ingénieur informatique et une linguiste, et

---

<sup>97</sup> Entrée blog de 2008 sur « Beta », site de recensions d'applications en version *beta*, <http://beta.seadvd.com/faroo-follow-up-p2p-search-engine-will-only-work-with-more-users/>

<sup>98</sup> Voir aussi Musiani, F. (2011). « Google et les autres: vers une recherche sociale et personnalisée », *Labs Hadopi*, <http://labs.hadopi.fr/actualites/google-et-les-autres-vers-une-recherche-sociale-et-personnalisee>

<sup>99</sup> Rafe Needleman (2007) « Faroo makes your PC a searchbot », CNET News, [http://news.cnet.com/8301-17939\\_109-9779710-2.html#ixzz1T8KBtcaL](http://news.cnet.com/8301-17939_109-9779710-2.html#ixzz1T8KBtcaL)

actuellement basée à Londres. Hiba, Jiechi et Szymon, respectivement une psychologue et deux ingénieurs informatiques, rejoignent l'équipe depuis ses débuts, en 2007. Ces cinq membres constituent la totalité de l'équipe pendant toute la durée de mon travail de terrain (Hiba quittera Faroo au début de 2012). Comme le fait noter Georg, « *we don't believe in big development teams if you are aiming radical change*<sup>100</sup> ».

Faroo développe l'un des rares moteurs de recherche distribuée, ou P2P, qui n'en soit pas resté à l'état de projet universitaire ou de prototype fermé. En trois ans, Faroo est devenu un des plus prometteurs parmi les dispositifs de recherche distribuée, et peut être considéré comme un service en fonction : il relie désormais (début 2011) deux millions de *peers*<sup>101</sup>, distribués globalement d'un point de vue géographique, ce qui en fait une grille de recherche en P2P dont les ressources approchent en quantité les ordinateurs à disposition des grandes firmes de la recherche en ligne dans leurs *data centers*.

En explorant le cas de Faroo, ce chapitre se propose de comprendre les contraintes, principalement techniques mais aussi économiques, qui s'exercent sur un dispositif dont les concepteurs veulent pousser la décentralisation jusqu'au bout, pour introduire une version de P2P « intégral » dans la recherche d'information en ligne. Pour cette analyse, on reprend certaines étapes clé du développement de Faroo, qui ressortent de discussions, débats et mises à l'épreuve du dispositif. En particulier, on se concentrera sur trois aspects : la mise en opération de « six niveaux de distribution » qui constituent l'ossature du service ; l'implémentation socio-technique du compromis entre personnalisation et *privacy* ; le rôle de l'utilisateur et de sa machine dans le système, en termes de mise à disposition de ressources matérielles et de calcul, et de contribution à une « ressource informationnelle » distribuée. Cette analyse s'appuie sur une série d'entretiens effectués avec les développeurs de Faroo et sur un suivi des débats qui ont accompagné le dispositif depuis son apparition sur le marché, et aussi, de façon ponctuelle, sur un entretien conduit avec la coordonnatrice d'un projet universitaire français de moteur de recherche décentralisé, Gossple, développé à l'INRIA.

La première section du chapitre suit le positionnement économique et « académique » difficile qui fait suite au choix, de la part de l'équipe Faroo, de faire en soi que les ordinateurs des usagers eux-mêmes fassent partie d'un large moteur de recherche collectif, dont l'architecture serait distribuée de telle sorte qu'il passerait à l'échelle en proportion de sa croissance. La deuxième section explore la recherche « six fois distribuée » – l'indexation, la découverte d'information, la recherche, la classification, l'amorçage, la mise à jour du logiciel – négociée et mise en œuvre dans Faroo ; la troisième montre, en suivant trois points de discussion et de controverse au sein de l'équipe, et entre l'équipe et certains utilisateurs, comment la « centralité » de l'utilisateur (et des ressources que celui-ci apportera à l'ensemble du système) qui découle de cette démarche de décentralisation « intégrale » est à la fois la plus grande

---

<sup>100</sup> <http://www.faroo.com/hp/p2p/faq.html#opensource>

<sup>101</sup> D'après le White Paper de Faroo.

opportunité et le risque majeur pour le dispositif. Enfin, la quatrième section décrit l'une des tensions centrales dans le développement du dispositif, l'équilibre entre personnalisation des interrogations et protection de la vie privée de l'utilisateur, et elle montre comment cette tension a contribué à élaborer une définition de *privacy* de l'utilisateur qui découle « par l'architecture » du modèle distribué.

<p>2006 : Georg et Susanne Schneider, un informaticien et une linguiste, fondent Faroo.</p> <p>2007 : Hiba, Jiechi et Szymon (une psychologue et deux ingénieurs informatiques) rejoignent l'équipe. Une version closed alpha du client P2P est réalisée. Georg cherche des investisseurs.</p> <p>2007 : Faroo est finaliste à TechCrunch40. La version public beta de Faroo P2P est publiée. L'équipe annonce sur le blog de Faroo son intention de travailler à une décentralisation « intégrale ».</p> <p>2007 (octobre) : Article de Bernard Lunn sur ReadWriteWeb, sur la problématique du lien entre personnalisation et <i>privacy</i>.</p> <p>2008 : Faroo devient une entreprise (Faroo Ltd.) et est basée à Londres. Sur ReadWriteWeb, Bernard Lunn s'exprime positivement sur les chances de succès de Faroo.</p> <p>2009 : Faroo reçoit un financement de la fondation chinoise Li Ka Shing. La composante <i>real-time search</i> de Faroo est publiée.</p> <p>2009 : Révision des termes d'utilisation de Faroo avec l'argument du « pur instrument ».</p> <p>2010 : Faroo relie deux millions et demi de pairs, et revendique le statut de « <i>largest P2P search engine</i> ». Une application pour iPhone et iPad est publiée.</p> <p>2010 (février) : Bernard Lunn devient consultant pour Faroo.</p> <p>2011 : L'équipe travaille à l'harmonisation de la recherche en temps réel et de la recherche P2P.</p> <p>[2011, septembre : mon enquête de terrain prend fin]</p> <p>2012 (janvier) : Hiba quitte l'équipe Faroo.</p> <p>2012 : Développement et publication de composantes spécifiquement destinées au marché chinois (application, page d'accueil).</p>
---

Tableau 2. Chronologie du développement de Faroo, 2006-2012.

## SECTION 1. PASSER À L'ÉCHELLE : LE CHOIX DE DÉCENTRALISER LA RECHERCHE D'INFORMATION EN LIGNE

La naissance de Faroo s'inscrit dans un débat récent, mais de plus en plus fréquent dans ces arènes académiques et politiques où l'on explore les côtés plus anarchistes et libertaires de la gouvernance de l'Internet<sup>102</sup> : celui sur les systèmes de recherche « *open process* ». En mars 2010, Toni Prug, fondateur de l'initiative « Hack the State », pose explicitement la question de la centralité du choix de l'architecture des systèmes de recherche pour la qualité, la complétude et la transparence des résultats de ces recherches, et pour l'expérience de l'utilisateur avec le dispositif :

*« The architecture and protocols of the Internet and the Web might be open, developed by IETF via open process, running mostly Free Software, but the architecture of search systems remains closed. This is not good enough. As part of the democratic practice of the common, we have to have search systems built on the basis of IETF and Free Software principles. We need Open Process search systems.*

*Search systems have four distinct components: Crawler, Index, Search&Rank, and GUI [Graphic User Interface, l'interface utilisateur]. We could and should build a public infrastructure where (the) first two components are shared, and on top of the indexed Web, open interfaces to various Search and Rank algorithms and user interfaces are provided. »<sup>103</sup>*

La question de la « démocratisation » de la recherche en ligne – qui renvoie à l'idée du contrôle par les utilisateurs des informations nécessaires à l'opération de recherche, et à la possibilité de les sauvegarder sur leurs propres machines – se voit également abordée par des chercheurs en socio-économie (Lowe, 2008 ; Rieder, 2008 ; Coleman, 2009 ; Franceschet, 2010) qui, en plaidant pour une (ré)introduction de pluralité dans le marché de la recherche d'information en ligne<sup>104</sup>, soulignent que

*« such attempts can only succeed if political support, technological feasibility and economic viability can be combined in a satisfactory fashion »* (Rieder, 2008).

Tout en soulignant l'influence que ces questionnements ont acquis au cours des

---

<sup>102</sup> Un exemple parmi tous, d'où est tirée l'intervention de Toni Prug qui suit, est le site Web « Hack the State » (<http://hackthestate.org/>: « Commu(o)nism, we could call it, is a new emerging form of communism hacked with open process and new commons. The small (o) in the middle stands for open »).

<sup>103</sup> Toni Prug (5 mars 2010), « Series on Commu(o)nism: Open Process, the organizational spirit of the Internet Model, pt 2 », (<http://hackthestate.org/2010/03/05/series-on-commuonism-open-process-the-organizational-spirit-of-the-internet-model-2/>)

<sup>104</sup> Pour de nombreux auteurs, le monopole dans la recherche d'information en ligne est par ailleurs inévitable (voir par exemple Pollock, 2009).

phases suivantes de sa démarche, Georg Schneider, le fondateur et actuel directeur général de Faroo, souligne quant à lui qu'une question a été largement dominante aux débuts de la conception de Faroo :

*« [It was a question that] had very little to do, at the moment, with political equality and democratic access: 'cui bono?', I kept asking myself. If one has to try and enter a monopoly-dominated market, he better ask himself who's going to take advantage of his product. »<sup>105</sup>*

Non seulement il est important de choisir une architecture qui apportera des changements et des avantages significatifs pour l'utilisateur, mais encore faut-il que ceux-ci soient assez transparents et clairs pour que ce dernier puisse les identifier et faire un choix informé, qui découlerait de meilleures performances techniques (vitesse, souplesse) et meilleurs résultats (la qualité et la pertinence de l'information obtenue). Et encore faut-il que l'utilisateur soit suffisamment « intéressé » pour « s'enrôler » (installer le logiciel Faroo) et pour que la traduction s'opère (Callon, 2006) : qu'il confie ou délègue au nouveau moteur le soin d'organiser les informations pour lui.

La couche P2P sur laquelle Georg se propose de construire le système doit donc être en tout premier lieu un avantage technique et économique comparatif, sans quoi l'aspect égalitaire et démocratique sous-tendu par un projet de ce genre risque de ne pas être suffisant pour convaincre un nombre important d'utilisateurs de migrer. D'ailleurs, souligne Jiechi Chang, responsable des aspects P2P de Faroo, cet aspect ne *devrait pas* être suffisant : « l'idéologie n'alimente pas la recherche à elle seule »<sup>106</sup>, remarque-t-il succinctement, tandis que Georg renchérit :

*« If nothing tangible comes into the picture, between a decentralized net from activists and a shiny, full-featured, commercial, centralized internet with all the popular music, films and media and where all of your friends are, most people will a priori choose the gilded cage »<sup>107</sup>.*

Quelle est donc la « valeur » d'une architecture P2P, et comment rendre sa spécificité « tangible » ?

Une première difficulté de taille se pose ici : comme on a pu le voir lors du précédent chapitre, travailler « aux marges du réseau » n'est pas chose facile pour les développeurs en P2P étant donné le caractère transitoire du réseau sur lequel le système se base : il est question d'un nombre de pairs très large, distribués d'une façon pleinement globalisée et reliés en un réseau extrêmement dynamique que les pairs rejoignent et abandonnent sans préavis. En ces circonstances, il est pourtant nécessaire d'arriver à limiter les effets de ce qui se dessine progressivement comme

---

<sup>105</sup> Entretien, Georg Schneider.

<sup>106</sup> Entretien avec Jiechi Chang (ma traduction).

<sup>107</sup> Entretien, Georg Schneider.

étant le principal obstacle à une performance satisfaisante du système : la période de latence des réponses aux requêtes, c'est à dire le temps qui passe entre la demande d'informations qu'un usager fait au système et la réponse du moteur de recherche. Ce n'est d'ailleurs pas pour rien que Google indique, lors de tout affichage des résultats d'une interrogation de recherche, combien de temps il a fallu au moteur pour y répondre : le chiffre sert d' « inscription » au sens de l'ANT (Latour, 1987, 1991), il est la représentation visuelle du (bref) temps de latence, venant prouver aux utilisateurs qu'ils ont bien raison d'utiliser ce moteur de recherche.

Dans le cas de Faroo, pour limiter la période de latence, il est nécessaire qu'un nombre suffisamment important d'utilisateurs rejoigne et « soutienne » la performance du réseau ; jusqu'à ce que cette masse critique d'utilisateurs soit obtenue, il est nécessaire de configurer le dispositif de manière à ce que les usagers pionniers n'obtiennent pas des performances insuffisantes, qui pourraient les amener à quitter le système avant qu'il ne fonctionne à plein régime. Il faut aussi que chaque information soit sauvegardée de manière redondante, et qu'il en existe au moins une copie ou réplique (exécutée automatiquement) au moment, non prévisible, où le pair qui la stocke quitte le réseau.

C'est avec la solution de ces problèmes que les spécificités de la couche P2P pourront être mises à profit pour le bénéfice des utilisateurs. Plusieurs arguments techniques sont débattus par l'équipe au moment où s'engage l'expérimentation P2P. En premier lieu, l'architecture du moteur de recherche devra être conçue pour tolérer l'arrivée et le départ d'utilisateurs à tout moment. De cette façon, une défaillance éventuelle pourra être traitée comme une partie intégrante du système, plutôt que en tant qu'événement inhabituel et négatif qu'il faut neutraliser (« *dealing with failure must be a permanent part of the operation(s), and not an extraordinary event* »<sup>108</sup>).

En deuxième lieu, une architecture distribuée devrait permettre d'éliminer l'éventualité qu'un seul et unique point de défaillance mette le réseau en difficulté, de la même manière que la décentralisation de l'Internet d'antan devait servir, pour ses pères fondateurs, de protection contre les dysfonctionnements ponctuels (« *we are returning to the roots of the Internet, which was designed in a decentralized manner to survive any global disaster* »<sup>109</sup>).

Troisièmement, il sera peut-être nécessaire d'éliminer les « super-pairs », au profit d'une architecture dans laquelle les opérations de routage seraient totalement distribuées. Les super-pairs sont ici perçus, comme pourrait l'être un serveur, en tant que possibles faiblesses du réseau plutôt que comme sauvegarde (« *super-peers work fine until there is a "black swan" event, like the one that took down Skype for two days in 2007 !* »<sup>110</sup>)

---

<sup>108</sup> Extrait des notes de travail de l'équipe Faroo, 2006.

<sup>109</sup> Id., 2006.

<sup>110</sup> Id., 2008.

Surtout, l'équipe Faroo voit dans une architecture de type P2P une opportunité d'explorer les moyens d'améliorer techniquement la recherche et découverte d'information en temps réel (« *real time* » *search and discovery*)<sup>111</sup>. Des discussions entre les membres de l'équipe ressort l'impression que, bien que la découverte d'informations nouvelles, récentes et pertinentes ait toujours été un aspect important de la recherche en ligne, la perception par le public de ce qui est récent a changé radicalement avec la popularité croissante de services tels que Twitter. Avant Twitter, un index d'informations était considéré comme à jour si les pages étaient indexées une fois par semaine, ou dans le meilleur des cas, une fois par jour. Mais d'après Twitter, en viennent à conclure Georg et son équipe, la « fraîcheur » attendue des informations est devenue de quelques minutes, quelques heures tout au plus.

Nombre de start-up se définissant comme des moteurs de recherche *real-time* ont donc été créées, à la faveur de certaines évolutions récentes du marché : le public est de plus en plus intéressé par l'information en temps réel, tandis que les grands acteurs de la recherche en ligne restent relativement lents pour initier un processus d'intégration de dispositifs *real-time* à leurs produits. L'*Application Programming Interface (API)* de Twitter, en accès libre, fournissait à la plupart de ces startups des données suffisantes, depuis le premier jour, pour commencer leurs activités. Concentrer son attention sur l'information récente et populaire a permis aux fournisseurs de services de garder des coûts d'infrastructure modestes, en baissant temporairement la barrière d'entrée dans le marché de la recherche d'information en ligne.

Mais cette situation, dans laquelle certaines entreprises se spécialisent uniquement dans la recherche *real-time*, ne peut pas durer longtemps. Les grands acteurs de la recherche en ligne investissent désormais sérieusement cette niche de développement, tandis que certains utilisateurs commencent à souhaiter un retour d'information à la fois historique et en temps réel, qui puisse être obtenu à la suite d'une seule et unique interrogation et/ou expérience avec l'outil – ce qui pose un défi de développement à la fois pour les start-up de la recherche en temps réel et pour les grands acteurs.

Pour faire face à ce défi, Google fait savoir qu'il envisage d'augmenter ses effectifs d'un à dix millions de serveurs<sup>112</sup>, une réponse par la puissance, que Georg définit comme « *a brute force approach of copying yet more information, and yet faster, on the Internet to a central system* » et à laquelle les start-up ne peuvent répondre avec les mêmes armes. Entre ces deux tendances, l'équipe Faroo manifeste l'intention de développer une architecture P2P décentralisée qui permette de dépasser la « *primacy of money* » au moyen d'un passage à l'échelle intégré du système, et de la recherche d'information auprès de l'utilisateur<sup>113</sup>.

---

<sup>111</sup> Entretien avec Georg Schneider.

<sup>112</sup> Rich Miller (20 octobre 2009), « Google Envisions 10 Million Servers », Data Center Knowledge (<http://www.datacenterknowledge.com/archives/2009/10/20/google-envisions-10-million-servers/>)

<sup>113</sup> Les informations contenues dans ce paragraphe et dans les deux précédents sont une réélaboration d'informations exposées dans le modèle d'affaires de Faroo, qu'il ne m'est pas possible de citer *verbatim*. Les expressions entre guillemets sont dérivées d'entretiens avec Georg Schneider.

Tels sont les points de départ du travail de Georg et son équipe, sous-tendus par un principe assez simple : si on veut que l'Internet soit interrogeable sans le « copier tout entier sur un seul système », il faut que les ordinateurs des usagers eux-mêmes fassent partie d'un large moteur de recherche collectif, dont l'architecture serait distribuée tout comme l'Internet lui-même, de telle sorte qu'il passerait à l'échelle en proportion de sa croissance. Simple à envisager sur le principe, l'idée n'en oblige pas moins l'équipe Faroo, lors de la phase de développement concret du logiciel qui commence en 2006, à se poser un certain nombre de questions dont les réponses et les solutions possibles sont beaucoup moins simples, tant d'un point de vue technique qu'économique. Les prochaines sections explorent ces « étapes critiques » du développement de Faroo.

#### UN POSITIONNEMENT ÉCONOMIQUE ET « ACADÉMIQUE » DIFFICILE

Premier point problématique : les tentatives de la firme de se positionner dans le secteur économique de la recherche d'information en ligne. Aux yeux de bien des observateurs s'exprimant sur des blogs très appréciés par les spécialistes, comme *ReadWriteWeb*<sup>114</sup>, *TheNextWeb*<sup>115</sup> ou *Ars Technica*<sup>116</sup>, Georg est en train de devenir l'un des entrepreneurs qui, en essayant de convaincre des investisseurs potentiels de s'intéresser à un nouveau moteur de recherche, sera obligé de se rendre compte, à ses dépens, que l'argument « on peut battre Google » est destiné à la faillite :

*« All I was getting from people was that what I wanted to do closely and dangerously resembled pitching against Microsoft in the PC world of yesteryear. What I was getting was, none of the high profile, well-funded search engine start-ups with cool new interfaces, social search or natural language technology have made a dent in the real world. Not even Microsoft, with its army of smart researchers and piles of cash, has been able to halt Google's relentless market share gains in search. And the like. »<sup>117</sup>*

Georg échoue donc pendant plus d'un an à obtenir des soutiens financiers pour Faroo, les investisseurs potentiels refusant de s'attaquer à la puissance technique et économique de Google. Ainsi Bernard Lunn, l'un des commentateurs P2P les plus réputés du populaire site « nerd » *ReadWriteWeb*, provoque quelques réactions perplexes parmi ses lecteurs lorsqu'il annonce, début 2008 :

*« So why do I think that a bunch of engineers in Germany at a low profile*

---

<sup>114</sup> <http://www.readwriteweb.com/> [Voir aussi la section 2 de l'Introduction générale].

<sup>115</sup> <http://thenextweb.com/> [Id.]

<sup>116</sup> <http://arstechnica.com/> [Id.]

<sup>117</sup> Entretien téléphonique avec Georg Schneider, fondateur de Faroo, 5 mai 2010.

*company called Faroo have a chance? »<sup>118</sup>*

Dans ses phases initiales, Faroo a joué la carte de son statut de logiciel libre, avec les connotations d'égalité et d'ouverture qui y sont généralement associées (contre celles de monopole et de domination associées à Google). Cependant, selon Lunn, l'enjeu essentiel de cette controverse est ailleurs : l'avantage de Faroo qui est susceptible de changer le jeu, et en passe de devenir le point central du débat, est l'architecture P2P du système. Une éventuelle alternative à Google, pour avoir une chance de succès, doit pouvoir opérer un déplacement dans le modèle économique de la recherche d'information en ligne. Or, d'après Lunn, le problème auquel doit faire face un grand système de recherche centralisé est celui des coûts : pour pouvoir fonctionner avec toujours plus d'utilisateurs, la puissance du système doit progressivement augmenter afin de pouvoir « copier tout l'Internet sur un seul système », en y ajoutant constamment des serveurs et de l'argent pour les acheter<sup>119</sup>.

Quelques temps après, Georg Schneider publie un billet de blog sur le site de Faroo (Tableau 3), qui concerne des questions techniques de passage à l'échelle. Mais celles-ci sont mises en relation directe, pour la première fois dans l'histoire du dispositif, avec des questions de « barrières à l'entrée du marché » :

#### ***Scaling & Market Entry Barrier***

*In web search we have three different types of scaling issues:*

*Search load grows with user number. **P2P scales organically, as every additional user also provides additional infrastructure.***

*With the growth of the internet more documents needs be indexed (requiring more index space). **P2P scales, as the average hard disk size of the users grows, and the number of users who might provide disk space grows as well.***

*With the growth of the internet more documents needs to be crawled in the same time. **P2P scales as the average bandwidth per user grows, and the number of users who might take part in crawling grows as well. Additionally P2P users help to smarten up the crawling by discovering the most relevant and recently changed documents.***

*For market dominating incumbents the scaling in web search is not so much a problem. For now they solve it just with money, derived from a quasi advertising monopoly and its giant existing user base. But this brute force approach of replicating the whole internet into one system doesn't leave the Internet unchanged. It*

---

<sup>118</sup> Bernard Lunn, « Faroo : Could P2P Search Change the Game ? », en ligne sur [http://www.readwriteweb.com/archives/faroo\\_could\\_p2p\\_search\\_change\\_the\\_game.php](http://www.readwriteweb.com/archives/faroo_could_p2p_search_change_the_game.php). La compagnie était à ce moment basée en Allemagne et au stade de closed alpha.

<sup>119</sup> En 2008, estimés par Faroo à 450000 serveurs et 2 millions de dollars (d'après le White Paper, 2007).

*bears the danger that one day the original is replaced by its copy. But for small companies the huge infrastructure costs are posing an effective market entry barrier. Opposite to other services, where the infrastructure requirements are proportional to the user number, for web search you have to index the whole internet from the first user on, to provide competitive search results. This is where P2P comes in, effectively reducing the infrastructure costs and lowering the market entry barrier.*

Tableau 3. Post « *Scaling & Market Entry Barrier* » sur le blog de Faroo, <http://blog.faroo.com/>, mai 2008.

Si l'idée a ses détracteurs, elle trouve aussi beaucoup de soutien au sein des discussions parmi les spécialistes, les développeurs et les usagers pionniers ou potentiels s'exprimant sur les sites IT – et les termes du débat se déplacent : l'avantage économique est possible pour une petite entreprise s'appuyant sur le P2P. La question est désormais : que faut-il techniquement à la recherche sur P2P pour être opérationnelle, non pas au niveau de la démonstration ou du prototype, mais à celui d'application pour le marché de masse. Les moyens semblent assez controversés. Comme le note Jiechi Chang, le responsable R&D Peer-to-Peer de l'équipe,

*« In fact, the idea of P2P Search has been around for a while and has been investigated in academic circles, but with contradictory findings about the feasibility. Until now, it has been very difficult to demonstrate the necessary combination of scaling, critical mass of users, search speed, quantity and relevancy. »<sup>120</sup>*

L'équipe Faroo se trouve face à une masse importante de travaux scientifiques concernant la recherche d'information en ligne sur des architectures P2P. Une partie de cet ensemble de documents a trait au grand potentiel du modèle P2P en termes de passage à l'échelle, d'efficacité et d'endurance, et aux manières dont un moteur de recherche P2P peut bénéficier de « l'input » intellectuel d'une grande communauté d'usagers. La littérature discute également assez largement les difficultés techniques pour identifier et sélectionner efficacement les pairs qui pourront mieux servir un besoin informationnel particulier.

Le fait de manier une architecture P2P modifie les points de vue sur certains aspects essentiels. Par exemple, l'absence d'une fonction centralisée d'indexation, ajoutée aux difficultés de calculer des paramètres globaux dans un vaste réseau très dynamique, gêne l'utilisation des méthodes traditionnelles de sélection et d'extraction de l'information (Bender et al., 2005). En outre, l'absence d'une instance de contrôle central peut faire en soi que les pairs se renseignent sur certaines portions du Web d'une manière largement non coordonnée (Chernov et al., 2007). Pour certains auteurs, il semble difficile de faire en sorte que ces pairs autonomes ne forment pas un ensemble d'informations se superposant ou des cloisons disjointes d'un espace de recherche combiné (Patterson, 2004).

---

<sup>120</sup> Entretien téléphonique avec Jiechi Chang, 13 octobre 2010.

Georg Schneider et son équipe avancent ici un argument central : la caractéristique spécifique de leur application, comme ce qui peut en assurer le succès, tient au fait de décentraliser l'architecture du dispositif, et de modifier ainsi le rôle de l'utilisateur dans le processus de recherche. Ils s'appuient sur plusieurs points.

En premier lieu, considèrent-ils, l'ère de la recherche centrée sur le document doit et peut être dépassée. Il est nécessaire de prendre en compte l'importance des usagers, à la fois comme cible de la recherche d'information et comme contributeurs à sa découverte, et ces aspects doivent se refléter dans l'infrastructure du système.

Deuxièmement, ils choisissent de « prendre au sérieux », au niveau de la conception, le principe selon lequel l'origine et le destinataire d'une information se trouvent distribués aux marges du réseau, comme c'est (encore) le cas dans l'Internet d'aujourd'hui. Un choix de design plus approprié à cette topologie de réseau comporterait donc que le processus de recherche, lui aussi, soit distribué aux marges du réseau – bien que cette distribution soit quotidiennement mise à l'épreuve, aujourd'hui, par des services et des systèmes d'exploitation éloignés de l'utilisateur.

En troisième lieu, l'architecture P2P ou distribuée peut être un avantage lorsqu'il s'agit d'indexer la totalité du Web – et donc au delà de sa « tête » (c'est à dire les documents récents d'un point de vue chronologique et auxquels on a prêté beaucoup d'attention) – car elle peut prendre en compte, de manière différente, la dimension locale de l'information à rechercher.

#### CE QUE « FAIT » L'ARCHITECTURE D'UN MOTEUR P2P

Par ces réflexions, l'équipe Faroo ouvre encore une fois la discussion aux usagers potentiels<sup>121</sup>. Plusieurs d'entre eux demandent aux développeurs de comparer le service envisagé à d'autres moteurs de recherche, décentralisés ou non. L'équipe prend alors conscience du fait que la question principale qu'ils doivent affronter remonte d'un cran : il n'est pas (encore) question du fait qu'ils sont « réellement » en train de construire un moteur de recherche Web décentralisé, mais de savoir ce qui qualifie, ce qui « fait » l'architecture P2P d'un moteur de recherche<sup>122</sup>. La décentralisation de Faroo prend donc forme, à la fois comme un avantage compétitif et un aspect originel/différent du système, à travers des réponses que l'équipe donne à la communauté.

Georg met en avant l'idée qu'un moteur de recherche, pour être décentralisé, doit « vraiment » utiliser le modèle P2P et les tables de hachage distribuées (*Distributed*

---

<sup>121</sup> A la fois sur le blog de Faroo (<http://blog.faroo.com>), sur le forum utilisateurs de Faroo (<http://forum.faroo.com>), et à l'occasion de discussions et révisions du produit sur des sites tiers, comme ReadWriteWeb et ReviewSaurus. Ces discussions seront citées plus précisément au fur et à mesure que l'on en mobilisera des extraits.

<sup>122</sup> Entretien avec Georg Schneider, 20 juillet 2010.

Hash Tables, DHT)<sup>123</sup> « in practice ». Dans une DHT, chaque pair (ou groupes de pairs, pour assurer la redondance) peut être responsable des différentes parties de l'index. La réponse à différentes interrogations dérive de différents pairs ; le client P2P Faroo de chaque utilisateur compte et liste les pairs avec qui il entre en contact lors de chaque requête, et l'utilisateur peut consulter cette liste (Image 1).

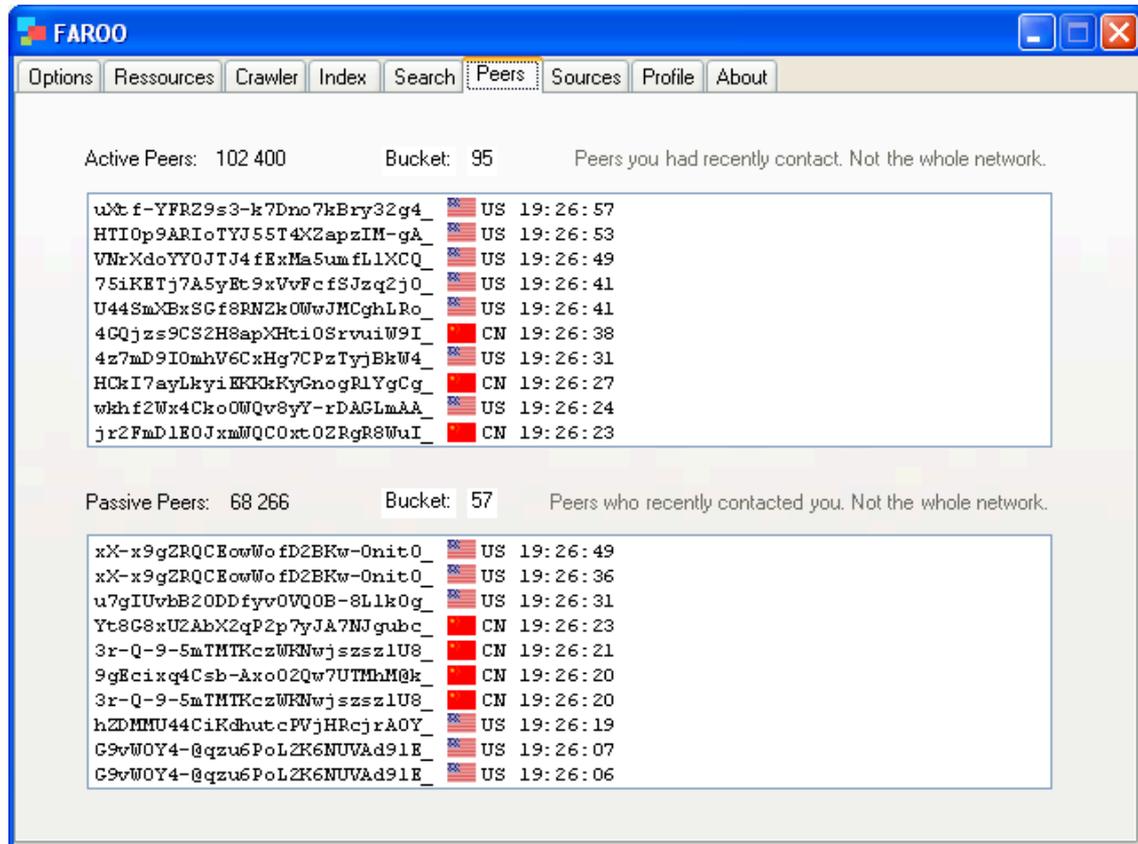


Image 1. Capture d'écran du « compteur » des pairs contactés par le client P2P d'un utilisateur au cours d'une interrogation de Faroo.

Un modèle de recherche P2P doit aussi être capable de passer à l'échelle dans une situation où l'on interroge non pas quelques dizaines de pairs, mais plusieurs centaines de milliers. Si, indépendamment du terme recherché, ce sont toujours les mêmes pairs (c'est-à-dire toujours les mêmes adresses IP) qui sont interrogés, le modèle sous-jacent n'est pas du P2P, mais se rapproche plutôt d'une « ferme de serveurs »<sup>124</sup> traditionnelle, ou d'un métamoteur.<sup>125</sup> De même, si les réponses aux

<sup>123</sup> Une table de hachage distribuée (ou DHT pour *Distributed Hash Table*) est une technologie permettant l'identification et l'obtention, dans un système reparté tel que certains réseaux P2P, d'une information. L'ensemble de la table de hachage est constituée virtuellement par tous ses constituants répartis sur tous les éléments du réseau, qui en possèdent chacun une partie.

<sup>124</sup> Une grappe de serveurs, ferme de serveurs, ou ferme de calcul (*computer cluster* ou *server farm* en

interrogations du moteur de recherche ne proviennent que d'un nombre limité d'adresses IP proches les unes des autres, c'est aussi une indication de la présence d'une ferme de serveurs.

Enfin, pour être P2P, un moteur de recherche doit impliquer le téléchargement d'un client, une application qui constitue l'interface entre l'utilisateur et le réseau de pairs ; cet aspect, en particulier, représente, comme on le verra plus tard, une source de perplexités et de doutes pour les utilisateurs, dans la mesure où il représente une « barrière à l'entrée » de taille pour l'utilisateur lambda accoutumé à un moteur de recherche accessible depuis un simple navigateur Web.

---

anglais) est un regroupement de plusieurs ordinateurs indépendants afin de permettre une gestion globale et de dépasser les limitations matérielles d'un ordinateur unique.

<sup>125</sup> Un métamoteur (ou méta-moteur ou méta-chercheur) est un logiciel qui puise ses informations à travers plusieurs moteurs de recherche, c'est à dire qu'il envoie ses requêtes à ces moteurs, et retourne les résultats de chacun d'entre eux. Le métamoteur permet aux utilisateurs de n'entrer le sujet de leur recherche qu'une seule fois mais d'accéder aux réponses de plusieurs moteurs de recherche différents. Un métamoteur élimine les résultats similaires ; par exemple, si Google et Yahoo! renvoient sur les deux mêmes liens, le métamoteur ne va l'afficher qu'une seule fois dans la liste des résultats.

## SECTION 2. EXPLORER LES NIVEAUX DE DISTRIBUTION

Comme Georg et Susanne me l'ont expliqué par la suite, c'est cette période de mise à l'épreuve de la définition même d'un moteur de recherche P2P, entre 2006 et 2007, qui leur fait prendre conscience de la nécessité de pousser l'approche P2P jusqu'au bout afin de se démarquer de leurs concurrents, à la fois centralisés et décentralisés. Mais comment façonner cette distribution ?

L'équipe Faroo s'est progressivement rendue compte que des aspects tels que l'arrivée et le départ des pairs du système, les mises à jour du dispositif, et la capacité du système à récupérer en cas de défaillances doivent être traités de façon très différente dès lors que l'on entre dans la phase de leur implémentation à un niveau technique. Selon les mots de Szymon,

*« joining peers, updates and recovery look different from a bird's eye view than from the ground »<sup>126</sup>.*

Des questions qui n'avaient pas été auparavant prises en compte commencent à s'imposer dans les discussions et le travail du groupe. Parmi elles figurent la capacité de l'algorithme d'amorçage<sup>127</sup> du système à passer à l'échelle parallèlement à l'augmentation des dimensions du réseau ; la capacité du système à récupérer ses fonctionnalités en cas d'importantes fluctuations dans le nombre de pairs connectés, en minimisant l'impact visible sur les usagers (« et avec grâce »<sup>128</sup>) ; la capacité du système à évoluer dans le temps, donc la possibilité pour le système de bénéficier d'une distribution des mises à jour du logiciel fréquente et rapide, peut-être immédiate.

Ces arguments amènent Georg et Jiechi à formuler la « philosophie de développement » de Faroo d'une façon qu'ils n'avaient pas envisagée au début de leur démarche, et qui relie indissolublement l'« expérience utilisateur » distribuée et l'architecture technique distribuée qui la sous-tendra :

*« The benefits of a distributed architecture, such as low cost, high availability and autonomy, will be fully realised only if the operational side is also fully distributed »<sup>129</sup>.*

Pour que la distribution et la décentralisation du dispositif se configurent en tant que points de force grâce à la flexibilité du système, plutôt que comme une faiblesse issue

---

<sup>126</sup> Entretien avec Szymon.

<sup>127</sup> *Bootstrapping* : le processus d'« entrée » initiale de pairs dans le réseau (voir le chapitre 2, section 2).

<sup>128</sup> « *Gracefully* » : Georg utilise cet adjectif lors d'une réponse à un usager ayant lu et commenté l'article de Bernard Lunn ([http://www.readwriteweb.com/archives/faroo\\_could\\_p2p\\_search\\_change\\_the\\_game.php](http://www.readwriteweb.com/archives/faroo_could_p2p_search_change_the_game.php)).

<sup>129</sup> Entrée blog de Faroo, février 2007 (<http://blog.faroo.com/2007/02/>).

de sa fragilité ou de son statut transitoire, l'équipe Faroo doit travailler avec la distribution à plusieurs niveaux, en incluant les aspects *opérationnels*, c'est-à-dire tout ce qui concerne le fonctionnement en profondeur du dispositif, au delà de la fonctionnalité principale de recherche.

L'intention de se libérer complètement de serveurs, d'éléments centraux et de points de passage obligés, est peu après rendue publique par Faroo, dans un communiqué de presse de début 2007 :

*It is important to look at the scaling of all operational aspects, not only at the main search functionality. The weakest element defines the overall scalability and reliability of a system. (...) There should be no centralized element that can fail, be attacked or blocked as a single point of failure or which simply does not scale; not for crawling, not for indexing and search, not for ranking and discovery, not for bootstrap, and not for update.*<sup>130</sup>

Dans les coulisses, l'équipe Faroo s'attache entre temps à considérer les éléments qui, dans les moteurs de recherche déjà présents sur le marché, contiennent des aspects décentralisés ; quelles leçons peuvent-ils en tirer, et sur quels fronts doivent-ils concentrer leurs efforts ?

#### LA FORCE DE LA COLLABORATION, OU LE *CRAWLER* DISTRIBUÉ

Un trait commun à plusieurs services similaires présents sur le marché est que le robot d'indexation, ou *crawler*,<sup>131</sup> est distribué, alors que l'index lui-même, et le moteur de recherche, sont toujours centralisés – un modèle utilisé, par exemple, par Grub, le crawler de Wikia Search<sup>132</sup>.

Szymon, l'ingénieur informatique conseiller scientifique du projet, s'aperçoit cependant que le *crawler* distribué n'offre, à lui seul, que des avantages limités. Tout d'abord, parce que transférer les pages explorées en retour à un serveur central n'économise pas beaucoup de bande passante, par rapport à la quantité nécessaire pour que le serveur les télécharge lui-même. En outre, la distribution de la charge de travail d'une manière fiable entre les éléments du *crawler* (c'est-à-dire les pairs eux-mêmes, qui sont donc présents de façon transitoire dans le système) présente des coûts additionnels.

Jiechi, qui est plus particulièrement responsable de la stabilité et de l'efficacité de la

---

<sup>130</sup> Faroo Product Brochure, version de 2007.

<sup>131</sup> Logiciel qui explore automatiquement le Web et en collecte les ressources afin de permettre à un moteur de recherche de les indexer.

<sup>132</sup> Wikia Search a été un moteur de recherche Web libre et open-source, lancé par Wikia, une société créée à la fin de 2004 par le fondateur de Wikipédia, Jimmy Wales, et par Angela Beesley. Wikia Search a été lancé officiellement en tant que *public alpha* le 7 janvier 2008, mais après avoir échoué à attirer un public important, le site a fermé en 2009.

technologie P2P qui sous-tend le système, observe en revanche qu'il y a des avantages à l'approche hybride (*crawler* distribué/index et moteur centralisés) dans la mesure où cela assure un meilleur fonctionnement de certaines composantes du moteur de recherche. Par exemple, le robot d'indexation est plus difficile à détecter et à bloquer si la charge est distribuée et provient de différentes adresses IP. Le *crawling* distribué ne sera à la hauteur de ses promesses que dans le cadre d'un moteur de recherche dont l'architecture sera entièrement distribuée. Cela ne fonctionne que si les robots d'exploration ne sont pas contrôlés par une seule instance de serveur centralisé, mais qu'ils agissent de façon autonome, dirigés par les actions des utilisateurs, leur proximité géographique, les affinités entre leurs requêtes au système :

*« network wide effects can be achieved by utilizing the geographic or contextual proximity between distributed index and crawler parts »*<sup>133</sup>.

Par exemple, résume Georg, les pages qui changent souvent – comme celles d'un article d'actualité qui serait fréquemment mis à jour avec des résultats électoraux, intéressant donc un grand nombre d'utilisateurs – sont également réindexées plus souvent. On peut donc dire que les utilisateurs :

*« implicitly control the distributed crawler in a way that frequently changing pages are kept fresh in the distributed index, while preventing unnecessary traffic on more static pages. »*<sup>134</sup>

Jiechi précise également la raison pour laquelle le crawler semble, par moments, lent, ce qui lui est reproché par quelques-uns des premiers utilisateurs du forum Faroo. Selon lui, il est nécessaire que les utilisateurs fassent la distinction entre la force du crawler collaboratif pris dans son ensemble, et l'impression de manque de rapidité que chaque utilisateur peut en retirer en étant – et donc en ne gérant – directement qu'une petite partie du système :

*FAROO's collaborative crawler swarm is actually very fast. The number of crawled pages per single client is relatively low for two reasons: [...] It is designed as low impact crawler for a smooth user experience, without using too much bandwidth and processor load from a single user. The power is rather defined by its massive scalability [ ; second,] the crawler speed is determined by the distribution of the data to the decentralized index, not by fetching and parsing the web pages*<sup>135</sup>.

Est donc demandé<sup>136</sup> à l'utilisateur de laisser, au moins temporairement, de côté l'impression de ce qu'il considère habituellement comme un moteur de recherche

---

<sup>133</sup> White Paper, p. 10.

<sup>134</sup> Entretien avec Georg Schneider.

<sup>135</sup> Faroo, Frequently Answered Questions, <http://www.faroo.com/hp/p2p/faq.html#results>.

<sup>136</sup> A la fois comme attente implicite suivant les mots des FAQ, et comme attribut du dispositif technique.

rapide *et donc* efficace – pour se considérer comme une brique dans une construction dont la stabilité et la force ne viennent pas des pics de performance « à l'unité » de l'outil, mais plutôt du fait de participer à un effort global. En même temps, le robot d'indexation est redéfini, non plus en termes de nombre de pages indexées par personne/machine, mais en termes de sa capacité de passage à l'échelle.

#### LA « SAGESSE DES FOULES », OU LA DÉCOUVERTE DISTRIBUÉE D'INFORMATION

L'efficacité d'un système d'indexation distribuée a deux faces : la première est que les pages qui sont fréquemment modifiées ou changent au vu de mises à jour automatiques soient maintenues « fraîches » dans l'index distribué ; la deuxième est d'éviter un gaspillage de ressources et de trafic sur les pages plus statiques, où aucun changement ne s'est produit. Les développeurs de Faroo cherchent à traduire dans un dispositif décentralisé la réponse à une question que les grands moteurs de recherche du marché se posent aussi – comment analyser l'ensemble du Web, chaque jour plus complexe, en quelques minutes, et découvrir de nouveaux contenus en très peu de temps ? – et que ceux-ci résolvent, partiellement, en augmentant leur « puissance de feu », c'est-à-dire en ajoutant toujours plus de serveurs à leurs fermes. Ce n'est que si le robot est dirigé sélectivement vers les pages nouvellement créées que la recherche en temps réel devient réalisable : il s'agit de mettre l'approche distribuée et le passage à l'échelle à l'épreuve des 100 milliards de pages qui constituent le Web. Jiechi commence donc à travailler sur les façons possibles d'utiliser la « sagesse des foules » (*wisdom of crowds*), c'est-à-dire, ici, les connaissances et les préférences que les utilisateurs expriment et valident implicitement en ouvrant des pages Web :

*« By aggregating and analyzing all visited web pages of our users for discovery, we started to think of our users as our scouts. »<sup>137</sup>*

Et Szymon renchérit:

*« By the very action of search and page selection, they bring in their collective intelligence and turn the crawler there where new pages emerge. The tough part is getting enough users to form a collective intelligence that is articulate enough. [We're] getting there though. »<sup>138</sup>*

Une conception de l'utilisateur en tant que « éclaireur » (*scout*) de la fiabilité, de la pertinence et de la sélection elle-même des résultats se dessine donc dans les paroles des acteurs. Dans l'algorithme de découverte et présentation des pages de la part du moteur de recherche, l'importance que l'utilisateur lui-même attribue à une nouvelle page devient le critère plus important – et se trouve davantage mise en relief si d'autres utilisateurs partagent ce même intérêt. Les requêtes au moteur de recherche se reconfigurent en données dynamiques pour « chercher l'information là où elle est

---

<sup>137</sup> Entretien avec Jiechi Chang, ref. Cette question, ainsi que le classement *attention-based* distribué, sera reprise plus en détail ci-dessous.

<sup>138</sup> Entretien avec Szymon.

finalement : chez l'utilisateur »<sup>139</sup>.

La pertinence des résultats dépend en même temps de la dimension collective de cette exploration « *bottom-up* », car la présence d'un nombre suffisant de ces éclaircisseurs est déterminante pour assurer que les résultats seront assez variés, nombreux et complexes pour être un miroir fiable des priorités et des affinités (préférences communes) entre les utilisateurs, et faire du moteur Faroo un outil attirant et pas simplement « original ». Il s'agit là d'un problème courant en micro-économie, celui « de la poule et de l'œuf » (Caillaud & Jullien, 2002), qui se retrouve régulièrement lors des suivis de la transformation des normes techniques dans les contextes les plus hétérogènes. Pour Faroo, il s'agit également d'un moment d'épreuve très délicat :

*« It's a chicken and egg problem: without users, no accurate results ; without accurate results, more users won't want to join<sup>140</sup>. »*

Pendant la phase *beta* (2007 et début 2008) la faible quantité des résultats se révèle être en effet un des inconvénients les plus importants aux yeux des utilisateurs/testeurs<sup>141</sup> : plusieurs d'entre eux se manifestent avec des questions qui peuvent se résumer en « pourquoi y a-t-il si peu de résultats ? ». La réponse des développeurs en première instance est de requalifier le dispositif comme une expérimentation au stade de développement en cours, et en quelque sorte de solliciter l'indulgence de ses usagers pionniers :

*« We are still in beta. This is mainly intended to test and improve our distributed architecture. Due to the relatively small number of users during the beta test also the number of pages indexed is limited. (...) So while trying FAROO you should rather judge the quality of the technology than the current quantity of the index<sup>142</sup>. »*

Si les discussions entre l'équipe et les utilisateurs pionniers se poursuivent sur la mise en pratique de ces propos (comment exactement, demandent certains, ces derniers peuvent-ils « juger de la qualité de la technologie » si le résultat visible est peu édifiant ?), une solution technique est envisagée au cours de 2008 pour répondre à ce point. Elle entraîne une modification du *crawler* : en plus de l'indexation des pages visitées par les usagers, celui-ci est mis en condition de conduire une partie de l'indexation de façon autonome<sup>143</sup>.

---

<sup>139</sup> Pour reprendre les mots d'Anne-Marie Kermarrec, responsable de Gossple, un projet universitaire de moteur de recherche collaboratif en P2P, basé à l'université de Rennes et financé au niveau européen (Entretien, 10 octobre 2009).

<sup>140</sup> Memo par Georg à Szymon, cité par Georg lors de notre deuxième entretien.

<sup>141</sup> Qui ont été mobilisés et ont rejoint le système à la suite de plusieurs annonces et articles parus, on le rappellera, sur le blog de Faroo (<http://blog.faroo.com>), sur le forum utilisateurs de Faroo (<http://forum.faroo.com>), et sur des sites tiers, comme ReadWriteWeb et ReviewSaurus.

<sup>142</sup> Faroo, Frequently Answered Questions, <http://www.faroo.com/hp/p2p/faq.html#results>.

<sup>143</sup> Entrée de blog Faroo, janvier 2008 (<http://blog.faroo.com/2008/01/08/new-active-community-directed-crawler/>).

*« The crawler became more 'active' while remaining community directed. Crawler start points are still derived from visited pages and searches of the FAROO users, but if a search returns only few or no results, other pages are crawled in real-time and included in the results of that search. (...) Active crawling increases the index size at faster pace and overcomes the chicken-egg problem when crawling only visited pages with relatively few users. By active crawling also passive peers may contribute. Increasing the index becomes independent from browsing activity, in this way also pages get indexed which nobody from the current FAROO community visited before. »<sup>144</sup>*

Les pages visitées et les requêtes des usagers Faroo gardent donc leur place centrale dans le dispositif de recherche, mais au lieu d'être les seuls résultats, ils constituent désormais un point de départ pour l'indexation qui se poursuit de façon automatique auprès de pages similaires, ce qui permet d'épaissir et de compléter l'index plus rapidement tout en limitant les effets réducteurs d'un index qui comprendrait exclusivement des pages déjà visitées. De façon intéressante, Jiechi souligne aussi comment le crawler prend en compte, désormais, les contributions à l'index collectif fournies par des utilisateurs qui ont différents niveaux d'engagement dans le système : puisque la complexification de l'index devient partiellement indépendante des activités de recherche d'information de la part des pairs actifs, cela permet d'inclure dans l'index à la fois global et distribué la contribution des « pairs passifs », ceux qui ne lancent pas, ou qui lancent très peu de requêtes mais qui, du fait qu'ils ont le client P2P Faroo installé et mettent dès lors une partie de leur ressources CPU (entre autres) à disposition, font quand même partie de sa « communauté ».

Dans les endroits de discussion et de mobilisation des utilisateurs, tels que les forums des sites spécialisés ReadWriteWeb ou TheNextWeb, la recherche distribuée est discutée comme un compromis entre donner la prééminence aux résultats venant de peu de sources – mais connues et fiables – et à ceux venant de l'agrégation des préférences collectives :

*« Even if the results of P2P are more accurate in a real-time format, sometimes I want to know a trusted source instead of a crowd source »*

*« Real-time, socially relevant results do indeed have their place for my own use, and I'm sure they would for others as well ».*

La principale opportunité, et en même temps, le principal problème perçu par les utilisateurs semblent être, quand même, la radicalité de l'approche P2P promu par Faroo. Le fait que la start-up propose un changement fondamental de l'infrastructure de base sous-tendant le processus de recherche – infrastructure qui n'a pas été changée depuis quinze ans – pourrait être responsable de sa fin si la technologie P2P est appliquée prématurément à un secteur qui n'est pas encore prêt :

*« I [...] think that it's far too early in the P2P search game to rule out a*

---

<sup>144</sup> Entretien avec Jiechi Chang.

*hybrid-style engine*<sup>145</sup>. »

Surtout, la portabilité et la mobilité des dispositifs de connexion à l'Internet pourraient constituer des obstacles de taille pour Faroo. C'est ce qui pousse un commentateur à demander si l'équipe Faroo est en train d'envisager l'intégration de son produit avec un ou plusieurs acteurs de la recherche Web déjà existants, comme Google – en suggérant peut-être implicitement que si Faroo marche vraiment bien, il est destiné à... être acheté par Google :

*« Right now, we're seeing a paradigm shift towards mobile devices. Many people have already given up on desktop machines, in favor of lighter and more mobile choices. But with that shift comes a lack of resources for peer-to-peer networks, don't you think? »*<sup>146</sup>

Georg explique alors comment la technologie Faroo peut faciliter l'intégration à un autre système afin d'assurer la transition depuis un modèle centralisé, et reformule la question de la mobilité d'une façon rassurante pour le futur du dispositif :

*« The point is that tomorrow's lightweight, mobile devices are as powerful as yesterday's high end desktop hardware. In this respect, P2P technology is universal, far beyond the actual implementation for a specific device type, and most importantly, it is removing the market entry barrier of high data center costs. »*<sup>147</sup>

#### MOINS DE COÛTS D'INFRASTRUCTURE, OU L'INDEX DE RECHERCHE DISTRIBUÉ

Les coûts liés à l'infrastructure d'un modèle centralisé sont, en effet, un des aspects stratégiques envisagés dans le modèle Faroo. Le stockage d'information en soi ne pose pas de problème technique de taille ; ce qui est cher, et qui constitue une barrière à l'entrée considérable pour un nouvel acteur dans le marché des moteurs de recherche, ce sont les énormes centres de données nécessaires pour stocker l'index de recherche et répondre à des millions de requêtes en parallèle. La décentralisation peut considérablement réduire ou supprimer ces coûts, car un index distribué passe à l'échelle au fur et à mesure que le nombre des utilisateurs augmente, tout en fournissant l'infrastructure supplémentaire nécessaire pour leurs interrogations ; il bénéficie également de l'augmentation et de l'abattement des prix des ressources matérielles, doublant tous les deux ans selon la loi de Moore. Les index des grands moteurs de recherche sont à ce jour distribués sur des centaines de milliers d'ordinateurs, concentrés dans de puissants centres de données ; en distribuant l'index de recherche à la périphérie du réseau, où résident à la fois les utilisateurs et les

---

<sup>145</sup> Discussion du post de B. McCarthy, <http://thenextweb.com/apps/2010/05/25/p2p-search-is-here-but-who-will-use-it/>

<sup>146</sup> Id.

<sup>147</sup> Id.

contenus, les données ne doivent pas être transportées depuis une instance de recherche centrale, et y retourner (Image 2).

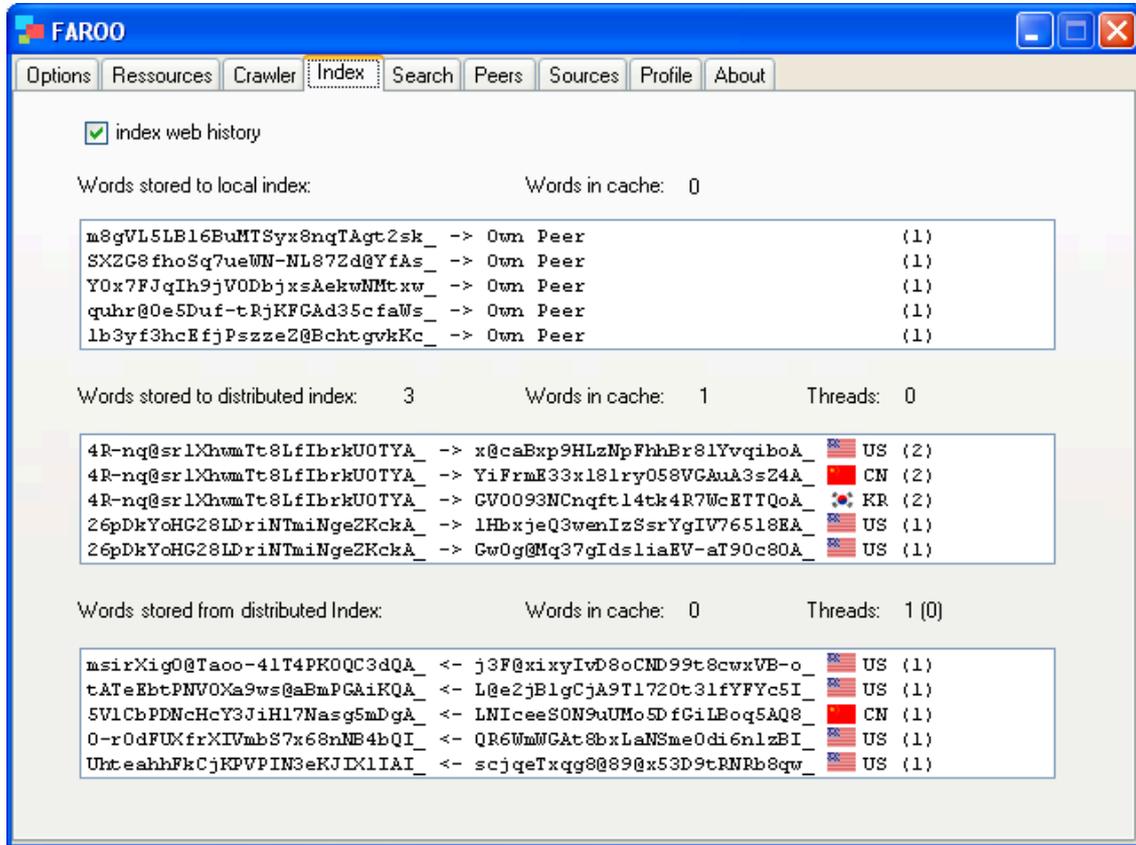


Image 2. Capture d'écran montrant à l'utilisateur la construction de l'index distribué, et son stockage en local par chaque nœud/pair. Chaque utilisateur peut suivre la manière dont l'index se construit et quel rôle a sa machine.

Jiechi accorde une grande importance – c'est l'aspect sur lequel il travaille en priorité au moment où j'écris – au fait que l'architecture du moteur prenne en compte la proximité géographique des pairs, offrant la possibilité d'une réduction ultérieure des ressources nécessaires à leur mise en relation et à leur interaction. Cet aspect n'est apparu que relativement tard dans la conception du dispositif, quand, lors d'une conférence en septembre 2008, on lui a suggéré que cela pouvait constituer un avantage comparatif par rapport aux grands centres de données dispersés qui font l'ossature des services centralisés :

*So this guy comes to me and says, 'besides preventing the single point of failure, combining the index distribution across multiple computers, in fact, it helps to leverage the geographic proximity that you normally get by spreading multiple data centers across the globe, right? It looks quite clever.' I didn't*

*really know how to tell him that I hadn't really given that much importance to the proximity aspect, but I knew Georg would have loved to hear this when I came back. It was, indeed, an economic advantage for the business that I could work on by addressing issues of technical infrastructure.*<sup>148</sup>

#### LE « POUVOIR DES DÉCLICS », OU LE CLASSEMENT DISTRIBUÉ PAR LE WEB IMPLICITE

Un quatrième niveau de distribution sur lequel l'équipe Faroo commence à travailler, en parallèle avec la découverte d'information distribuée, est celui qui concerne le classement (*ranking*) de l'information. Si la phase de recherche et découverte d'information est conçue afin de mettre à profit la « sagesse des foules », la manière dont l'information est classée, son importance, l'attention qu'on lui dédie, devrait aussi pouvoir en profiter<sup>149</sup>. Explorer les solutions existantes à ce sujet, et essayer de comprendre comment elles pourraient s'articuler à l'algorithme Faroo, est plus particulièrement le travail d'Hiba, une psychologue cognitive qui, au sein de l'équipe, est en charge de la « qualité » de l'expérience de recherche. La « sagesse des foules » est-elle traduisible techniquement ? La discussion et les essais s'engagent au sein de l'équipe, début 2008.

Pour trouver l'information la plus pertinente possible à la requête d'un usager du système, en principe, il faudrait pouvoir classer à chaque fois la totalité du Web. Afin d'assurer une classification la plus précise possible, chaque document devrait être classé par un nombre d'utilisateurs le plus important possible. Mais le temps nécessaire à un vote « manuel », explicite, constituerait un frein trop important pour la plupart des visiteurs, car il s'agit d'attribuer un vote à chaque document visité. Seul un classement automatique et implicite permet d'avoir un retour de chaque visiteur pour chaque page, s'accorde enfin l'équipe<sup>150</sup>.

Georg, Jiechi et Szymon s'intéressent donc de plus en plus, dans la lignée des « géants » de la recherche en ligne, à l'approche du Web implicite (*Implicit Web*) : il s'agit de l'analyse du comportement des utilisateurs pendant la navigation ou la recherche et la récolte des « traces » que ceux-ci laissent au cours de leur voyage à travers le Web ; celles-ci peuvent être comprises comme un « vote » de préférence implicite, automatique, sans action manuelle de la part de l'utilisateur (Goldstein, 2005 ; Wilson, 2006 ; Guillaud, 2007 ; Ertzscheid, 2007). Georg me renvoie à Alex Iskold, commentateur de ReadWriteWeb, pour m'expliquer la manière dont il envisage d'incorporer le Web implicite à l'architecture de Faroo :

*The basic concept of the Implicit Web is simple. As we touch information, we vote. When we come across an article we like, we spend time reading. When we like a movie, we recommend it to our friends and family. And if a piece of*

---

<sup>148</sup> Entretien avec Jiechi Chang.

<sup>149</sup> Entretien avec Hiba, 13 février 2011.

<sup>150</sup> Notes de travail Faroo, 2008 et entretien avec Georg Schneider.

*music resonates with us, we listen to it over and over again. We do this automatically, subconsciously or implicitly. But the consequences of our behavior are important. The things that we are paying attention to have great value to us, because we like them*<sup>151</sup>.

Voulant se servir du web implicite pour diriger le crawler vers les endroits qui intéressent les usagers, afin de sélectionner, classifier et personnaliser les résultats en se basant sur l'attention que les usagers donnent aux contenus visités, l'équipe a ensuite discuté de l'opportunité de programmer un nouvel algorithme ou d'en utiliser un qui existe déjà. Le choix tombe enfin sur PeerRank<sup>152</sup>, un algorithme d'indexation basé sur la mesure de l'attention : si un utilisateur passe un temps relativement long sur une page, la visite souvent, y applique un marque-pages ou l'imprime, cette page grimpera dans le classement du moteur de recherche. Plusieurs observations guident ce choix, dont une, faite par Szymon, qui touche au fonctionnement technique et à l'expérience de l'utilisateur : presque 40 % des recherches que les gens font sur Internet sont des doublons de requêtes qu'ils ont fait au moins une fois dans le passé. Faroo doit donc pouvoir assister l'utilisateur dans la tâche chronophage de garder ses informations à jour, et la collecte implicite des « données d'attention » permet d'identifier les requêtes susceptibles d'avoir une pertinence à long terme pour l'utilisateur<sup>153</sup>. De son côté, Hiba remet sur la table la question des connotations politiques et économiques d'un moteur de recherche qui fonctionnerait sur l'agrégation distribuée de votes implicites :

*Monitoring the browsing habits of the users and aggregating those “implicit” votes across the whole web promises a more democratic and timely ranking, (and t)his also eliminates shortcomings of a Wikipedia like approach where content is contributed in a highly distributed way, but the audit is still centralized. Implicit voting automatically involves everybody in a truly democratic ranking, as the groups of adjudicators and users become*

---

<sup>151</sup> Alex Iskold, « The Implicit Web : Last.fm, Amazon, Google, Attention Trust » ReadWriteWeb, juin 2007 ([http://www.readwriteweb.com/archives/the\\_implicit\\_web\\_lastfm\\_amazon\\_google.php](http://www.readwriteweb.com/archives/the_implicit_web_lastfm_amazon_google.php)).

<sup>152</sup> PeerPong Blog, « A Closer Look at PeerRank », <http://blog.peerpong.com/post/816152003/a-closer-look-at-peerrank>: « PeerRank (technology) is made up of our proprietary semantic, natural language processing (NLP) indexing algorithms, which determine a numerical weighting for a person's knowledge and expertise. That is paired with our massive ontology (fancy word for the representation of a body of knowledge) of “n-grams,” which associates groups of key words and phrases with a common idea or topic – like ski conditions, for example. The “n-gram” for ski conditions might include words like: snow, sun, mushy, humidity, icy, powder, etc. The combo of the algorithms and ontology lets us score and rank individuals according to their demonstrated interest and knowledge on any given topic. People demonstrate their interest and expertise with their Twitter and Facebook status updates (the relative importance and relationships of updates is part of our secret sauce). In other words, the PeerRank is a huge list of people, cross-referenced by a huge list of topics, and then ranked by relevance. (... PeerRank's fast growth means that) every month, our ability to find the very best person for your question gets better and better. As we add people to the PeerRank, we can find more suited people to answer a broader variety of inquiries. »

<sup>153</sup> Entretien avec Szymon.

*identical*<sup>154</sup>.

En proposant une définition de la « démocratie » basée sur des votes implicites et automatiques, qui repose sur l'action de l'utilisateur, l'équipe Faroo fait donc de son moteur de recherche un système « qui fonctionne sur les clics<sup>155</sup> » : un dispositif qui prend note d'un dé clic de souris, du temps passé sur une page, d'un copier/coller, et le traduit en votes de préférence des utilisateurs. Les dé clics, le temps et les actions des utilisateurs deviennent des *inputs* fournis à un algorithme d'optimisation qui, en assignant aux différents paramètres des poids différents, produit un résultat qui établit combien l'utilisateur a apprécié un contenu spécifique.

Par ailleurs, les développeurs se trouvent ici face à un problème de taille, car la collecte et la mise à profit des traces plus ou moins implicites que l'utilisateur laisse lors de sa navigation, pour une meilleure personnalisation de l'expérience de recherche, les confrontent à la gestion technique d'un autre aspect de la recherche distribuée auquel les développeurs tiennent beaucoup : la protection de la *privacy* de l'utilisateur (Shaker, 2006 ; Zimmer, 2008). Il s'agit là d'une des principales tensions à laquelle l'équipe a dû faire face depuis le début de sa démarche<sup>156</sup>.

#### PAS DE « SUPER-PAIRS », OU UN AMORÇAGE DISTRIBUÉ

Le processus d'amorçage initial des nouveaux pairs (*bootstrap*) est l'autre aspect, concernant la décentralisation du dispositif, sur lequel l'équipe Faroo a beaucoup travaillé en 2009 – dès que la phase *beta* a pris fin et qu'il a été question de l'incorporation au système d'une masse, potentiellement importante, d'utilisateurs. La première fois qu'un nouveau pair veut se connecter à un réseau P2P, il doit prendre contact avec les pairs connus (appelés « super pairs », ou « pairs racine », ou encore pairs « de rendez-vous ») afin de connaître les adresses des autres. Les adresses des pairs connus sont envoyées à ce nouveau pair dans une liste avec le logiciel client, ou sont téléchargées dynamiquement à partir de caches web. Le nouveau pair stocke les adresses des pairs qu'il a appris des super pairs, et pourra donc se connecter par la suite directement à ces adresses, sans contact *a priori* avec les super pairs. Mais si un pair a été hors ligne pendant un certain temps, la plupart de ses adresses stockées deviennent invalides, puisqu'il s'agit d'adresses IP dynamiques. Si le pair ne parvient pas à se connecter au réseau P2P en utilisant les adresses stockées, il commence à nouveau le processus d'amorçage, à l'aide des super pairs.

Pendant une période de forte croissance initiale du réseau (comme peut l'être celle pendant laquelle le dispositif est ouvert pour la première fois à des utilisateurs sans restrictions de statut), de nombreux pairs accèdent aux super pairs afin de se

---

<sup>154</sup> Entretien avec Hiba.

<sup>155</sup> « Powered by clicks » : entretien avec Georg Schneider.

<sup>156</sup> Cette tension, et la controverse qui s'en est ensuivie, feront l'objet de la section « Personnalisation et privacy » ci-dessous.

connecter au réseau P2P : ainsi, le super pair devient le « goulot d'étranglement », et un point de défaillance possible dans un système par ailleurs totalement décentralisé. Si ces pairs sont surchargés, de nouveaux pairs ne peuvent pas non plus rejoindre le système, empêchant ainsi la croissance du réseau. Aussi, au cas où l'ensemble du réseau P2P se décompose en raison d'un incident web et où tous les pairs essaient de se reconnecter en même temps, il existe un risque de surcharge pour les super pairs, qui limite une reprise rapide (car les pairs n'arrivent pas à se connecter mais continuent à essayer), et provoque une charge supplémentaire pour le réseau. Georg et son équipe ont à ce propos pu bénéficier des leçons dérivées des problèmes rencontrés par le populaire réseau P2P hybride de téléphonie et voix sur IP, Skype, qui en août 2007 a été en mode hors connexion pendant « deux longs, incroyables jours » à cause d'une surcharge du réseau<sup>157</sup>.

De longues discussions dans l'équipe de travail ont porté sur le fait que la présence de super pairs dans le réseau le rendait dans l'ensemble vulnérable, à cause de leur caractère centralisé. En effet, si les super pairs permettent plus facilement de se libérer du serveur comme unique point de passage obligé du trafic du réseau, ils ne restent pas moins, à leur tour, points de passage obligés. Un « mal nécessaire », selon Szymon et Jiechi, pour assurer la stabilité du dispositif dans sa phase d'amorçage ; pour Georg, cela constitue cependant une limite, car la surveillance et éventuellement le blocage de l'ensemble du réseau P2P deviennent possibles simplement en bloquant ou en observant ces quelques nœuds de première importance :

*« [these peers still qualify as a] bottleneck and single point of failure in an otherwise distributed system »<sup>158</sup>.*

Les ingénieurs de l'équipe ont donc commencé à travailler à un algorithme d'amorçage distribué, permettant d'éliminer les super pairs comme derniers éléments centraux et

*« provide an organic scaling for the bootstrap procedure, ensure a fast recovery in case of a system wide incident, make the P2P network immune to the blocking or monitoring of super peers »<sup>159</sup>.*

« COMME DES CELLULES », OU LA MISE À JOUR DISTRIBUÉE

Le « sixième niveau de distribution », se remémore Georg, n'était pas du tout envisagé lors des premières phases de conception du dispositif, les développeurs s'étant concentrés en priorité sur les façons d'agréger les données relatives aux

---

<sup>157</sup> Todd Georg, « Skype Failed the Boot Scalability Test : Is P2P Fundamentally Flawed ? », *HighScalability*, <http://highscalability.com/skype-failed-boot-scalability-test-p2p-fundamentally-flawed>

<sup>158</sup> Entretien, Georg Schneider.

<sup>159</sup> Id.

préférences et aux traces d'intérêt laissées par les usagers, afin de rendre les résultats de recherche plus pertinents, et en conséquence, le système plus fiable. Cependant – et encore une fois, c'est l'ouverture à une masse potentiellement importante d'usagers qui impose des nouvelles questions et change les priorités – la question de la sécurité et de l'intégrité du réseau, associée à la nécessité de fréquentes mises à jour du client P2P, s'impose bientôt. Comme remarque un utilisateur sur le forum Faroo :

*« Sure, the distributed system becomes automatically smarter just by the increasing relevance of the collected attention data, but at any point you may want to refine the underlying algorithms, to improve the efficiency of the P2P overlay, to extend the data model, or to add new functions, the example of Windows shows that it might be necessary to apply security patches, network wide, frequently and immediately. And updating P2P clients requires a very efficient software update distribution. »*<sup>160</sup>

Jiechi calcule que, une fois le système fonctionnant à plein régime, des mises à jour conduites de façon centralisée utiliseraient trop de temps :

*« 10 million peers and 5 Mbyte client software size would require the system to distribute 50 Terabytes for a full network update. Even for a 100 Mbit/s network connection a central update would last 50 days, if you manage to evenly distribute updates over time »*<sup>161</sup>.

Encore une fois, cette incompatibilité vient du fait d'avoir poussé à fond l'approche distribuée pour les autres composantes du système : puisque le système est construit pour pouvoir facilement passer à l'échelle, et que chaque utilisateur/pair du système a installé un logiciel client, chacun de ces clients est susceptible d'avoir besoin de nombreuses et fréquentes mises à jour. La solution envisagée par l'équipe Faroo est donc de répartir la transmission des mises à jour sur un modèle, lui aussi, distribué. Les pairs se transmettraient l'un l'autre, « *cell division like* » (de façon similaire à la reproduction cellulaire, dit Georg)<sup>162</sup>, les mises à jour – les délais de mise à jour se trouvant de ce fait considérablement réduits.

Distribuer le système de mises à jour implique, par ailleurs, l'intégration à ces mises à jour d'une étiquette ou une « signature » que le système peut reconnaître, légitimant ainsi le paquet, avant d'autoriser la mise à jour<sup>163</sup>. Faroo n'est pas imperméable aux vulnérabilités qui affectent généralement les systèmes P2P : en particulier, les codes non sécurisés et non signés qui sont susceptibles d'autoriser l'accès à distance aux ordinateurs du réseau, risquant ainsi de compromettre sa stabilité. La fonction même de « facilitateur d'échanges » qui est une des forces majeures du P2P rend aussi très difficile le contrôle de la propagation de virus ou de vers. Cette faille préoccupait d'ailleurs beaucoup certains utilisateurs lors de l'annonce, de la part des

---

<sup>160</sup> Post de forum Faroo, avant la migration sur forum.faroo.com [sauvegardé par moi].

<sup>161</sup> Notes de travail Faroo, 2008.

<sup>162</sup> Entretien, Georg Schneider.

<sup>163</sup> « *Some signature stuff to ensure the integrity of the network* », id.

développeurs, du système de mises à jour distribuées. En avril 2008, l'utilisateur Darren Stuart donne voix à ces doutes :

*« Spam, whats [sic] to stop spammers from hooking up a ton of bots and just surfing spam? »*<sup>164</sup>

La réponse de Georg dans cette même discussion reconnaît la difficulté de la tâche tout en soulignant que le problème ne se pose pas seulement pour les mises à jour, mais plus généralement pour la qualité et la pertinence des résultats obtenus par les utilisateurs lors des requêtes – et que dans les deux cas, il est possible de faire face à ce problème en associant, dans l'algorithme de recherche distribuée, des critères de classement basés sur l'attention de l'utilisateur et l'analyse des liens entrants et sortants d'une page, afin d'en évaluer le niveau de confiance :

*« We are aware of the fact that for every search engine spam is a major issue. Spam fighting is done mainly by ranking, not index filtering. This addresses consistently also the levels of gray, which are between spam and high quality content.*

*While we use attention for ranking, we still use link analysis for trust. Additionally, by statistical sampling we can keep our attention based ranking unbiased, even if we would have 50% rogue peers. »*<sup>165</sup>

En dévoilant progressivement la manière dont la décentralisation et la distribution sont agencées dans les différentes « couches » de l'architecture de Faroo, cette section a souhaité montrer comment l'équipe de développeurs du dispositif configure le P2P comme un point de force dérivant de la flexibilité du système, plutôt que comme une faiblesse dérivant de sa fragilité ou de son statut transitoire. Dans la poursuite de cet objectif, il s'avère nécessaire pour l'équipe Faroo de travailler avec la distribution à plusieurs niveaux, incluant les aspects *opérationnels*, c'est-à-dire tout ce qui concerne le fonctionnement en profondeur du dispositif, au delà de la fonctionnalité principale de recherche – qui est la plus visible à l'utilisateur.

---

<sup>164</sup> Commentaire de l'entrée blog de B. Lunn, [http://www.readwriteweb.com/archives/faroo\\_could\\_p2p\\_search\\_change\\_the\\_game.php#comment-110471960](http://www.readwriteweb.com/archives/faroo_could_p2p_search_change_the_game.php#comment-110471960).

<sup>165</sup> Id. à [http://www.readwriteweb.com/archives/faroo\\_could\\_p2p\\_search\\_change\\_the\\_game.php#comment-110471971](http://www.readwriteweb.com/archives/faroo_could_p2p_search_change_the_game.php#comment-110471971)

### SECTION 3. LES « SIX NIVEAUX DE DISTRIBUTION » ET LA CENTRALITÉ DE L'USAGER

Avec sa recherche « six fois distribuée » – l'indexation, la découverte d'information, la recherche, la classification, l'amorçage, la mise à jour du logiciel – analysée dans la section précédente, Faroo travaille à formater un réseau décentralisé capable de se libérer, aux différents niveaux qui composent son architecture, de la dichotomie entre le client et le serveur. Les pairs sont mobilisés en tant que composants d'un système sociotechnique reliant les usagers, les algorithmes et les ressources techniques mises à disposition. Le fait de ne pas considérer les pairs comme des avant-postes d'un système centralisé<sup>166</sup>, mais de « prendre au sérieux » leur statut d'égaux aux marges du réseau et de travailler à partir de cette conception, se présente comme un enjeu extrêmement délicat à ce stade de l'expansion du système. Comme le souligne le blogueur Brad McCarthy en 2010,

*« Peer-to-peer search has some huge advantages. Not being reliant on a single (or even multiple) data centers can clearly free up a lot of concern. Socially relevant results do indeed have their place for my own use, and I'm sure they would for others as well. [...] At the same time, Faroo] is not being marketed toward anyone. This leaves the entire future of the project (at this moment) in the hands of the user. [...] if you're relying on the world at large to be your network, and nearly refusing to entertain the idea of hybrid technology, you may be on the way to kill your company's shot. »<sup>167</sup>*

Cette centralité de l'ensemble des usagers apparaît à la fois comme la plus grande chance et le risque majeur auquel le dispositif Faroo doit faire face : il s'agit d'un système qui ne comporte pas d'importants investissements de ressources de la part de l'entreprise, et qui est en principe suffisamment flexible pour travailler à la fois pendant des situations stables et instables, voire mobiles. Le système est construit pour se ressaisir de manière autonome après un événement qui en aurait modifié la structure : en absence d'une entité centrale, les pairs « survivants » sont faits pour se retrouver et assurer, en des délais courts, la recombinaison d'une entité distribuée pouvant encore assurer ses fonctions<sup>168</sup>.

En même temps, la centralité de l'utilisateur et des ressources que celui-ci apportera à

---

<sup>166</sup> Certains usagers considèrent que tel est le cas de Skype : « *a new node in Skype goes through 2 big steps before it fully joins the network. The first is that it has to authenticate to the central servers, and the second is that it has to find a good "supernode" to initialize/bootstrap and maintain connectivity to the central servers. (...) due to the synchronized restart of clients, when all the nodes started banging on the central servers, the supernodes started losing connectivity and dropped out of the "supernode" mode. Creating a chicken and egg problem and making the situation worse than before. Is it possible that the real reason why Skype network failed was because of its dependence on central servers ?* » Voir la discussion du post cité dans la note précédente.

<sup>167</sup> Brad McCarthy (2010), « A look at Faroo's P2P Search. Will this one make it? », <http://thenextweb.com/apps/2010/05/25/p2p-search-is-here-but-who-will-use-it/>.

<sup>168</sup> Georg reprend ici la métaphore des cellules : « *In biology organisms naturally deal with the rise and falls of its cells; simple elements form superior systems. We believe that evolution works in search too, and that the future belongs to multi-cellular organisms/systems.* » (White Paper, 22).

l'ensemble du système peut aussi être problématique pour la durabilité du dispositif, particulièrement dans ses phases initiales. Faroo « compte sur le monde entier pour être son réseau », comme l'a résumé de façon efficace Brad McCarthy<sup>169</sup>. Cette posture amène à prendre un gros risque, car parvenir à intéresser un nombre important d'utilisateurs n'est pas seulement une question de « réclamer sa part du gâteau » dans le marché (en tant qu'outil suffisamment répandu), mais est quelque chose de vital pour le bon fonctionnement technique du dispositif lui-même, et pour sa croissance et stabilité au fil du temps : pas d'utilisateurs, pas de ressources, de tous points de vue (Image 3).

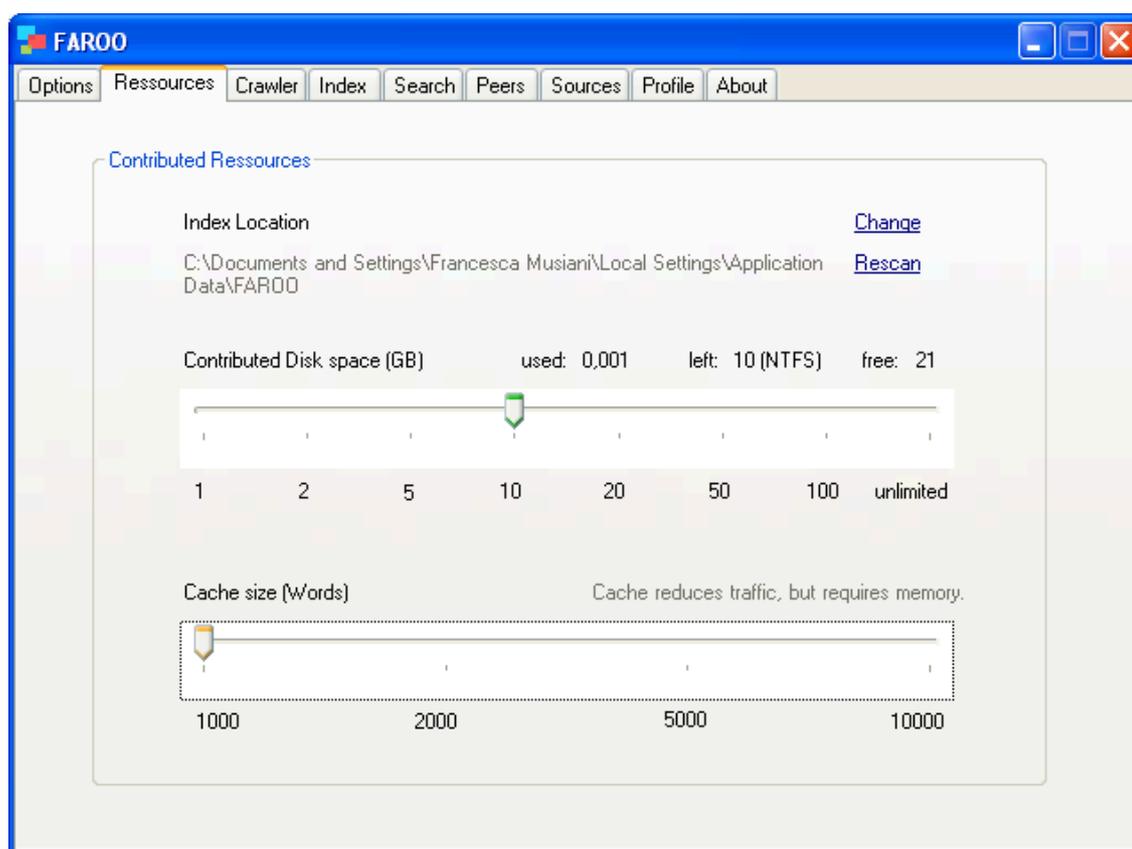


Image 3. Capture d'écran du panneau de configuration, qui permet à l'utilisateur de voir et surveiller ses contributions de ressources au moteur de recherche distribué.

En revanche, davantage d'utilisateurs signifie non seulement plus de ressources matérielles (mémoire, CPU, bande passante) à disposition de l'ensemble, mais aussi, on l'a vu, un affinement croissant de la variété et de la fiabilité des résultats obtenus par chacun d'entre eux, et une meilleure expérience utilisateur pour l'ensemble du réseau. Cette tension provoque plusieurs points de discussion et de controverse au

<sup>169</sup> McCarthy, *ibid.*

cours des premières années d'existence de Faroo.

QUI VIENT EN PREMIER, L'ŒUF (L'INDEX) OU LA POULE (LES USAGERS) ?

La pertinence et la précision des résultats pouvant être obtenus pendant la phase d'amorçage – donc, leur « qualité » dans la définition des développeurs de Faroo<sup>170</sup> – est le plus disputé de ces points : la quantité, et, conséquence directe, la qualité de l'index de recherche en P2P dépendent de la quantité de personnes/pairs qui utilisent Faroo. Pendant la première phase du *beta*, en 2007, la qualité de l'index est qualifié de problème principal par les utilisateurs de la première heure, au point de compromettre, dans certains cas, l'expérience utilisateur :

*« the P2P search index scales with the number of people using FAROO. It looks like a show-stopper. As the number of searchers is very low right now, people are likely to try it, get lousy results and forget about it<sup>171</sup>. »*

Le fait que l'indexation des sites se limite aux pages qui ont réellement été visitées par un ou plusieurs pairs est une façon de limiter le spam et d'améliorer la qualité des résultats, mais avec peu d'utilisateurs, peu de résultats sélectionnés aboutit souvent, en pratique, à un résultat décevant ou même à une absence de résultat :

*« When you go to perform a search, Faroo will only search sites that real people have visited. This is both the service's strength and its weakness. You don't get a lot of junk sites this way. But since Faroo is in private beta with a limited user pool, you don't get many sites at all that way. A major letdown. »*

Un autre utilisateur résume ce dilemme en le renvoyant, comme l'avaient déjà fait les développeurs, à la « vieille question de l'œuf et de la poule » :

*« as with any social/user-generated software, there's a chicken and egg problem. It's hard to market a product that has a small user base, because it just doesn't perform very well yet. At the same time, it's hard to attract new users with a product that doesn't deliver right away all it promises.<sup>172</sup> »*

Un autre site d'évaluation envisage entre les lignes une autre solution, tout en soulignant qu'elle devrait être temporaire et transitoire si la valeur novatrice de l'index P2P doit être maintenue :

---

<sup>170</sup> Pour les développeurs de Faroo, la qualité des résultats est liée en tout premier lieu à la quantité d'utilisateurs qui font part du réseau de recherche, car elle se construit à partir de leurs contributions – et plus celles-ci sont, plus les résultats peuvent être « affinés » progressivement. La définition de la « qualité » des résultats d'un moteur de recherche est par ailleurs un point sur lequel beaucoup a été écrit et continue à être écrit. Voir à titre d'exemple [Simonnot, 2008].

<sup>171</sup> McCarthy, *ibid.*

<sup>172</sup> Les extraits de ce paragraphe sont issus d'une discussion de la phase *private beta* de Faroo, sur un site d'évaluation de logiciels en beta, SeaDVD (<http://beta.seadvd.com/faroo-follow-up-p2p-search-engine-will-only-work-with-more-users/>).

*« Mahalo<sup>173</sup> displays search results from different search engines. Until they don't have many users logged in at one time, why not ? »<sup>174</sup>*

Face à la possibilité qu'une majorité des utilisateurs pionniers perdent rapidement tout intérêt pour le projet, l'équipe Faroo revoit ses priorités immédiates : le but principal dans la phase de « recrutement » d'une masse critique d'utilisateurs<sup>175</sup> doit être le passage à l'échelle de l'index, pour une amélioration rapidement visible de la qualité des résultats. De façon temporaire, la solution technique de réalisation plus immédiate est celle d'agréger, et de montrer à l'utilisateur, des résultats obtenus par d'autres moteurs de recherche en plus des résultats obtenus grâce à Faroo. Comme le souligne Georg :

*« (such a solution) is not by itself a big deal and used by other alternative engines, and I wasn't thrilled. But we had to get stuff that worked immediately, if we wanted to climb on the shoulders of our users, so to speak – and, by getting some results, make the next results better. So, we made a tool with which at some point you could get some results no matter what you were looking for, and you were more willing to continue to use the software. In the meantime, our index was scaling faster<sup>176</sup>. »*

La solution temporaire produit quelques effets bénéfiques pour la motivation des utilisateurs, et le « recrutement » d'utilisateurs reprend : en mars 2008, les contributeurs de SeaDVD en arrivent à la conclusion qu'« ils ont effectivement implémenté des mesures pour y remédier<sup>177</sup> », tandis que ReviewSaurus reporte triomphalement que

*« they took our suggestion and have started to display the search results from Yahoo or Microsoft Live, Wikia and few others if at all Faroo can't deliver the search result on its own<sup>178</sup>. »*

Certains utilisateurs apprécient le fait que l'équipe de développeurs soit à l'écoute de son public potentiel (*« This also shows that they listen to their customers and are ready to customize the application as their user wants »<sup>179</sup>*), tandis que d'autres

---

<sup>173</sup> <http://www.mahalo.com/>

<sup>174</sup> ReviewSaurus, « Faroo : P2P Web search engine, impressive idea! », janvier 2008, <http://www.reviewsaurus.com/web-applications-reviews/faroo-p2p-based-search-engine-impressive-idea/>

<sup>175</sup> Il ne semble pas déplacé de rappeler ici que certains auteurs doutent que ce recrutement et intéressement des usagers soit effectivement possible à l'heure actuelle, étant données les caractéristiques de « monopole naturel » du marché de la recherche en ligne. Voir [Pollock, 2009] et [Rieder & Sire, 2012].

<sup>176</sup> Entretien, Georg Schneider.

<sup>177</sup> SeaDVD, cit.

<sup>178</sup> ReviewSaurus, « Faroo opens for public and takes our suggestion ! », mars 2008, <http://www.reviewsaurus.com/software-reviews/faroo-opens-for-public-and-takes-our-suggestion/>

<sup>179</sup> Id.

invitent explicitement des usagers potentiels, parmi les lecteurs du blog, à joindre le système, en rappelant que cela peut se faire avec différents degrés d'engagement, à différents moments, tout en aidant le système à démarrer.

Dès le début, même en tant qu'utilisateur passif (c'est-à-dire en installant le client P2P Faroo sans pour autant lancer des interrogations de recherche), un pair contribue déjà à apporter des ressources à un système qui a grand besoin de se globaliser :

*« a suggestion for all of you too, go ahead and install Faroo and use it as much as you can, it's high time and we need an alternative of Google and you can help in doing that by just running this client. »*<sup>180</sup>

Il est donc possible – suggère ce blogueur – si une fois qu'on a installé le client on n'est pas satisfait de la performance du moteur de recherche, de le laisser agir à l'arrière-plan pendant quelques mois et d'y revenir ensuite, quand plus de pairs auront rejoint le système et que l'index aura augmenté de taille et de complexité. Cela mérite de faire l'essai, estime ReviewSaurus, d'autant plus que selon son expérience, la quantité de ressources mobilisée par le client Faroo ne réduit pas le bon fonctionnement du terminal de l'utilisateur, ce dont on pourrait douter pour un logiciel fonctionnant en « local ».

Il s'agit là d'une mesure de « politesse technique » qui plaît aux utilisateurs/recruteurs. Le blogueur de SeaDVD, qu'on a déjà cité, renchérit :

*« the client doesn't seem to eat much RAM, so you could always install it now, go about your business and check out the search engine feature in a few weeks or months. »*<sup>181</sup>

Du côté de l'équipe Faroo, la question de la gestion « courtoise » des ressources informatiques des utilisateurs est reprise avec espièglerie par Susanne, la linguiste et responsable de communication de l'équipe Faroo, lors d'une importante réécriture et remise en forme du site web qui se passe en 2010 :

*« FAROO does not require a dedicated crawler [...] so that there is no additional load for you. [...] FAROO is a well behaving search engine. »*<sup>182</sup>

## LE TÉLÉCHARGEMENT DU CLIENT P2P, UNE BARRIÈRE À L'ENTRÉE ?

Un deuxième point de controverse concerne le statut du client<sup>183</sup> P2P, que les

---

<sup>180</sup> Id.

<sup>181</sup> SeaDVD, cit.

<sup>182</sup> Faroo FAQs, <http://www.faroo.com/hp/p2p/faq.html#monitoring>

<sup>183</sup> Pour clarifier toute ambiguïté sémantique, on souhaite préciser qu'on utilise le mot "client" dans cette section pour indiquer le logiciel que l'utilisateur télécharge sur sa machine afin que celle-ci participe au réseau P2P – et non dans l'acception d' "acheteur d'un service ou d'un bien".

utilisateurs de Faroo doivent installer comme précondition à l'entrée dans le système. On a déjà brièvement expliqué comment le client P2P acquiert, dans la démarche des membres de l'équipe Faroo, une grande importance, en tant qu'endroit matériel et dispositif technique où se déroule le cœur des opérations d'indexation et de recherche, et en tant que symbole de la « centralité distribuée » de l'utilisateur dans le système. Cependant, à la sortie du *closed beta* (version relativement stabilisée, accessible sur invitation des développeurs, début 2008), un argument revient fréquemment parmi les réactions des usagers, qui invite à reconfigurer ultérieurement le client P2P, à la lumière de deux tendances récentes dans le marché des services Internet : la montée en puissance du « *cloud* » et le déplacement de plus en plus d'utilisateurs vers des connexions en mobilité.

*Faroo, being P2P, requires a client download. That can be an incredibly big hurdle to adoption. At first I thought your best shot was to build on top of existing P2P services such as LimeWire or Gnutella<sup>184</sup>. But your strategy appears to be different as you are building on top of .Net.... Oops there go all the Mac and Linux heads... bye. This is potentially a serious issue as lots of early adopters tend to be on the Mac. If I am a Mac user, then I may not use Faroo without borrowing somebody's PC<sup>185</sup>.*

Le fait de développer le client P2P sur .Net (un kit de composants logiciels structurels développés par Microsoft et par conséquent étroitement liés à l'environnement Windows) semble en effet poser un problème de taille à certains utilisateurs pionniers. Fabian remarque tout simplement :

*No Mac, No Linux? Then not for me<sup>186</sup>.*

Samidh, de façon moins lapidaire mais également critique, note que la multiplication des dispositifs reliés à l'Internet désormais possédés par les utilisateurs, et l'actuelle tendance à rendre toujours plus « légères » et mobiles les applications installées sur leurs terminaux, pourraient poser de problèmes supplémentaires pour la diffusion d'une application qui doit être obligatoirement téléchargée :

*« I think the biggest obstacle is that they require a download. Startups need to realize that the age of the download is dead. Skype was the end of an era. People are far too mobile and use too many computers to worry about specific downloads anymore. Everybody now needs to keep repeating the mantra, 'The Internet is the Platform'.<sup>187</sup> »*

Et Danah Boyd, spécialiste des réseaux sociaux pour Microsoft, contribue à présenter l'obligation d'installation d'un client P2P comme « poids » sur les épaules des usagers potentiels, plutôt que comme opportunité de sauvegarde et de gestion

---

<sup>184</sup> Voir chapitre 2, section 2.

<sup>185</sup> Discussion de l'entrée Web de B. Lunn, cfr.

<sup>186</sup> Id.

<sup>187</sup> Id.

différentes des données circulantes dans le système :

*« I'm also curious about the future of peer-to-peer systems in light of the move towards the cloud, but I'm not convinced that decentralization is a panacea to all of our contemporary woes. Realistically, I don't think that most users around the globe will find a peer-to-peer solution worth the hassle. The cost/benefit analysis isn't in their favor.<sup>188</sup> »*

Le symbole de l'articulation entre les dimensions locale et globale du système, « vendue » par l'équipe Faroo comme un des points novateurs et originaux du système, se transforme dans les réactions de certains usagers et commentateurs en un obstacle, une barrière à l'entrée.

Le développeur P2P de l'équipe Faroo, Jiechi, signale à son tour les difficultés que certains des prédécesseurs de Faroo ont dû gérer, et auxquelles il doit faire face :

*« Indeed, it is not easy even for successful ones, look at Joost<sup>189</sup> – they've been forced to move to a pure browser model because nobody took the download plunge. The only people who can successfully get downloads distributed are people who already have wide distribution. »<sup>190</sup>*

Le téléchargement du client P2P est par ailleurs, dans la vision de Georg, un des aspects incontournables et distinctifs d'un moteur décentralisé, la façon dont l'utilisateur peut garder à l'œil le trafic entrant et sortant de son client P2P – trafic indiquant à la fois qu'une interrogation est en train d'être indexée par le pair qu'il représente, et que ce pair est en train de répondre à des interrogations. Il formule à ce propos un intéressant renversement de perspective, une tentative de « traduction » – et définit le client P2P en termes de bénéfice pour l'utilisateur :

*« There have been the SecondLive client, the StumbleUpon toolbar and Oovoo after Skype, to name a few. With the bandwidth rise downloads are not that uncomfortable anymore. [...] 'The Internet is the Platform'. So do we, referring to the decentralized structure of the internet, including users as part of the internet, not only reducing it to a bunch of central servers ;- ) I believe that people will jump over this 'download hurdle' if they get something in return, which they can't get in a centralized system...»<sup>191</sup>*

Point de vue partagé et clairement expliqué par Bernard Lunn, qui deviendra plus tard consultant pour Faroo, et qui répond à Danah Boyd ainsi :

---

<sup>188</sup> Intervention de Danah Boyd sur la liste de diffusion de la P2P Foundation, 27 mai 2010 ([http://listcultures.org/pipermail/p2presearch\\_listcultures.org/2010-June/008941.html](http://listcultures.org/pipermail/p2presearch_listcultures.org/2010-June/008941.html)).

<sup>189</sup> Joost (nom de code The Venice Project) est un projet de Niklas Zennström et Janus Friis (les fondateurs de KaZaA et Skype), dont le but est de créer un logiciel permettant la distribution de programmes de télévision et d'autres formes de vidéos sur le web en utilisant la technologie P2P.

<sup>190</sup> Entretien avec Jiechi Chang, cfr.

<sup>191</sup> Discussion de l'entrée Web de B. Lunn, cfr.

*« The download hurdle is a common reservation. It has not proved a hurdle when the payoff is big enough – think Skype and Spotify. And more recently the iPhone and Android app stores have totally changed the mindset around downloading software. The key issue I think is what you mention – “hassle”. It is a hassle free experience on iPhone but people have found the experience less than hassle-free on the PC. »*<sup>192</sup>

Selon les mots de Georg et plus tard de Bernard Lunn, le téléchargement d'un client qu'on héberge sur son propre terminal, et qui assure les connexions P2P, est une opportunité de marché, le symbole d'une autre façon de concevoir la recherche d'information sur le Web et peut-être, dans un deuxième temps, l'Internet *at large*. En effet, il s'agit d'incorporer l'interopérabilité des applications P2P avec le reste du réseau plus en profondeur dans celui-ci, au lieu de rendre les systèmes P2P des « îles » ne s'intégrant qu'avec difficulté :

*« The first P2P application was file-sharing, the next Skype, now we are building P2P search. What could be the next? Instead of another isolated P2P application, we would like to see P2P built into the OS and Internet stack in a standardized manner. »*<sup>193</sup>

Point de vue partagé par un utilisateur (et développeur à son tour), Sam Rose, qui à l'occasion de la discussion entre Danah Boyd et Bernard Lunn, définit la question de l'acceptabilité du téléchargement du client P2P comme l'investissement nécessaire d'un peu de temps et d'énergie de la part de l'utilisateur, dans le cadre d'une intégration plus « globale » du P2P dans l'infrastructure Internet :

*« This all boils down to the way the application is made. In the case of skype, the control of sharing is automatic. I assume it is the same in spotify. If control of sharing is automatic, I believe many people will find cost/benefit in their favor (as they have with skype). Indeed it is not simple to do this. But, in the end, if p2p is to become part of our computing infrastructure, it will be a necessary investment. We have a chance now to think about how to make this very open and accessible. »*<sup>194</sup>

Les évolutions du marché de la bande passante – notamment le fait qu'on peut, d'année en année, obtenir plus de bande pour un prix moins élevé – seraient en train d'ouvrir les portes à la dynamique du téléchargement, renchérit Georg, au moment même où la mobilité et le *cloud* peuvent l'entraver. Ce qui fera pencher la balance d'un côté ou de l'autre est la valeur ajoutée que le système décentralisé sera capable de fournir, à la fois en termes de modèle économique et d'expérience utilisateur :

*« ... this will be privacy protected personalization, attention based ranking,*

---

<sup>192</sup> Intervention de Bernard Lunn sur la liste de diffusion de la P2P Foundation, 27 mai 2010.

<sup>193</sup> White Paper, p. 30.

<sup>194</sup> Intervention de Sam Rose sur la liste de diffusion de la P2P Foundation, 13 juin 2010.

*and revenue sharing based on the saved infrastructure costs<sup>195</sup>. »*

« LÉGALITÉ » DU DISPOSITIF, ENTRE RESPONSABILITÉ INDIVIDUELLE ET PARTAGÉ

Un troisième point de discussion parmi les usagers, et une question qui est souvent posée aux membres de l'équipe, concerne le statut juridique du dispositif et par conséquent, de ses utilisateurs/pairs. Deux points, principalement, préoccupent ces utilisateurs de la première heure : les techniques d'anonymisation et d'encryptage utilisées dans Faroo<sup>196</sup> sont-elles susceptibles de favoriser des usages frauduleux, comme la propagation d'informations inexacts ou propriétaires, et si oui, en quoi la mise en œuvre de ces usages illicites de la part d'un utilisateur pourrait-elle entraîner la responsabilité d'un ou plusieurs autres utilisateurs qui, comme lui, auraient mis à disposition une partie de leurs ressources de calcul dans un « cerveau de recherche » distribué dont les composants sont, par design, difficilement séparables et identifiables ?

Ces doutes, raconte Georg, avaient été en partie anticipés par l'équipe. Le statut de dispositif pair à pair de Faroo (bien que « sagement minimisé par moments », note au passage le développeur) faisait du système une cible facile pour des questionnements d'ordre légal et juridique. Certes, dans un système où les contenus « échangés » et distribués sont les interrogations des usagers et l'index qui se construit à partir d'elles, la question classique des « tristement célèbres » échanges en pair à pair de fichiers couverts par le droit d'auteur ne se pose pas. Cependant, l'équipe s'attendait à un certain nombre d'interrogations, venant principalement du public et de la presse non spécialisée, à propos de l'association fréquente entre P2P et partage de fichiers protégés. Ces interrogations amènent donc Faroo Ltd. à introduire, dans la troisième version des termes de licence à l'utilisateur (février 2009), le point suivant :

5. INDEMNIFICATION OF FAROO WITH REGARD TO SEARCH RESULTS GENERATED BY THE USER

5.1 With this software, FAROO merely provides a tool with which the user can create and run search routines. The software itself does not provide any content for the Internet. Consequently, FAROO assumes no responsibility for content that becomes accessible to the user on the screen or otherwise as the result of executing search routines. The same shall apply to any and all violations of rights of third parties on or by any pages displayed on the screen or otherwise made accessible to the end user during execution of search routines.

5.2 In relations with FAROO, the user alone shall bear (...) responsibility for the

<sup>195</sup> Entretien, Georg Schneider. Certains de ces aspects ont déjà été abordés, certains seront repris par la suite.

<sup>196</sup> Il s'agit des techniques de sauvegarde de la privacy mises en œuvre dans Faroo : voir la section suivante.

search results and content generated by him with the help of the software.

*Tableau 4. Clause 5 des termes de licence Faroo (à partir de février 2009).*

Peu après, la page des “Questions fréquentes” du site de Faroo s’enrichit d’une réponse à la question : « *Does FAROO’s anonymization and encryption endorse fraudulent use?* » :

*Responsible is solely the person who is publishing the content or consuming the content. FAROO (is) not responsible for a possible misuse of its technology, as the developer of a web server, a browser, the HTTPS-protocol or the AES-encryption, or the manufacturer of a monitor, a hard disk or a memory chip is not responsible for storing, transferring or displaying illegal content by this systems. (...) FAROO is not publishing content to the Internet, but makes it available or provides anonymous access to it (...) assist(s) each other locating of information which already exists in the Internet*<sup>197</sup>.

Bernard Lunn, au cours de la discussion sur la liste de la P2P Foundation, remarque la difficulté, voire l’impossibilité, d’attribuer une responsabilité individuelle à un pair dans un système comme Faroo, Skype ou d’autres systèmes distribués, dans lequel des ressources, matérielles et de calcul, sont mises à la disposition d’un « groupe » dont on ignore la composition. L’attribution de responsabilité qu’on pourrait obtenir en limitant ou fermant le réseau s’oppose au même principe de mise en commun des ressources :

*« I have heard this worry [un participant à la discussion avait suggéré que Faroo, comme d’autres systèmes décentralisés, aurait bientôt dû faire face à une masse de “déviant”], and what many people do not realize is that this could be happening now with systems such as skype, etc. In shared systems, individuals simply cannot be responsible for how the many are using those resources that they share. It is also possible to restrict your sharing to trust networks. This will decrease the resources available, but come closer to ensuring that you know who participants are. (This is only a worry when you are pooling resources for data storage and passing.) »*<sup>198</sup>

Les conditions d’utilisation, et les détails donnés dans la foire aux questions du site Faroo, sont pour Georg et son équipe une occasion de préciser le statut juridique, notamment la responsabilité, de Faroo et de ses usagers. Mais cette occasion devient aussi une possibilité de négocier et de définir ultérieurement ce que le dispositif fait, la nature des contenus qui y sont « échangés » ou montrés, et le rôle assumé par le service.

---

<sup>197</sup> Faroo FAQs, <http://www.faroo.com/hp/p2p/faq.html#anonymization>. Je souligne. De façon intéressante, cette formulation renvoie la responsabilité vers l’utilisateur, mais ne mobilise pas explicitement le statut d’hébergeur et le droit à l’« havre de paix » (Elkin-Koren, 2006) qui en découlerait.

<sup>198</sup> Intervention de Bernard Lunn sur la liste de diffusion de la P2P Foundation, 27 mai 2010.

Comme beaucoup de leurs prédécesseurs – et notamment Google – les développeurs de Faroo revendiquent une fonction d'« instrument » ou d'« outil » pour leur service : les résultats et les interrogations par lesquels les usagers contribuent au système, ou l'utilisent, n'engagent qu'eux-mêmes d'un point de vue juridique. Mais dans cette même formulation, les termes de licence clarifient ce que l'outil n'est pas fait pour faire... rappelant ce à quoi il *est* censé servir : le logiciel (et donc le fournisseur du service) n'est pas un éditeur de contenus, mais un facilitateur de mise à disposition, de recherche et de localisation de la part de l'utilisateur. Par conséquent, l'utilisateur peut créer et exécuter des interrogations de recherche, qui, en étant intégrées à l'index distribué, auront une influence sur les résultats de recherche obtenus par les autres utilisateurs.

C'est ici que, une fois établie la responsabilité de l'utilisateur vis-à-vis de Faroo, se profile l'idée d'une responsabilité « partagée » vis-à-vis du système – et donc, des autres utilisateurs. La centralité de l'utilisateur est, encore une fois, réaffirmée en lui attribuant la responsabilité de la durabilité et de la fiabilité du système. Par contre, les développeurs de Faroo mettent en avant l'idée d'un facilitateur bien plus discret et effacé qu'il ne l'est effectivement, en sous-estimant le rôle fondamental des intermédiaires et *gatewatchers* de l'Internet (Caillaud & Jullien, 2002 ; Elkin-Koren, 2006). Si les « six niveaux de distribution » contribuent à façonner une entité décentralisée, omniprésente, *pervasive*, la fonction d'intermédiation exercée par cette entité ne s'efface pas pour autant. Les développeurs eux-mêmes, d'ailleurs, fournissent à l'utilisateur les moyens d'explicitier et de donner visibilité à cette intermédiation : le logiciel est programmé pour proposer à l'utilisateur nombre d'instruments (dont les captures d'écran que parsèment ce chapitre ont voulu donner un aperçu) afin de monitorer les ressources qu'il met à disposition du système, les pairs auxquels son client se connecte, les requêtes qu'il héberge.

Cependant, ces instruments – au moment même qu'ils réaffirment la possibilité de contrôle de l'utilisateur sur les ressources et sur les dynamiques du système – sollicitent de l'utilisateur certaines compétences techniques, parfois poussées, afin de comprendre comment se façonnent les réponses aux interrogations qu'il pose au moteur Faroo. En posant des conditions à l'engagement des utilisateurs – notamment, en cherchant à « enrôler » de manière prioritaire le sous-ensemble d'utilisateurs qui a un intérêt spécifique pour ces dynamiques décentralisées – les développeurs contribuent à façonner des formes de solidarité technique (Dodier, 1995) complexes, qui ne font guère – contrairement à ce que la tendance à la décentralisation intégrale, et au recentrage sur l'utilisateur, pourrait laisser supposer – l'économie de l'intermédiation.

Par ces exemples, on a voulu montrer comment la centralité attribuée dans le dispositif à l'ensemble des usagers, et à leurs contributions au système, apparaît à la fois comme la plus grande opportunité et le risque majeur auquel Faroo doit faire face, particulièrement dans ses phases initiales. En « comptant sur le monde entier pour être leur réseau », les développeurs de Faroo sont conscients que parvenir à intéresser un nombre important d'utilisateurs est vital pour le bon fonctionnement technique du dispositif lui-même, et pour sa croissance et stabilité au fil du temps.

Les développeurs de Faroo se concentrent donc en priorité sur le fait d'attirer et de garder un nombre important d'utilisateurs, aussi bien en phase d'amorçage que, ensuite, « à plein régime », lors de son fonctionnement optimal. Mais que signifie développer un outil de « recherche centrée sur l'utilisateur », et comment peut-on communiquer cette idée à ses utilisateurs/ressources/pairs afin qu'ils s'investissent avec confiance dans un système, qui devient plus efficace précisément grâce leur propre contribution? La prochaine section s'attache à cette question, en analysant une des tensions centrales à laquelle l'équipe Faroo s'est attaquée dans les premiers pas du dispositif : l'équilibre entre personnalisation des requêtes et *privacy* de l'utilisateur.

#### SECTION 4. LIER PERSONNALISATION ET VIE PRIVÉE : LE P2P COMME « RECHERCHE CENTRÉE SUR L'USAGER » ?

En octobre 2007, un article sur ReadWriteWeb a d'importantes répercussions sur le développement de Faroo. Il y est question du « bruit et de la confusion qui surviennent quand une force impétueuse se confronte à un objet qui ne peut pas bouger », et des « opportunités entrepreneuriales » qui s'ensuivent<sup>199</sup>. Selon l'auteur de l'article, Bernard Lunn, les technologies de l'information se trouvent actuellement face à une tension ne pouvant être apaisée qu'au moyen d'importants changements et de projets novateurs, voir perturbateurs (*disruptive*) : la tension entre personnalisation et *privacy* de l'information et des données personnelles concernant l'utilisateur. La personnalisation est une force irrésistible, dit Lunn, pour des raisons qui vont de l'économie de l'attention et du temps de l'utilisateur (à qui seront proposés dans les plus brefs délais des contenus pertinents pour lui) à la fidélisation du client :

*« ...the key to productivity. Personalization technology cuts through the clutter and saves time. The firm that delivers personalized content sits at the top of the attention economy food chain [...] Personalization leads to relevancy in advertising; and loyal customers. »*<sup>200</sup>

Cette force en mouvement se confronte à celle, plus stable, de la *privacy* des informations personnelles concernant l'utilisateur. Une personnalisation efficace ne se fait pas sans connaître ces informations et sans savoir les utiliser (ce que quelques controverses autour du traitement des données personnelles, de la part d'un réseau social de plus en plus répandu, sont en train de montrer à ce moment-là<sup>201</sup>). Mais elle est destinée à devenir une tendance bien plus envahissante :

*« You cannot do personalization effectively without knowing an awful lot of information about an enormous number of people. The privacy backlash is building. Today it is only techies who are aware of the issue and where it is headed, but when mainstream users get spooked by a few more high profile cases, we will see consumer backlash and then, with politicians on the bandwagon, more regulation. »*<sup>202</sup>

Ces deux tendances sont depuis longtemps confrontées l'une à l'autre, mais ce n'est que récemment que les moyens techniques visant à les réconcilier sont devenus envisageables. Lunn conclut donc avec une suggestion et une question :

*« This will make a loud noise and will be messy; and in that mess will be big*

---

<sup>199</sup> Bernard Lunn (2007), « Attention to Intention to VRM – Opportunities for Entrepreneurs », [http://www.readwriteweb.com/archives/attention\\_to\\_intention\\_to\\_vrm.php](http://www.readwriteweb.com/archives/attention_to_intention_to_vrm.php)

<sup>200</sup> Id.

<sup>201</sup> Harrison Hoffman (2007). « Facebook's code source goes public », CNet News, [http://news.cnet.com/8301-10784\\_3-9758702-7.html](http://news.cnet.com/8301-10784_3-9758702-7.html)

<sup>202</sup> Id.

*opportunities for entrepreneurs. What are the alternative models that can combine the benefits of both personalization and privacy?*<sup>203</sup> »

Deux ans plus tard, en 2009, Georg se rémémore comment la lecture de cet article a marqué un moment important dans la « philosophie » de Faroo, car

*« there it was, in clear words : it was one of these opportunities we were trying to leverage : we were using P2P to try and solve both privacy and search scalability*<sup>204</sup>. »

Dans les premières phases de développement du dispositif, en effet, deux aspects se dessinent comme fondamentaux pour la valeur ajoutée de la « recette Faroo » : la personnalisation des recherches impulsée par la communauté, et une meilleure sauvegarde de la *privacy* par rapport aux services les plus répandus du marché. Le P2P, permettant d'éviter toute centralisation de données (au niveau du stockage ou du contrôle) et d'encrypter l'information quand elle « voyage » dans le système, peut se configurer comme une solution technique possible pour la tension entre propriété, appropriation et exposition des données personnelles (Katyal, 2005), dans une optique d'intégration du principe de *privacy* dans la technologie dès sa conception (Cavoukian, 2006 ; Schaar, 2010)<sup>205</sup>.

Cette section revisite certains moments du développement de Faroo qui ont déjà été introduits au cours des sections précédentes – et reprend en partie quelques-uns des arguments qui y ont été faits – à l'aune de la problématique du lien entre personnalisation et protection de la vie privée. En effet, la personnalisation – l'adaptation progressive des réponses du moteur suite à la détection des préférences de l'utilisateur – doit trouver ses racines, pour les développeurs de Faroo, dans la démarche de recentrage du processus de recherche sur l'utilisateur qu'on a déjà évoqué, et est mise en œuvre avec les outils du Web implicite et de l'économie de l'attention. Cependant, cette solution, qui suppose une observation et un enregistrement constants de la navigation de l'utilisateur, doit s'accompagner à des mesures techniques et juridiques susceptibles de « rassurer » les utilisateurs quant à la gestion de leurs données personnelles ; ce que les développeurs cherchent à obtenir en mettant à profit certaines caractéristiques de l'architecture décentralisée du service.

DU « DOCUMENT-CENTRE » À L' « USAGER-CENTRE »

---

<sup>203</sup> Id.

<sup>204</sup> Entretien, Georg Schneider.

<sup>205</sup> Id. (« *This is the big conflict: we need community powered personalization to filter information but we also need privacy. It is possible that P2P, where no data is stored/controlled centrally and is encrypted on the wire, is the only way to do this.* »)

Dès 2005, on peut noter dans les conversations entre Georg et ses collaborateurs la volonté d'un déplacement de paradigme qui aiderait ou précipiterait la fin de l'ère de la « *document-centered search* » en faveur d'une priorité attribuée aux usagers, aux contenus qu'ils créent ou suggèrent, à leurs conversations<sup>206</sup>. Au fil des mois, puis des années, se dessine au sein de l'équipe un changement de perspective par rapport à la centralité du document (Salaün, 2012)<sup>207</sup>, qui souhaite amener les conversations et les usagers à avoir la priorité – soit dans la phase de recherche elle-même, soit en tant qu'objets de cette recherche :

*« Users and conversation, both as the target of search as well as by contributing to discovery and ranking, must be core to the next generation of search. »*<sup>208</sup>

Sur cette base commune, la centralité de l'utilisateur est redéfinie par ailleurs de plusieurs façons. L'accent est mis selon les cas, au fil des discussions dans lesquelles l'équipe s'engage jusqu'à la fin de 2007, sur plusieurs « obstacles-problèmes » (Callon, 1986) que ce déplacement de perspective dans la conception d'un moteur de recherche peut contribuer à résoudre :

- la *rapidité* de l'obtention de résultats que la recherche d'information chez l'utilisateur rend possible – et comment cela satisfait, en même temps, la demande de cet usager pour les mises à jour les plus récentes (« *To satisfy consumer demand for real-time, documents need to be ranked as soon as they appear* ») ;
- la façon dont le *formatage* de l'information et des liens de confiance établis entre utilisateurs dans les médias sociaux remet en discussion, à un niveau technique, le paradigme dominant de la recherche en ligne consistant à recopier les documents qui forment le Web sur d'autres machines que celles des utilisateurs (« *we have had the explosion in content creation from social media. This new scale undermines the legacy search idea of copying the whole Web to giant server farms* ») ;
- l'excessive quantité de « bruit », d'informations non pertinentes ou de mauvaise qualité, de plus en plus présentes sur le Web, en parallèle à sa diffusion ; mettre l'utilisateur au centre du dispositif signifierait dans ce cas lui fournir de meilleurs outils de *filtrage*, un défi auquel d'autres dispositifs ne répondent que partiellement, en donnant la priorité aux documents qui

---

<sup>206</sup> Entretien, Hiba.

<sup>207</sup> Dans un ouvrage récent, Jean-Michel Salaün propose de considérer le Web comme une étape dans une « histoire longue et méconnue, celle du document ». Il suggère que les « nouveaux venus » dans le domaine des services Internet, comme Apple, Google ou Facebook, privilégient chacun une dimension différente du document pour prendre une position dominante dans la construction d'un « néodocument », et propose enfin aux acteurs indépendants du Web de développer leurs formes spécifiques d'« architecture de l'information » pour contrer l'hégémonie des géants de la Toile.

<sup>208</sup> Entretien avec Georg Schneider, 2010, reparcourant des notes de travail et discussions de l'équipe de la période 2005-2007. Les extraits cités dans la liste qui suit sont également tirés de cet entretien.

construisent leur popularité au fil du temps plutôt qu'à l'information potentiellement très pertinente par rapport à une requête mais pas encore suffisamment populaire (« *users need much better filters. Traditional ranking methods based on links fail to meet this new filtering challenge, as new documents naturally have no history and no record of incoming links. Ranking based on the absolute number of votes penalizes new documents* ») ;

- la possible valeur ajoutée, du point de vue de l'utilisateur, d'indexer la « longue traîne du Web », ces pages qui, bien qu'étant visitées (et étant, donc, pertinentes ou intéressantes pour un ou plusieurs usagers) n'arrivent pas à obtenir suffisamment de votes explicites pour être découvertes et classifiées par popularité (« *The long tail is the source of most real time search. Most pages never appear within the stream and cannot be discovered and ranked by popularity through explicit voting.* »)

Sur la base de ces arguments, l'équipe en vient à la conclusion que la réponse à l'ensemble de ces défis est de rendre opérationnel un « *crowd sourced approach to search* », une approche de la recherche déléguée aux « foules », et basée sur la détection et la traduction en algorithme de leurs intérêts agrégés. Opérer le déplacement d'un *document-centre* à un *usager-centre* passe donc par la construction d'un algorithme qui prenne en compte, en priorité, la découverte et la classification, de la part des usagers, de documents à la fois nouveaux et pertinents.

#### L'ÉCONOMIE DE L'ATTENTION ET LA RECHERCHE « SOCIALE »

De façon intéressante, l'équipe Faroo se positionne dès lors, explicitement<sup>209</sup>, parmi ces développeurs et entrepreneurs qui veulent traduire en pratique... et en *business plan*<sup>210</sup> – de l'architecture technique jusqu'aux motivations des utilisateurs et des usages eux-mêmes – l'approche socio-économique dite de l'« économie de l'attention ». Celle-ci cherche à rendre compte du fonctionnement de marchés où l'offre est abondante, et par conséquent économiquement dévalorisée, la ressource rare devenant le temps et l'attention des consommateurs. Cette théorie se traduit en deux points : dans le marché de l'information en ligne, où l'offre est certainement surabondante, la richesse de l'information crée un manque d'attention ; donc la séparation entre information pertinente et bruit non pertinent devient de plus en plus importante (Simon, 1971 ; Goldhaber, 1997 ; Kaplan, 2007 ; Kessous, Mellet et Zouinar, 2010).

Pour « construire » un moteur de recherche, remarque Georg, il faut donc considérer que les usagers

*« made search experiences [sic], sacrificed their attention before, and voted*

---

<sup>209</sup> White Paper, p. 19.

<sup>210</sup> Pour paraphraser Daniel Kaplan (2007).

*implicitly on the content they visited. This saves other users' time as they can focus on consuming already preselected, relevant information instead of searching for the needle in the haystack again. »<sup>211</sup>*

Le moment où plusieurs membres de l'équipe Faroo associent dans leur récit les concepts de Web implicite<sup>212</sup> et d'économie de l'attention, sera le seul moment pendant mon enquête où les développeurs seront extrêmement « quantitatifs » dans leur discours<sup>213</sup> ; très probablement, afin de « prouver » les différences d'efficacité entre le vote explicite, dont se servent beaucoup de moteurs de recherche en temps réel basés sur Twitter qui ont vu le jour en même temps que Faroo, et le vote implicite qu'ils implémentent dans leur système.

Georg précise par ailleurs que cela est dû à l'habitude que Szymon, Jiechi et lui-même ont, depuis deux ans, de répondre à des points de controverse soulevés par des collègues lors de plusieurs conférences entre 2008 et 2009 : ces critiques portent régulièrement sur la faisabilité de la mise en opération de l'approche « sagesse des foules ». Ces objections ont mis en avant soit le manque de représentativité d'une minorité d'utilisateurs actifs :

*« as with other social services, only a small fraction of the users may be actively contributing, and this lack of representativeness may put promising ideas like the wisdom of crowds into question »<sup>214</sup>,*

soit la disproportion entre l'information que ces utilisateurs actifs apportent et celle qu'ils connaissent effectivement :

*« also those people who are contributing, publish an even smaller fraction of the information they know »<sup>215</sup>.*

Le passage du vote explicite au vote implicite a été identifié comme le moyen de résoudre ces deux problèmes, pouvant proposer des résultats plus complets tout en maximisant l'économie de l'effort pour l'utilisateur, sans exiger une action de classification ou de vote explicite<sup>216</sup>.

---

<sup>211</sup> Entretien, Georg Schneider.

<sup>212</sup> Voir la section 2 de ce chapitre : il s'agit de l'analyse du comportement des utilisateurs pendant la navigation et la recherche, et de la récolte des « traces » que ceux-ci laissent au cours de la navigation. Ces traces sont comprises et traitées comme des « votes de préférence implicite » d'un contenu par les utilisateurs.

<sup>213</sup> Jiechi s'en excuse presque : « *It might be worth to have a closer look at the quantity structure of the underlying foundation, just to explore the feasibility and limits of this approach.* »

<sup>214</sup> Entretien, Szymon.

<sup>215</sup> Id.

<sup>216</sup> C'est là que Szymon estime opportun de quantifier ses propos, à ma quelque surprise : « *Twitter now has 44.5 million users and provides about 20,000 Tweets per minute. If every second tweet contains a URL this would be 10,000 URLs per minute. According to Nielsen the number of visited Web pages per*

Au fil du temps, les développeurs explorent d'autres dynamiques qui concernent l'attention des usagers à l'information, et la façon dont le système peut en « rendre compte » aux autres usagers. Le temps est identifié comme un facteur crucial, surtout pour la recherche en temps réel : l'utilisateur doit pouvoir découvrir une nouvelle page dès que possible, et comprendre dans des délais très courts la popularité de cette nouvelle page. Mais ces deux objectifs sont en conflit, dans le cas d'un mécanisme de vote explicite, avec le fait que le nombre de votes est proportionnel au temps d'observation de la page ; pour des pages introduites récemment, le nombre très limité de votes explicites n'est pas suffisamment représentatif pour donner un classement de pages fiable. Ce que l'équipe cherche à réduire, encore une fois, au moyen du mécanisme de vote implicite.

En même temps, Szymon (en particulier) travaille sur l'amélioration du mécanisme de vote explicite, qui sera complémentaire au vote implicite :

*« We just should not treat them as equal – because they are not. Some of them we [les usagers] trust more than others, and with some we share more common interest than with others, for the very same reason why we follow some people and some not. This helps us to get more value and meaning out of the very first vote. »<sup>217</sup>*

La valeur et la signification du vote explicite sont établies de façon relative, par rapport à ce que les autres usagers pensent de la page et de la confiance qu'ils lui attribuent. Le geste de l'interrogation au moteur de recherche n'est plus, dès lors, une activité solitaire, entre l'utilisateur et le dispositif : derrière les résultats obtenus se révèlent des préférences partagées, des affinités, des suggestions de la part d'autres usagers – en bref, des connexions :

*« In the real world you do not rely only on silently reading through heaps of documents; you also ask your colleagues and get hints from your friends. Yet, when it comes to online search, you are basically on your own again, in front of the search engine. But in fact, the engine is recommending and connecting you to people who are working on the same topic, or looking for information about it, at the same time. (...) Why is context sensitive search advertising so successful? Because you are presented the right ads at just the moment when you are interested in a specific topic. How would it be if you were presented*

---

*person per month is 1,591. The 44.5 million users visiting 1.6 million web pages per minute, while explicitly voting only for 10,000 per minute. Implicit voting and discovery provides 160 times more attention data than explicit. This means that 280,000 users with implicit voting could provide the same amount of information as 44.5 million users with explicit voting. Or that implicit discovery during one day finds as much web pages as explicit discovery during a half year. This shows drastically the limits of a web search which is based solely on explicit votes and mentions, and which potential can be leveraged by using the implicit web. »* Une discussion au sujet de la recherche "tweet-based" et de ses limites, où Georg reprend ces chiffres, diminuera quelque peu ma surprise ([http://www.readwriteweb.com/archives/the\\_limits\\_of\\_tweet-based\\_web\\_search.php](http://www.readwriteweb.com/archives/the_limits_of_tweet-based_web_search.php)).

<sup>217</sup> Entretien, Szymon.

*not with ads, but with like-minded people? »*<sup>218</sup>

C'est là que le choix de l'architecture technique du système, que nous avons momentanément abandonnée dans notre récit (et les acteurs dans le leur), refait surface :

*« As long as both source and recipients of information are distributed the natural design for search is distributed. [...] A peer-to-peer client allows the implicit discovery and attention ranking of every visited web page. In real time search the amount of index data is limited, because only recent documents, with high attention and reputation need to be indexed. This allows a centralized infrastructure at moderate cost. But as soon as search [...] aims to fully index the long tail of the whole web, then a distributed peer-to-peer architecture provides a huge cost advantage. »*<sup>219</sup>

Afin de mêler pertinence des requêtes et mise en relation des usagers qui en ont lancées de semblables, la recherche en temps réel se doit donc, selon Georg, d'être combinée à une architecture distribuée, en pair à pair. Le rôle de l'utilisateur, de ses interactions et conversations, doit « correspondre à une infrastructure adéquate ». D'un côté, pour des raisons d'efficacité où le technique se mêle au « social », dans le sens décrit ci-dessus ; de l'autre côté, parce qu'il s'agit de choisir un modèle qui répond mieux à la nécessité de récolte et distribution d'informations rapide, *pervasive*, directe et « aux marges »<sup>220</sup>. Le pair à pair peut fournir une telle structure.

Une opposition encore plus radicale entre les moteurs classiques « à la Google », basés sur un navigateur Web, et les moteurs P2P, se dessine également dans cette approche. Une des possibles valeurs ajoutées du modèle P2P/Web implicite appliqué à un moteur de recherche semble être le fait de capter une information sur les préférences et les pertinences d'un utilisateur « à la source » et indépendamment de la plateforme de recherche – cela étant rendu possible uniquement par un logiciel, incorporé au client P2P, qui observe l'ensemble de la navigation – tandis que les moteurs classiques ne captent les préférences et la pertinence des ajustements que dans l'usage de la plateforme moteur. Il s'agit d'une dimension ultérieure de l'approche centré sur l'utilisateur, qui rend par ailleurs la question de la *privacy* encore plus sensible, et nécessite d'une solution technique de protection de la vie privée

---

<sup>218</sup> Certaines caractéristiques de Faroo seront implémentées au cours de 2009 qui explorent cette vision, comme le *Social Tagging* (destiné à capitaliser sur la capacité des usagers à catégoriser l'information, en rassemblant les étiquettes – tags – appliquées par l'ensemble des usagers à un document spécifique, et proposant à l'utilisateur de chercher ces étiquettes et les utiliser comme filtres de recherche), la *Visual Preview* (la capacité du système à associer, où possible, une image dérivant de l'article en face du résumé de l'article, plutôt qu'un aperçu de la page web dans son entièreté), et la *Sentiment Detection* (principe basé sur l'idée que ce n'est pas seulement les « nouvelles » en soi, mais aussi les émotions et les sentiments qui concernent les usagers et qui attribuent une signification ultérieure à l'information, ce que l'algorithme de Faroo cherche à détecter et à visualiser dans les « conversations » entre usagers). (White Paper, p. 27).

<sup>219</sup> Entretien, Georg Schneider.

<sup>220</sup> Id.

encore plus « paradoxale », parce qu'elle doit s'accompagner à une observation constante de la navigation de l'utilisateur. Comme le montre la fin de cette section, les développeurs de Faroo se rendent bientôt compte de l'importance de résoudre cet enjeu de manière à ce que la personnalisation ne soit pas perçue et vécue comme « espionnage perpétuel », mais comme valeur ajoutée du moteur.

#### DE LA « PRIVACY BY POLICY » À LA « PRIVACY BY ARCHITECTURE »

Au fur et à mesure que les détails de la « personnalisation » des recherches dans Faroo sont rendus explicites, testés, et discutés par les usagers pionniers, certains d'entre eux commencent à exprimer une perplexité latente : certes, remarque le site de recension collaborative de start-up ReviewSaurus, l'historique des recherches (*search history*) de chaque usager contient des informations de valeur pour le classement et la personnalisation des recherches effectuées par les autres usagers. L'utilisation de cette information pour améliorer la recherche et donner aux résultats obtenus une meilleure pertinence est d'une grande importance pour gérer la surcharge informationnelle qui ne cesse d'augmenter. Faroo se démarque ainsi de façon intéressante de ses concurrents : les premiers usagers se trouvent généralement d'accord pour reconnaître dans ce mécanisme la principale valeur ajoutée du système<sup>221</sup>.

Mais qu'en est-il de ces informations ? Sous quelles conditions l'élaboration d'informations personnelles est-elle susceptible de diminuer ou reconfigurer la *privacy* des usagers ?

*« Personal search history is key to relevance, we got it. But we are not willing to trade our privacy to get this, »*

remarque Osias, tandis qu'Endre renchérit :

*« While this seems a useful thing, it raises privacy concerns. We feel and fear our privacy is once more fading away. »*

Un autre usager, surnommé *stumped*, est parmi les plus sceptiques : il compare Faroo à des logiciels novateurs mais controversés tels que *myware*<sup>222</sup>, se demandant – et demandant aux développeurs – comment les deux se distinguent :

---

<sup>221</sup> ReviewSaurus, « Faroo, P2P Web Search : Impressive Idea ! », <http://www.reviewsaurus.com/web-applications-reviews/faroo-p2p-based-search-engine-impressive-idea/>.

<sup>222</sup> Sorte de spyware installés volontairement par l'utilisateur, les logiciels *myware* gardent des traces des activités que celui-ci exerce sur et au moyen de son ordinateur, et en font bénéficier les autres utilisateurs, en s'appuyant sur des technologies d'agrégation de données et de métadonnées. Voir Erick Schonfeld, « With myware, soon you may be spying on yourself online », CNNMoney, janvier 2006, <http://money.cnn.com/2006/01/19/technology/futureboy/index.htm>, et en français, Laurent Neyssensas, « Social Software, partage et myware », septembre 2006, <http://turing.lecolededesign.com/neyssensas/index.php?2006/09/16/79-oh-no-not-me-i-did-it-my-war>

*« All in all, we install stuff that tracks our behavior. It can be for more or less noble reasons, but in the end that's the point: people are going to know stuff about me. I'd like to know something more about how that happens in the Faroo case. »*

Et Mark Johnson, un autre utilisateur, résume d'un ton neutre :

*« FAROO indexes the webpages I visit in my browser. Does this hurt my privacy?<sup>223</sup> »*

Bernard Lunn, de son blog sur ReadWriteWeb, souligne également l'importance de réconcilier la protection des données personnelles et la recherche basée sur la personnalisation, qu'il définit « révélation » des préférences et des intentions des usagers. Lunn signale qu'un certain degré de protection des données personnelles est selon lui une condition nécessaire pour que suffisamment d'utilisateurs pionniers aient confiance dans les mécanismes de personnalisation, et acceptent d'être part de la « masse critique » nécessaire au démarrage du dispositif. Cette confiance peut être renforcée à la fois par des considérations politiques et économiques, ayant trait à l'indépendance de Faroo vis-à-vis des acteurs dominants du secteur, ainsi qu'à la clarté, la fiabilité et le potentiel novateur du modèle technique :

*« [Faroo] needs to show the world that they are not just Microsoft acquisition-bait. That fear would alienate much of the early adopter crowd who they need to help them gain initial traction. [...] If users feel 100% confident that their privacy is being protected, they will reveal more and thus create a more accurate database of intentions. But that confidence is one part technical and one part institutional trust. So if they get bought by Microsoft, too many people will fear that their privacy will not be as safe as advertised. »<sup>224</sup>*

De leur côté, Georg et son équipe avaient longtemps travaillé sur les possibilités de réconcilier personnalisation et vie privée. Pendant un certain temps, il a semblé que la seule solution était de travailler à l'obtention d'un équilibre entre les deux, au moyen d'un compromis : le moteur de recherche aurait été utilisé par des usagers/volontaires,

*« primarily interested in personal and social search, and willing to experiment with it »,*

en laissant temporairement au second plan des questions plus épineuses concernant la protection des informations personnelles. Un autre aspect intéressant surgit ici, celui de la « délégation de la sérendipité<sup>225</sup> » : c'est à dire, ce qui est confié au moteur n'est

---

<sup>223</sup> ReviewSaurus, id. Cette formulation sera reprise par Faroo sur le site web de la firme, dans la partie des FAQs dédiée à la discussion de la privacy.

<sup>224</sup> Bernard Lunn, « Faroo: Could P2P Search Change the Game ? », cit.

<sup>225</sup> J'emprunte cette formule à Guillaume Sire (correspondance personnelle).

plus la hiérarchisation du web, mais l'interprétation des données relatives à chaque utilisateur ; c'est du croisement de ces interprétations que découle, en un deuxième temps, la hiérarchisation du Web.

L'équipe se rend progressivement compte que les utilisateurs potentiels de Faroo pendant la phase d'amorçage, intéressés et avertis d'un point de vue technique, ne s'arrêteront pas là : Georg se remémore que, quand, au début de 2008, des discussions comme celle du site ReviewSaurus ont commencé à émerger, l'équipe était « quelque peu préoccupée, mais pas vraiment surprise<sup>226</sup> ». Jiechi, en particulier, soutenait depuis déjà quelques temps que « cela [la réconciliation de la personnalisation et de la vie privée] ne doit pas forcément être un *trade-off*<sup>227</sup>. » Comment faire pour que Faroo puisse

*« make better use of personal information than current search engines can, and at the same time protect privacy better »<sup>228?</sup>*

En analysant la situation des produits sur le marché, l'équipe observe que certains moteurs de recherche ne sauvegardent pas les journaux (*logs*) des requêtes, tandis que d'autres systèmes les effacent ou les anonymisent après une durée de temps déterminée, ou à la demande de l'utilisateur. Mais ces mesures présentent plusieurs limites, car cela expose les *logs* à la traçabilité de la part de plusieurs acteurs : de la firme proposant le service elle-même, jusqu'au fournisseur d'accès Internet ou l'administrateur du système. C'est d'ailleurs pour cela que certains économistes plaident pour une visibilité publique des logs comme mesure en faveur de la concurrence (Argenton & Prüfer, 2011).

C'est sur ces limites de sécurité que l'on peut travailler pour apporter des changements à la conception et à la sauvegarde de la *privacy* dans un moteur de recherche. L'équipe se rend compte, en travaillant sur les couches P2P du système, que c'est encore une fois la décentralisation de l'architecture du système qui pose en même temps des opportunités et des défis de design :

*« if the balance of privacy and personalization is to be solved not as a compromise, but in a way that makes both aspects superior to current solutions, this is only possible through distributed search »,*

propose Georg à la communauté<sup>229</sup>. Avec la distribution, il devient possible d'éviter la traçabilité des *logs* de la part d'une entité centrale, et de faire en soi que les informations personnelles ne quittent pas le client P2P et le terminal de l'utilisateur, à moins qu'elles ne soient encryptées avant sur ce même terminal. Les risques de fuites

---

<sup>226</sup> Entretien, Georg Schneider.

<sup>227</sup> Entretien, Jiechi Chang.

<sup>228</sup> White Paper, p. 15.

<sup>229</sup> Au moyen d'une entrée de blog, <http://blog.faroo.com/2008/05/>

ou vols délibérés d'informations d'un « dépôt » central peuvent être minimisés. Le modèle distribué permet de reconfigurer le système Faroo comme « une sorte d'anonymiseur distribué » où les termes de recherche et d'interrogation ne peuvent pas être bloqués par des acteurs extérieurs... ou par l'entreprise<sup>230</sup>.

Georg et son équipe décident donc de poursuivre une stratégie de développement en trois volets : l'absence d'une entité coordonnatrice centrale (et sa distribution, ou répartition, sur le réseau<sup>231</sup>), l'absence de *search logs* (« carnets des interrogations ») et l'encryptage des requêtes faites au moteur de recherche.

En premier lieu, Jiechi travaille sur le client P2P de façon à ce que les données personnelles de l'utilisateur ne quittent pas son ordinateur. Georg, en reprenant tout particulièrement la question de Mark Johnson, explique que Faroo est programmé pour indexer seulement les pages publiques sur l'Internet, plutôt que des pages Intranet ou protégées par le protocole https. Tout en rappelant à ses usagers que l'action d'interroger un moteur de recherche, et d'en obtenir des résultats, est dans la plupart des cas une action « publique » vis-à-vis l'écosystème des fournisseurs de services Internet

*« it is important to be aware of the fact, that there is no privacy while visiting internet pages. The ISP, many intermediate stations in the Internet, and the visited site itself knows about your visit »,*

l'équipe Faroo met en avant dans son blog, en août 2009, les points principaux de sa « stratégie de *privacy* ». Entre coordination décentralisée, opérations en local et index distribué, les enjeux centraux du dispositif se dessinent : anonymisation et encryptage des interrogations.

*« In contrary, FAROO has no central institution, which would be able to collect the click streams. The (anyway public) web pages are yet hashed and encrypted at the computer of the user and then stored to the distributed index. The index contains only encrypted information. It does not contain any information about who stored the information into this index<sup>232</sup>. »*

---

<sup>230</sup> Entretien, Georg Schneider.

<sup>231</sup> La philosophie de la « distribution à plusieurs niveaux » revient dans le récit de Georg par rapport à la question de la *privacy* : « *In privacy, as in many other things, the whole system is as strong as its weakest element. Every centralized element is a point of failure, vulnerable to attacks, blocking and surveillance. That's why it's so important to apply the distributed principle to each vital element of the system. And to my knowledge Faroo is the only P2P system which has done this in practice.* » (Correspondance personnelle, mai 2011).

<sup>232</sup> Entrée de blog de Faroo, <http://blog.faroo.com/2009/08/>. Cette idée de « connecter » les usagers au moyen de leurs interrogations communes d'abord, et seulement ensuite, s'ils choisissent de le faire, de rendre l'interaction plus personnelle, se trouve aussi quand Georg aborde la réconciliation entre les caractéristiques de « social search » du système et sa conception de *privacy* : « *what we do is to unlock the collective intelligence of a search community of peers without sacrificing your privacy. To use the social features no registration is required. You are using an arbitrary alias or nickname. We just bring people with the same interest together. You communicate with like-minded people in the sense that you*

Le deuxième volet, l'encryptage des interrogations, doit permettre à l'utilisateur d'être protégé d'éventuels blocages ou filtrages de la part du fournisseur d'accès Internet ou de l'administrateur de système. Quant au troisième volet – l'anonymisation des interrogations – il s'agit là d'un aspect qui dérive de l'architecture distribuée et qui empêche le fournisseur de service de garder trace des *search logs*, même si c'était son intention :

*« no central instance for monitoring exists. Therefore not collecting search logs is not just a promise: it's technically infeasible<sup>233</sup>. »*

Bernard Lunn, commentateur de ReadWriteWeb qu'on a déjà rencontré dans notre récit – devenu en début 2010 consultant pour Faroo, sur la question, en particulier, de la tension personnalisation/privacy – reprend et détaille cet argument au cours d'une discussion avec Danah Boyd, par courriel sur la liste de diffusion de la P2P Foundation<sup>234</sup> :

*« FAROO only indexes information that is already publicly available. If something is bad, you (the user, NdR) can remove the original page, than it will disappear from the p2p system, automatically. Recently Google offered https encrypted search for more privacy (...) While Google's https search eliminates wiretapping by intermediate parties, they still collect log files. HTTPS does not protect against a leak of those log files by technical problems, human error, later changes of terms of services, criminal data theft, legal challenges, or silent cooperation with interested authorities. Google still collects personal data, but offers privacy by policy (which can easily be changed or challenged). The same is true of Facebook clearly! FAROO in turn does not collect personal data at all, and offers privacy by architectural design (attention data for ranking are anonymized /search queries stay encrypted all the time, because the index is encrypted itself / there is no central data repository which could be wiretapped). »<sup>235</sup>*

La stratégie opérée par Faroo peut donc se résumer en ce déplacement d'une « *privacy by policy* » à une « *privacy by architecture* ». Dans le récit de Georg et l'explication écrite de Bernard Lunn, se dessine donc, de façon intéressante, une idée de « *privacy par la technique* » qui met l'accent sur les contraintes posées par le modèle technique de choix des développeurs (Cavoukian, 2006 ; 2009 ; 2010 ;

---

*profit from their search experience, follow their discoveries and exchange ideas. But you may decide later if you want to become friends and when to reveal your identity. FAROO's Social Search is pure opt-in. It can bring you into conversation, but only if you opt to do that. »* Je souligne.

<sup>233</sup> Id. Je souligne.

<sup>234</sup> Groupe de recherche international qui étudie l'impact des technologies et de la pensée peer-to-peer sur la société : <http://p2pfoundation.net/>.

<sup>235</sup> Intervention de Bernard Lunn sur la liste de diffusion de la P2P Foundation, 27 mai 2010. Je souligne.

Schaar, 2010). Celle-ci souligne comment les termes d'utilisation et les politiques de confidentialité des concurrents de Faroo, promettant aux usagers que leurs données vont être traitées d'une certaine façon, peuvent être assez facilement changés, souvent à l'insu des usagers eux-mêmes<sup>236</sup>.

L'intervention de B. Lunn met par ailleurs l'accent sur un ensemble de contraintes qui, pour certaines opérations et traitements de données, ont obligatoirement ou nécessairement lieu au moment où se « fabrique » un réseau distribué et sans centre, avec ses limites, ses opportunités et ses conditions techniques. L'anonymisation des « données d'attention » issues des interrogations ; l'encryptage de ces interrogations et de l'index lui-même ; l'absence d'un point central qui pourrait être attaqué, constituent trois points qui contribuent à préciser une définition de protection de la vie privée de l'utilisateur/pair qui découle « *by architectural design* » du modèle distribué, et déplacent les termes de définition de la confidentialité des données de navigation comme ils avaient déplacé ceux de la recherche personnalisée.

---

<sup>236</sup> La discussion email dont est tiré l'extrait concerne l'opposition entre le réseau social distribué Diaspora\*, qui venait d'être publié en version *closed alpha* au moment de la discussion, et l'omniprésent Facebook, dont les politiques de confidentialité et les termes d'usage sont souvent critiqués pour leur fréquence de changement, et leur opacité pour l'utilisateur final (voir par exemple BBC News, « Facebook faces criticism on privacy change », <http://news.bbc.co.uk/2/hi/8405334.stm>).

## EN GUISE DE SYNTHÈSE

Ce chapitre s'est proposé d'explorer les contraintes, principalement techniques mais aussi de marché, qui s'exercent sur un dispositif dont les concepteurs veulent pousser la décentralisation jusqu'au bout, pour introduire une version de P2P « intégral » dans la recherche d'information en ligne. Pour ce faire, on a analysé plusieurs moments clés du développement de Faroo, start-up anglo-allemande née en 2005 et se proposant d'implémenter, sur la base d'une architecture technique en *peer-to-peer*, un moteur de recherche « réparti sur les utilisateurs » composé de nœuds-usagers qui, en lançant des interrogations de recherche, partagent en même temps des informations sur eux-mêmes, leurs préférences et leurs affinités avec les autres usagers. Le chapitre s'est plus particulièrement concentré sur trois aspects : la mise-en-opération des « six niveaux de distribution » qui constituent l'ossature du service ; l'implémentation socio-technique du compromis entre personnalisation et privacy ; le rôle de l'utilisateur et de sa machine dans le système, en termes de mise à disposition de ressources matérielles et de calcul, et de contribution à une « ressource informationnelle » distribuée.

Notre analyse de la recherche « six fois distribuée » – l'indexation, la découverte d'information, la recherche, la classification, l'amorçage, la mise à jour du logiciel – négociée et mise en œuvre dans Faroo, nous a permis d'observer comment l'équipe de développeurs de ce moteur de recherche s'est attachée à libérer le réseau sous-jacent à leur système, aux différents niveaux qui composent son architecture, de la dichotomie entre le client et le serveur. Les pairs y sont mobilisés en tant que composants d'un système sociotechnique reliant les usagers, les algorithmes et les ressources techniques mises à disposition. Mais on a vu comment le fait de ne pas considérer les pairs comme des avant-postes d'un système centralisé, mais de « prendre au sérieux » leur statut d'égaux aux marges du réseau et de travailler à partir de cette conception, se présente comme un enjeu extrêmement délicat pour les développeurs à ce stade de l'expansion du système. On a également vu comment la conception de la « centralité » de l'utilisateur (et des ressources que celui-ci apportera à l'ensemble du système) qui découle de cette démarche de distribution est apparue à la fois comme la plus grande opportunité et le risque majeur pour le dispositif. Il s'agit d'un système qui ne demande pas d'importants investissements de ressources de la part de l'entreprise, qui peut travailler dans des situations stables et instables. Il lui est également possible de se ressaisir de façon autonome après un événement ayant modifié sa structure, grâce aux pairs « survivants » qui garantissent la reconstitution d'une entité distribuée pouvant encore assurer ses fonctions.

En même temps, la centralité de l'utilisateur et des ressources que celui-ci apportera à l'ensemble du système peut aussi être problématique pour la durabilité du dispositif, particulièrement dans ses phases initiales. Ne compter sur d'autres ressources que celles fournies par les usagers, « compter sur le monde entier pour être son réseau », signifie aussi, pour les développeurs, prendre un gros risque, en pariant sur le fait qu'ils parviennent à intéresser un nombre important d'utilisateurs afin d'assurer le bon fonctionnement technique du dispositif lui-même, et de permettre sa croissance et stabilité au fil du temps. Plus d'utilisateurs signifient non seulement plus de ressources

matérielles (mémoire, CPU, bande passante) à disposition de l'ensemble, mais aussi, on l'a vu, un affinement progressif de la variété et de la fiabilité des résultats obtenus par chacun d'entre eux, et une meilleure expérience utilisateur pour l'ensemble du réseau. Dans ce chapitre, on a donc suivi comment cette tension s'est explicitée autour de trois points de discussion et de controverse au sein de l'équipe, et entre l'équipe et certains utilisateurs, au cours des premières années d'existence de Faroo.

Le premier de ces cas, concernant la qualité des résultats pendant la phase d'amorçage, a montré comment – la quantité et qualité de l'index de recherche en P2P dépendant de la quantité de personnes/pairs qui utilisent Faroo – la pénurie d'utilisateurs s'est révélée le problème principal pour les utilisateurs de la première heure, jusqu'au point de compromettre, dans certains cas, l'expérience utilisateur, ce qui a contraint les développeurs à des solutions techniques temporaires de repli. Le deuxième cas a suivi les discussions centrées sur le client P2P, que les utilisateurs de Faroo doivent installer comme précondition à l'entrée dans le système. Le symbole de l'articulation entre les dimensions locale et globale du système, « vendue » par l'équipe Faroo comme un des points novateurs et originaux de leur service, se transforme dans les réactions de certains usagers et commentateurs en un obstacle, une barrière à l'entrée, contraignant l'équipe Faroo à une renégociation du client P2P en tant qu'opportunité de marché, symbole d'une autre façon de concevoir la recherche d'information sur le Web et peut-être l'Internet dans son ensemble. Enfin, on a suivi la manière dont un troisième point de discussion parmi les usagers, concernant le statut juridique du dispositif et par conséquent de ses utilisateurs/pairs, a contribué à dessiner une conception de responsabilité socio-juridique partagée, basée sur la capacité de l'utilisateur de créer et exécuter des interrogations de recherche, qui, en s'intégrant à l'index distribué, auront une influence sur les résultats de recherche obtenus par les autres utilisateurs.

La dernière partie du chapitre s'est enfin attachée à décrire une des tensions centrales à laquelle l'équipe Faroo a fait face lors des premiers pas du dispositif : l'équilibre entre personnalisation des interrogations et protection de la vie privée de l'utilisateur. On a vu comment, dans sa démarche de « *crowd sourced approach to search* », approche de la recherche d'information déléguée aux « foules », l'équipe Faroo a traduit le déplacement, commun à plusieurs moteurs de recherche aujourd'hui, de la centralité du document à la centralité de l'utilisateur. Ce déplacement s'opère par un algorithme qui prend en compte, en priorité, la découverte et la classification, de la part des usagers, de documents à la fois nouveaux et pertinents. Pour ce faire, l'équipe a identifié dans le modèle P2P, pour des raisons d'efficacité où le technique se mêle au « social », l'architecture plus propre à refléter et prendre en compte le rôle de l'utilisateur, de ses interactions et conversations. Les développeurs ont également pris conscience de l'importance de réconcilier la protection des données personnelles et la recherche basée sur la personnalisation : la première est une condition nécessaire pour que suffisamment d'utilisateurs pionniers aient confiance dans la deuxième, et pour qu'ils acceptent d'être part de la « masse critique » nécessaire au démarrage du dispositif. En a résulté une stratégie de développement à trois volets : rechercher l'absence d'une entité coordonnatrice centrale (et sa distribution, ou répartition, sur le réseau) ; l'absence de *search logs* ; et l'encryptage des interrogations faites au moteur de recherche. Cette stratégie a contribué à préciser une définition de protection de la

vie privée de l'utilisateur/pair découlant « *by architectural design* » du modèle distribué.

En conclusion, l'analyse de Faroo a suivi une démarche d'innovation qui – en attribuant une importance centrale à l'utilisateur, tout en craignant sa défection – souhaite proposer un modèle de décentralisation « intégrale », à plusieurs niveaux, comme alternative économique – vision qui, bien que par moments liée à questions plus classiquement « politiques » de liberté et contrôle, ne peut se résumer à celles-ci. Comme le souligne Georg,

*« In the last ten years, people have often come to me and told me: 'for me as a user, a client, a customer, there have been several startups I was interested in until they made a point of their architecture being based on P2P. Customers really just want a service. Does basing a service on P2P matter?' I try to ask myself this question as often as I can. What it takes not to get blown away by the most dreamy and utopian aspects of a project like this, is to always keep in mind this question first: cui prodest? If we are thinking about a decentralized Internet, we should analyze why the Internet moved away from that idea in the first place.*

*Only from there can we move to consider how it would look like, who would use it, who will build it, who will support it, who will prevent it<sup>237</sup>. »*

La libération du serveur, par un P2P qui s'explique dans plusieurs « couches » du dispositif, n'est donc pas, dans la démarche de Faroo, une fin en soi, mais un moyen pour l'entreprise, les usagers, ou les deux à la fois, d'obtenir un avantage économique – un changement dans la protection de ses données personnelles, dans l'utilisation de ses ressources techniques et de celles des autres usagers, dans la pertinence des résultats de ses interrogations. Pour l'équipe Faroo, il faut donc que le modèle P2P soit poussé jusqu'au bout, non pas comme étiquette médiatique mais comme avantage comparatif et comme produit de qualité, comparable à ceux des grands acteurs<sup>238</sup>.

Cette stratégie, au delà des aspects plus « rêveurs » et utopistes liés à la complète distribution de l'information que le modèle P2P comporte, est susceptible de proposer de manière claire un modèle de contribution de ressources certes risqué – puisqu'il s'appuie entièrement sur le choix de « migration » délibéré de l'utilisateur – mais capable de mériter la définition de « technologie de rupture »<sup>239</sup>. Elle suggère aussi la

---

<sup>237</sup> Entretien avec Georg Schneider, 20 juillet 2010.

<sup>238</sup> Comme explique Georg dans un très récent post (15 juin 2011) sur le forum de Faroo, une étiquette P2P sur-exploitée dans une campagne de communication pourrait même être contre-productive : « *The idea of P2P search, we have announced some time ago, and we also had a good response in the press. The next news (...) is a P2P search, which can compete in quality with the result of traditional search engines. Of this, people must be able to be convinced immediately. Since we already are working on it with high pressure, to start a major public relations campaign prematurely is counterproductive, and burns the P2P idea.* » (<http://forum.faroo.com/viewtopic.php?f=3&t=67>)

<sup>239</sup> Ou « perturbatrice », *disruptive* (Christensen & Overdorf, 2000 ; Christensen et al., 2006); sur le P2P comme *disruptive technology*, voir [Oram, 2001] et [Subramanian & Goodman, 2005].

possibilité de co-existence de deux versions du *social turn* dans la recherche d'information en ligne (Noll & Meinel, 2007 ; Smyth et al., 2009) : une qui demeure algorithmique et centralisée avec Google, et une qui s'appuie sur la distribution, la décentralisation, le P2P.

S'il est l'une des expérimentations avec la recherche décentralisée les plus réussies jusqu'ici, Faroo n'est pas un cas isolé. D'autres projets et entreprises, à des stades de développement très différents – des projets universitaires<sup>240</sup>, des projets qui sont en train d'implémenter la couche P2P de leur service au même moment où ils deviennent une start-up<sup>241</sup>, ou encore des dispositifs en fonction, parfois dotés d'un nombre important d'utilisateurs<sup>242</sup> – contribuent à façonner la recherche d'information en P2P, et sa place présente et future dans le marché des moteurs de recherche.

D'une part, l'introduction de dynamiques décentralisées dans les moteurs de recherche peut sembler un phénomène de niche – ne pouvant intéresser que des utilisateurs pionniers, souvent développeurs eux-mêmes, pour qui l'exploration du substrat technique complexe du système représente un défi, et son utilisation une valeur ajoutée. Les moteurs de recherche en P2P resteraient alors apanage de cette minorité de *power users* pouvant démêler, à l'aide de leurs compétences techniques, des équations à l'allure parfois paradoxale : comme celle qu'on a analysé dans la dernière section de ce chapitre, qui relie une *privacy* augmentée à un suivi perpétuel des actions des utilisateurs – actions dont les traces resteraient locales et encryptées. D'autre part, le nombre important de clients/usagers de Faroo – deux millions et demi en 2010 – semble montrer que l'intérêt pour ce type de dispositif se développe sur des bases plus larges.

Le succès ou l'échec des moteurs en P2P pourrait enfin être liée à la capacité de ces systèmes à « monter sur les épaules des géants », au même moment où ils cherchent à s'en différencier : pour plusieurs de ces projets, Gossple et Seeks par exemple, la valeur ajoutée et la spécificité d'un moteur P2P ne consiste pas dans le fait de se présenter comme une « alternative » radicale à Google, mais en complémentarité avec celui-ci. Cela présente l'avantage, pour le moteur P2P, de « construire sur » la force d'un acteur comme Google au lieu d'essayer de la contrer, ce qui, surtout dans la phase d'amorçage, signifie sauf exceptions la défaite. Cette complémentarité pourrait être la clé permettant au grand public d'envisager la recherche en P2P non pas comme « meilleure » ou « pire » que la recherche centralisée en termes absolus, mais comme une technologie permettant d'explorer différents types de résultats, ou de donner la priorité à certains aspects de la recherche et de la découverte d'information : par exemple, à la recommandation, à l'identification des préférences, et au rapprochement des usagers partageant des intérêts communs.

---

<sup>240</sup> Gossple, <http://www.gossple.fr>

<sup>241</sup> Seeks, <http://www.seeks-project.info/>

<sup>242</sup> Faroo, mais voir aussi YaCy, <http://yacy.net/>

## CHAPITRE 5

### WUALA, OU L'« ART DU COMPROMIS » ENTRE DISTRIBUÉ ET CENTRALISÉ

#### INTRODUCTION

Troquer de l'espace disque en local contre de l'espace de stockage en ligne : en bref, telle est la proposition de Wuala, une start-up suisse proposant un service de stockage en ligne<sup>243</sup> communautaire. Le dispositif est basé sur une technologie mise au point par les co-fondateurs de l'entreprise, Dietrich Poell et Kurt Gueden, lors de leur doctorat en informatique à l'Institut fédéral de technologie (ETH) de Zurich. La version *closed alpha* du dispositif est sortie à la fin de 2007, et la version *beta* publique a été proposée en 2009, année où Wuala s'est associé à StockFast<sup>244</sup>, fabricant français de périphériques de stockage. En septembre 2011, l'équipe compte 17 personnes<sup>245</sup>, et continue de s'accroître ; dix des membres de l'équipe sont diplômés d'un master ou d'un doctorat en science informatique de l'ETH Zurich ou de l'Ecole Polytechnique Fédérale de Lausanne.

Aux débuts de 2007, la création de la petite start-up survient dans un secteur où « l'industrie du stockage en ligne des données ne s'est jamais sentie aussi bien », notera par la suite Yannick Guerrini du populaire site de recension de matériel informatique, Tom's Hardware<sup>246</sup>. Les disques durs montés sur la plupart des ordinateurs d'aujourd'hui dépassent souvent les cent gigaoctets de mémoire, tandis que ceux des serveurs destinés au stockage professionnel de données en comptent désormais plusieurs téraoctets. Amazon S3, Azure Storage, Sun Cloud Storage sont quelques exemples de ce que les grands noms de l'industrie informatique proposent aux entreprises, en mettant en avant un modèle de service où le fournisseur est en

---

<sup>243</sup> Service permettant à ses utilisateurs de stocker, sauvegarder, consulter et partager des fichiers sur une ou plusieurs machines qu'ils ne contrôlent pas directement, mais connectées à l'Internet.

<sup>244</sup> Le nom est fictif.

<sup>245</sup> En 2009, quand j'ai commencé à m'intéresser au dispositif, ils étaient une dizaine.

<sup>246</sup> Yannick Guerrini (2010). « Wuala : le P2P comme solution de stockage », <http://www.presence-pc.com/actualite/Wuala-stockage-cloud-P2P-39035/#xtor=RSS-11>

charge de l'infrastructure physique et du produit logiciel, abritant ainsi à la fois les applications et les données dans un lieu inconnu de l'utilisateur (le fameux « nuage », *cloud* en anglais) et interagit avec ce dernier grâce à une interface client. En parallèle, se développent de nombreuses offres de stockage en ligne pour le grand public, d'un volume généralement limité à quelques gigaoctets – c'est le cas de SkyDrive de Microsoft, une offre de stockage en ligne de 25 gigaoctets intégrée à l'espace Live<sup>247</sup>, lancé en août 2007.

C'est au sein de ce scénario foisonnant d'applications de stockage, qui reposent généralement sur des serveurs « classiques » ou des plates-formes en nuage, que Wuala doit se positionner et chercher une vocation lors de son démarrage en 2007. L'équipe de développeurs choisit de travailler sur l'architecture du service en tant qu'aspect différenciant et original par rapport à ses compétiteurs : l'idée centrale sous-tendant le dispositif est que les fichiers téléchargés par les utilisateurs dans le système seront stockés à la fois dans des centres de données classiques et sur un nuage de stockage composé d'une partie des disques durs de chaque utilisateur, reliés entre eux en architecture P2P.

Si dans le chapitre précédent il était question d'explorer les contraintes techniques et économiques qui s'inscrivent dans un dispositif complètement décentralisé, ce chapitre souhaite explorer une approche hybride au P2P, en prenant la construction du modèle de stockage distribué de Wuala comme cas d'étude. Comment et pourquoi se construit le compromis entre une approche largement dépendante d'une plateforme de serveurs, et une approche d'architecture distribuée et décentralisée, où le fonctionnement du dispositif repose sur la mise à disposition de ressources *hardware* de la part des usagers ? La courte mais dense histoire du dispositif qu'on étudie dans ce chapitre (Tableau 5) est celle d'une application dont les créateurs ont travaillé, selon les moments, avec différents modèles d'architecture technique – mais aussi avec la représentation de ces architectures auprès de leur public de clients potentiels, et avec la volonté de ces clients/usagers de s'engager dans un « jeu » de partage de ressources aux implications matérielles, juridiques et sociales multiples.

Dans le façonnage de leur service de stockage distribué, les développeurs de Wuala mènent une double expérimentation : avec la frontière entre centralisation et décentralisation, ainsi qu'avec des dynamiques de partage et d'utilisation mêlant P2P et réseautage social. On explorera cette double expérimentation pour analyser les manières dont est gérée une situation qui est à la fois incertitude et opportunité technique : la mobilisation des marges du réseau, où se déploient des ressources matérielles et « humaines » très riches, mais disponibles de façon temporaire et inconstante. Ressources qui peuvent aider l'entreprise à trouver sa niche au sein d'un secteur foisonnant de nouvelles solutions, et où la concurrence est aigue.

---

247

<http://www.microsoft.com/belux/fr/windows/windowslive/Views/productDetail.aspx?product=Skydrive>

<p>2004: Début de doctorat de Dietrich Poell (DP) et Kurt Gueden (KG) à l'ETH Zurich.</p> <p>2006: Participation de DP et KG à Venturelab avec le projet Kangoo (prédécesseur de Wuala).</p> <p>2007 (janvier-juin): DP et KG obtiennent leur doctorat et comptent parmi les « Innovateurs Suisses » (un prix établi par le gouvernement suisse) de 2007 ; création de Caleido, firme de développement de logiciels ; création d'Address-Book, le premier logiciel de Caleido</p> <p>2007 (juillet): Publication de la version <i>closed alpha</i> de Wuala</p> <p>2007 (octobre): Controverse sur la sécurité de Wuala (Colin Percival)</p> <p>2007 (octobre): Google Tech Talk de démonstration de la version <i>closed alpha</i> du dispositif</p> <p>2008 (mars): Distribution d'invitations à tester le <i>closed alpha</i></p> <p>2008 (octobre): Publication de la version <i>private beta</i></p> <p>2009 (mai): Révision des conditions d'utilisation/clarification sur l'utilisation des ressources partagées</p> <p>2009 (juillet): Publication de la version <i>public beta</i>.</p> <p>2010 (juin): Fusion avec StockFast « pour travailler sur l'avenir du cloud storage ».</p> <p>2011: l'équipe compte 17 personnes.</p> <p>2011 (28 septembre): la fonctionnalité de troc est arrêtée.</p> <p>[2011, septembre: mon enquête de terrain prend fin]</p> <p>2012 (juillet) : DP a quitté Caleido. « Wuala by StockFast » vient de publier Wuala Business, une solution <i>cloud</i> destinée aux entreprises.</p>
---

Tableau 5. Chronologie du développement de Wuala, 2004-2011.

## SECTION 1. UN STOCKAGE EN RÉSEAU P2P

En 2009, quand j'entre en contact avec l'équipe Wuala, Dietrich Poell a vingt-neuf ans et est titulaire, depuis deux ans, d'un doctorat en sciences informatiques de l'université de technologie de Zurich, au cours duquel il a fait des stages à l'université technique de Delft, aux Pays-Bas (2003) et à Siemens Corporate Research dans le New Jersey (2004). Une fois obtenu son doctorat, en 2007, il fonde la firme de développement de logiciels Caleido<sup>248</sup>, dont le premier produit est le Caleido Address-Book, un logiciel pour la gestion des contacts professionnels des petites et moyennes entreprises. Au cours de l'année qui suit, il vend près de 35000 licences pour Address-Book en Suisse, Allemagne et Autriche, ce qui vaut au logiciel une certaine notoriété dans la partie germanophone de l'Europe<sup>249</sup>. Avec les revenus de ce premier-né « d'un impact révolutionnaire douteux, mais doté d'une valeur stable et claire<sup>250</sup> », Caleido compte financer les premiers pas d'un projet plus ambitieux, se basant sur un « vieux rêve » de l'informatique pour lequel les pères fondateurs, comme le professeur américain Andrew Tanenbaum, envisageaient pourtant de nombreux défis :

*« It has been an old dream of computer science, actually Tanenbaum wrote in his book on distributed operating systems stats, in the mid-nineties, that the design of a world-wide fully transparent distributed file-system for simultaneous use by millions of mobile and frequently disconnected users was left as an exercise for the reader. Now, I really hope that he meant this ironically, because we have been working on this exercise for the last five years<sup>251</sup> ».*

L'idée du projet Wuala, qui à ce moment n'a pas encore de nom, prend forme dès le retour de Dietrich des Etats-Unis en 2004, son inscription en doctorat de sciences informatiques à Zurich, et sa rencontre avec Kurt Gueden. Les deux étudiants se découvrent un intérêt commun pour les projets de recherche à large échelle sur les systèmes de stockage distribué, conduits à ce moment-là dans plusieurs universités techniques et départements de recherche d'entreprise aux Etats-Unis : OceanStore à l'université de Californie à Berkeley<sup>252</sup>, Cooperative File System au Massachusetts Institute of Technology (MIT)<sup>253</sup>, Past à Microsoft Research<sup>254</sup>. Ils trouvent cependant qu'il y a une limite commune à ces projets – aucun d'entre eux n'est

---

<sup>248</sup> <http://www.caleido.com/>

<sup>249</sup> E.g. CHIP Online, [http://www.chip.de/downloads/Caleido-Address-Book-2008\\_12995385.html](http://www.chip.de/downloads/Caleido-Address-Book-2008_12995385.html)

<sup>250</sup> Entretien avec Dietrich Poell, co-fondateur et CEO de Wuala, le 23 septembre 2009.

<sup>251</sup> Entretien avec Kurt Gueden, co-fondateur et CTO de Wuala, le 24 novembre 2009.

<sup>252</sup> <http://oceanstore.cs.berkeley.edu/>

<sup>253</sup> <http://pdos.csail.mit.edu/papers/cfs:sosp01/>

<sup>254</sup> <http://research.microsoft.com/en-us/um/people/antr/past/>

orienté vers la fabrication d'un produit destiné à un marché de masse :

*Our goal was to build something that was both challenging from a research point of view but which also had a great potential. Distributed storage systems were a hot research topic at that time and building a system that can harness idle resources seemed very promising, as the need for online storage was rising and file sizes were increasing. But the research projects had an academic focus only – we wanted to realize it and make it work in practice for millions of users<sup>255</sup>.*

*There have been a lot of research projects around this topic (...) but of course they had a different focus, they had an academic focus, and they didn't want to build a product. (...) We were inspired by them, but we really wanted to make it work, we wanted to bridge the gap between fundamental research and what it takes to build a product<sup>256</sup>.*

Pendant qu'ils achèvent leur doctorat, Dietrich et Kurt cherchent à préciser leur idée « orientée produit » d'un système de stockage distribué. En mai 2006, ils participent aux formations de Venturelab<sup>257</sup>, un programme national suisse financé par la Commission pour la technologie et l'innovation. « Kangoo », leur proposition de projet, comporte déjà certaines caractéristiques du futur Wuala : un système de stockage global et distribué en ligne, qui met à profit les ressources des ordinateurs participants, et qui répond au besoin croissant d'espaces de stockage de l'utilisateur *lambda*. Le mémo présenté par Dietrich à Venturelab mentionne aussi l'accès en mobilité et l'« ouverture du système, donnant lieu à une multiplicité de possibilités<sup>258</sup> ». L'idée est sélectionnée et, quelques mois après, les créateurs de Kangoo comptent parmi les « Innovateurs Suisses » de Venturelab pour 2007<sup>259</sup>.

Quand Dietrich Poell et Kurt Gueden terminent leur doctorat à l'Institut de technologie de Zurich, en 2007, ils ont donc déjà travaillé depuis quelques années au stockage en ligne. Il leur semble clair à ce moment que l'originalité de « *encore un autre* » dispositif de stockage ne pourra pas être liée à la nature du service fourni. Celui-ci proposera *grosso modo* les mêmes fonctionnalités de plusieurs autres services de stockage : la possibilité de sauvegarder, stocker et accéder à des contenus en ligne, en attribuant différents degrés de confidentialité aux différents contenus<sup>260</sup>. Les développeurs expliquent ces fonctionnalités dès novembre 2007, dans une des premières entrées de blog sur le site du service :

---

<sup>255</sup> Entretien avec D. Poell.

<sup>256</sup> Entretien avec K. Gueden. Je souligne.

<sup>257</sup> <http://venturelab.ch/>

<sup>258</sup> Memo « Kangoo » présenté par Dietrich Poell à venturelab.ch en mai 2006, disponible sur le site de Venturelab.

<sup>259</sup> [www.rezonnance.ch/images/doc/pre\\_orderformSwissInnovatorsYB2007.pdf](http://www.rezonnance.ch/images/doc/pre_orderformSwissInnovatorsYB2007.pdf)

<sup>260</sup> Ces fonctionnalités s'enrichiront par la suite d'autres plus « sociales » de collaboration et partage ; voir par la suite et à <http://www.wuala.com/en/learn/features>.

*The main window looks very similar to any file system, you have different folders, and if I go into a folder you can see other subfolders that have different colours. The yellow folders are private, so only you have access to them, the red folders are shared with a definite number of friends, and the blue ones are published, these are the ones everybody can access. It's very easy to share, for example if I click on the "share" button at the top, I could just select another friend with whom I want to share it, or I can click on public to publish this folder, or I can set it private again. If you go to "My friends" then you can see all the files that your friends share with you. Then there are also groups, which are very similar to a Flickr pool and can be private or public. Then there is the World area, where you can search and browse for all the files that have been published. All the files in the blue folders are indexed here<sup>261</sup>.*

« *There is only so much a customer can actually do with a storage service<sup>262</sup>, » résume de façon lapidaire Dietrich. L'originalité du service ne sera donc pas celle du service en lui-même ; les deux développeurs considèrent que la manière dont ils pourront émerger et se différencier de leurs concurrents sera le travail sur les « tuyaux », la plateforme sur laquelle celui-ci est fourni – avec des implications pour la sécurité du système, la privacy des données stockées, la vitesse des téléchargements en amont et en aval, la rapidité des échanges effectués.*

Et si les ressources de cette plateforme étaient les terminaux, unités de mémoire, bande passante mises à disposition par les usagers ? Cette décentralisation permettrait d'un côté, du point de vue de l'entreprise, de diminuer les coûts d'entrée sur le marché en rendant cette entrée plus facile ; et, par ailleurs, cela permettrait de proposer aux clients un produit « alternatif » à plus d'un titre :

*« an online storage with the power of P2P, if well done, means fast downloads, no file size limit, no traffic limit (...) a decentralized system could harness idle resources of participating computers to build a large, secure, and reliable online storage. This will enable its users to trade parts of their local storage for online storage and allow us to provide a better service for free<sup>263</sup>. »*

C'est cependant l'« effet vitrine » donné par l'interface du service (Image 4) qui lui vaudra le nom qu'il porte à ce jour, Wuala :

*« The new name came right before Dietrich left for his Silicon Valley tour<sup>264</sup> ... it came out of the French word 'voilà', 'look there'. For us, it*

---

<sup>261</sup> Blog Wuala, <http://www.wuala.com/blog/2007/11>

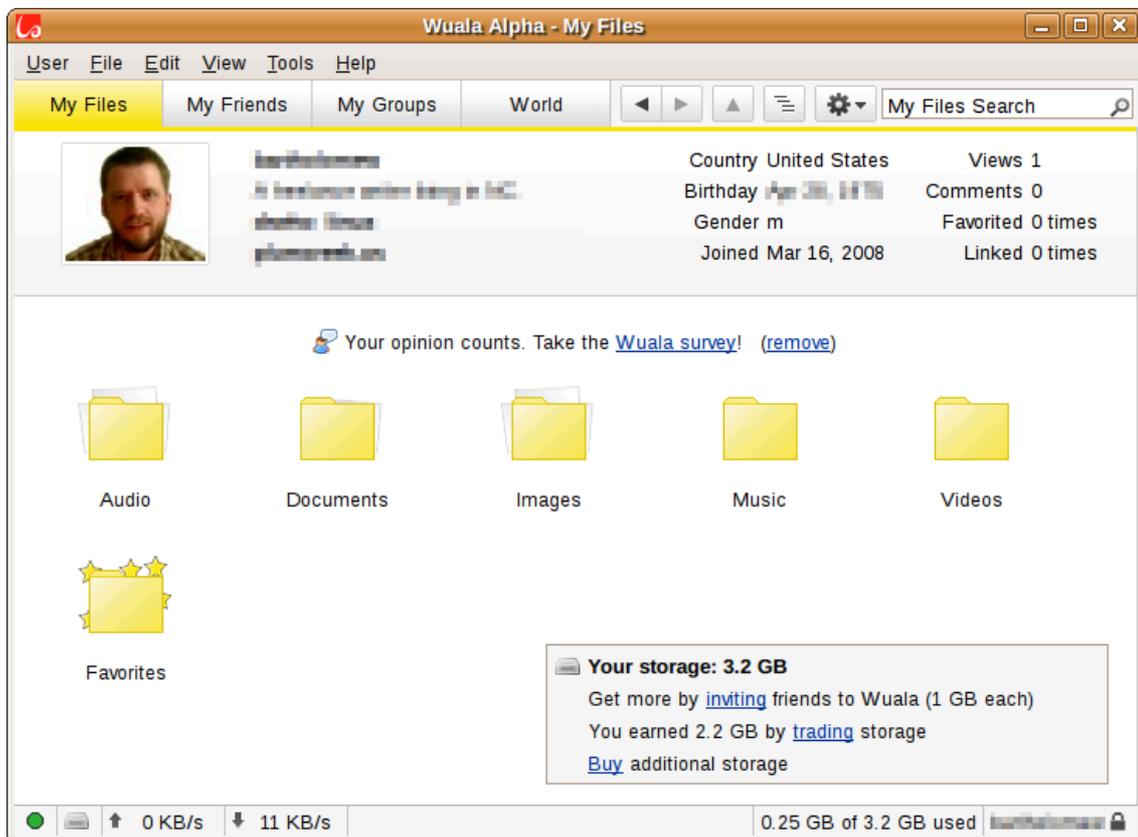
<sup>262</sup> Entretien avec D. Poell.

<sup>263</sup> Entretien avec K. Gueden.

<sup>264</sup> D'octobre à novembre 2007, Dietrich est aux Etats-Unis pour présenter Wuala et chercher des potentiels financeurs, programmeurs et contributeurs, notamment au Web 2.0 Summit de San

*means 'there, here's your file', simple and fast<sup>265</sup>. »*

L'idée d'un système où les fichiers sont stockés dans un réseau composé d'une partie des disques durs des utilisateurs, reliés entre eux directement en réseau P2P, commence alors à prendre forme et à faire le pont entre la recherche informatique et mathématique pure et un projet de service utilisable. Avec elle, les « problèmes » typiques du développeur de systèmes P2P commencent à faire surface et à nécessiter de solutions *ad hoc*. Dans la première version du système envisagée par Dietrich, lorsqu'un utilisateur télécharge un fichier, celui-ci doit d'abord être chiffré et mémorisé en bloc sur l'ordinateur de son utilisateur, puis divisé en fragments encryptés, qui quittent alors l'ordinateur de l'utilisateur pour être stockés de manière redondante dans l'ensemble des ordinateurs des utilisateurs, reliés les uns aux autres directement, de manière à former une plateforme.



*Image 4. Capture d'écran de l'interface Wuala vue par l'utilisateur, stade de closed alpha (début 2008) sur système d'exploitation Linux. Image par Daniel Bartholomew sur LinuxJournal.com.*

---

Francisco. (entretien du 23 septembre 2009 et <http://www.wuala.com/blog/2007/10>).

<sup>265</sup> Id.

Comme ce sera le cas plusieurs fois par la suite, remarque ironiquement Kurt Gueden, c'est lorsque Dietrich lui fait part de cette idée que son rôle dans la compagnie devient clair :

*« I started to ask difficult and less-than-dreamy questions, and sometimes he didn't exactly love me for it<sup>266</sup>. »*

Tous deux développeurs talentueux, l'un plus visionnaire et l'autre plus circonspect, Dietrich et Kurt forment un embryon d'équipe efficace... et ne tardent pas à se rendre compte de la difficulté de leur tâche. Comme souligne Dietrich :

*« Many times, we found that we needed to take a step back and wonder, is decentralizing it all really worth it? In most instances, it was about debating the pros and cons of decentralizing a specific aspect of the service's structure, and the answer we ended up giving ourselves was never twice the same<sup>267</sup>. »*

#### REDONDANCE : FAÇONNER UN STOCKAGE P2P PERMANENT

Les deux développeurs – qui ont entre temps été rejoints dans leur démarche par trois autres doctorants ou ex-étudiants de l'ETH Zurich – doivent répondre à une première question « difficile et peu séduisante ». Si les données sont stockées dans les ordinateurs formant le réseau P2P, les terminaux individuels de chaque pair peuvent donc passer hors ligne et réapparaître en ligne à tout moment. Comment peut-on assurer, dans cette situation de déconnexions et reconnexions fréquentes, que l'ensemble des données reste toujours disponible dans le réseau pour un utilisateur qui souhaiterait récupérer ses données ? En d'autres termes, comment construire un stockage P2P stable et permanent ?

Une première approche est tout d'abord discutée au sein du groupe de développeurs : le réseau peut être développé de façon à ce que les données y soient constamment présentes, ce qui implique de déplacer les données stockées d'un ordinateur, quand celui-ci passe hors ligne, vers un autre ordinateur. Cette solution ne paraît pas, après discussion, la plus appropriée dans un contexte distribué, pour deux raisons. Tout d'abord, le déplacement de fichiers volumineux ou d'un nombre important de fichiers prendrait beaucoup de temps ; cela poserait un problème pour l'utilisateur sur le point de se déconnecter, qui devrait attendre que cette opération soit effectuée. Ensuite, et surtout, les déconnexions et reconnexions dans un environnement P2P sont généralement fréquentes ; *de facto*, les ressources en bande passante à la disposition du système, aussi importantes soient-elles, seraient utilisées pour des opérations sans fin de copie de données, au lieu d'être dédiées à la rapidité

---

<sup>266</sup> Entretien avec Kurt Gueden, 24 septembre 2009.

<sup>267</sup> Id.

des téléchargements<sup>268</sup> :

*« We do want to have lots of data. So, suppose that you have 6 gigabytes stored on your computer, you don't want to wait until these 6 gigas are uploaded on another computer. It is also not a good idea if you have a high churn rate, which means that people frequently disconnect, reconnect, re-disconnect to the system, and this is what you typically have in such an environment<sup>269</sup>. »*

La deuxième approche consiste à laisser les données (ou, comme on verra d'ici peu, les fragments de données) là où ils sont, et à en assurer la redondance au sein du réseau. Celle-ci ne doit pas pour autant impliquer une réplication totale des fichiers, ni des fragments de fichiers, sur chaque ordinateur, ce qui nécessiterait un trop grand déploiement de ressources. Pour obtenir une redondance qui ne soit pas une réplication, l'équipe Wuala commence donc à travailler avec les codes d'effacement,

*« codes that transform a message of  $k$  symbols into a longer message with  $n$  symbols, such that the original message can be recovered from a subset of the  $n$  symbols<sup>270</sup>. »*

Ce qui, traduit dans la gestion des ressources à disposition dans un système – m'explique Dietrich – signifie que ce système de codage permet de télécharger des duplicata de morceaux de fichiers beaucoup moins de fois qu'il ne serait nécessaire s'il fallait télécharger le fichier dans sa totalité<sup>271</sup>. Dans cette approche, Dietrich et Kurt voient aussi une façon de contrer un des problèmes des systèmes P2P « traditionnels » de partage de fichiers, qui marchent d'autant mieux que les fichiers sont plus populaires (et donc plus répandus dans le réseau)<sup>272</sup>.

Au vu du grand nombre d'utilisateurs envisagés, et de données qui circuleront dans le dispositif selon les prévisions de l'équipe, c'est donc cette deuxième approche qui est finalement adoptée, comme explique Dietrich lors d'une première démonstration de la version *closed alpha* du dispositif, en octobre 2007, à l'occasion d'une Google Tech Talk<sup>273</sup> :

---

<sup>268</sup> Ce que l'équipe Wuala souhaite, au contraire, être l'un des points de force du service rendu possible par le réseau P2P.

<sup>269</sup> Entretien avec Dietrich Poell, 23 septembre 2009.

<sup>270</sup> Entretien avec K. Gueden.

<sup>271</sup> Le système de codage utilisé par Wuala se reporte aux codes de Reed-Solomon, un code correcteur dont le principe est de suréchantillonner un ensemble de symboles ou données à transmettre dans un système. La redondance que comporte ce suréchantillonnage permet au récepteur des données encodées de les reconstruire, même s'il y a eu des erreurs pendant la transmission ([http://en.wikipedia.org/wiki/Reed-Solomon\\_error\\_correction#Data\\_storage](http://en.wikipedia.org/wiki/Reed-Solomon_error_correction#Data_storage)).

<sup>272</sup> Entretien avec K. Gueden.

<sup>273</sup> La *Google Tech Talk Series* est une série de conférences techniques, hébergées par Google dans son site central de Mountain View, en Californie. La plupart d'entre elles sont enregistrées et disponibles

*« We ended up using erasure codes. Without erasure codes, it wouldn't work, because the necessary redundancy factor would be much too high. Erasure codes, on the other hand, have the nice property that you can encode n fragments into m, and then you need only n out of these m to reconstruct the file, but it doesn't matter which<sup>274</sup>. »*

Ce choix entraîne par ailleurs des questions d'« entretien » du système. Assurer la redondance du système oblige à gérer la disponibilité des données non seulement d'un point de vue spatial (tous les fragments présents dans au moins un ordinateur), mais aussi dans le temps : à cause des déconnexions prolongées et inattendues, et des déconnexions permanentes, la disponibilité des certaines ressources matérielles de stockage peut cesser définitivement à un moment. Comme les développeurs s'en rendront compte, à la suite des réactions des utilisateurs du *closed alpha* après ses premiers mois de vie, il s'agit là d'un aspect tout aussi important pour la stabilité du système et sa durabilité à moyen terme :

*« We then figured out how important maintenance was. When I store five hundred fragments today, it does not mean they will be the same five hundred fragments that will be available in a month from now, for instance. So, we introduced the possibility for Alice's computer to check and maintain her files, the computer periodically checks whether file availability is still good enough, and if it seems that some computer has left the system forever, then if necessary, it reconstructs new fragments and upload them to the network<sup>275</sup>. »*

« FRAGMENTS ENCRYPTÉS » ENTRE LOCAL ET GLOBAL, UNE PROMESSE DE SÉCURITÉ ?

Le choix de baser le stockage P2P sur le mécanisme des codes d'effacement déterminera aussi, de façon que Dietrich n'hésite pas à définir « naturelle<sup>276</sup> », un autre aspect de la technologie Wuala qui fera débat : l'encryptage en local, sur la propre machine de l'utilisateur et au moyen du client P2P Wuala, des contenus à stocker – avant que ceux-ci ne soient divisés en fragments, dupliqués et repartis au sein du système. Les « fragments encryptés » résultant de ce processus deviendront tour à tour – suivant les mots des développeurs Wuala, des experts techniques qui

---

au public à <http://research.google.com/video.html>.

<sup>274</sup> Extrait de la Google Tech Talk Wuala par Dietrich Poell, transcrite par l'auteure. Dietrich me donnera aussi un exemple « quantifié » au cours de nos entretiens : « *Say there are 100 original fragments which are encoded into 500. These 500 are stored onto so called storage nodes, which are computers that are online for at least 17% of their time and are willing to trade local storage (not everybody has to, you start with 1 GB which is provided by us, which is sufficient for a lot of users). At any point of time, you only need 100 out of these 500 to reconstruct the file. If you do the math, this gives you an almost-complete availability.* »

<sup>275</sup> Entretien avec K. Gueden.

<sup>276</sup> Entretien avec D. Poell.

donneront leur avis sur les premiers pas du dispositif, et de sa base d'utilisateurs pionniers – le symbole même de la solution de stockage permanent en P2P proposée par la firme, solution sur laquelle repose une bonne partie de la crédibilité et de la spécificité du produit ; le mécanisme technique qui assure la sécurité des données par rapport non seulement aux utilisateurs non autorisés, mais à la firme elle-même ; ou encore, l'arène où va se jouer la controverse entre utilisateurs « méfiants » et « confiants » sur la capacité du dispositif de maintenir la promesse de sécurité et confidentialité accrue faite par ses développeurs.

Les fonctions d'encryptage et de fragmentation des fichiers à stocker sont programmées par l'équipe Wuala pour entrer en fonction dans tout ordinateur qui, ayant installé le logiciel, a sélectionné l'option de stockage partagé. Le fait de sélectionner cette option n'est pas nécessairement un choix obligé : au cours de ses premiers mois d'existence, Wuala était programmé pour donner la possibilité aux utilisateurs ayant besoin de moins d'un gigaoctet de mémoire stockage de ne pas l'obtenir du réseau d'ordinateurs connectés, mais des serveurs de la firme (« *You start with 1 GB provided by us* ») : en conséquence, dans ce cas, les fragments encryptés n'étaient pas stockés sur l'ordinateur de cet utilisateur<sup>277</sup>. Cette option, faisant partie des mesures prises par l'équipe pendant la phase d'amorçage, ne rencontre cependant qu'un intérêt mitigé de la part des usagers de la première heure ; comme souligne Dietrich,

*« people who tried Wuala were interested in the P2P storage mechanism, and that involved the storage trading, thus, the encryption<sup>278</sup>. »*

Un profil type de client/utilisateur pionnier de Wuala commence à se dessiner à travers les mots de Dietrich : il s'agit de quelqu'un avec un intérêt pour le mécanisme de stockage en P2P, son fonctionnement technique et sa dimension communautaire à la fois – quelqu'un qui « suit » le dispositif, parce qu'il possède les capacités techniques et l'expertise pour le faire.

Le logiciel est programmé pour fonctionner de la manière suivante : quand un usager dépose un fichier dans Wuala, ce fichier est d'abord encrypté sur son ordinateur. Ensuite, le fichier crypté est divisé en fragments (« non exécutables, » note Dietrich – fragments qui, pris un par un, sont dépourvus de signification et d'autonomie), qui sont indexés dans le dossier « données » du logiciel et encodés de manière redondante, encore une fois sur la machine de l'utilisateur. A ce moment, les fragments encryptés sont téléchargés à différents endroits du réseau P2P, où ils sont stockés. Dietrich résume lors de la Google Tech Talk d'octobre 2007 :

*« Let's say that Alice wants to store a file, for example, a roadtrip videoclip. The first thing that happens when Alice drags the file into Wuala is that the file gets encrypted on her computer. Then the encrypted file is split into fragments, and these fragments are redundantly encoded into other*

---

<sup>277</sup> Foire aux questions, site Wuala, <http://www.wuala.com/en/support/faq/c/16>.

<sup>278</sup> Entretien, D. Poell.

*fragments<sup>279</sup>. Now let's say that Alice wants to share the file with Bob. What Bob then tries to do is to find end fragments, a subset of all fragments, from the P2P network. (Now, at the moment of the closed alpha, if he doesn't find enough fragments in the P2P network, he gets the remaining fragments from our servers. Then it will be self-supporting, making it possible to retrieve all the data from the network). Then the application would decode, decrypt, and open the file. So these are the basic steps when you upload and download a file<sup>280</sup>. »*

On voit dans ce processus une articulation intéressante des dimensions locale et globale qui formate le rôle de l'utilisateur envisagé par le dispositif. Jusqu'à ce que l'encryptage et la fragmentation des données ne soient accomplis, le cœur du service est le terminal de l'utilisateur : là ont lieu, au moyen du client P2P, les opérations d'encryptage et de division en fragments. Comme le souligne Kurt,

*« All encryption and decryption is performed locally – the advantage of having your software running on the client<sup>281</sup>. »*

Une fois que les données sont sous forme de fragments encryptés, ceux-ci sont stockés et partagés, « éparpillés » sur le réseau. Lorsque s'effectue le passage de la dimension locale de la machine utilisateur à l'ensemble du réseau, l'emplacement physique des contenus stockés (ou plutôt les multiples emplacements de ces contenus) est constitué par des sous-ensembles d'ordinateurs, qui forment un réseau constamment provisoire et évolutif. Plutôt que l'ensemble des ordinateurs, la dimension « globale » du système est donc l'algorithme qui permet à chaque moment de définir quels ordinateurs, à chaque moment et pour quelle durée, recueillent les données d'un utilisateur en particulier.

En même temps, à moins que des opérations de partage et interrogation ne soient déclenchées explicitement, chaque utilisateur ignore quels fragments, à un moment, sont stockés sur sa propre machine parmi les autres. Un fichier partagé de cette façon sur le réseau est donc, en fait, invisible par les utilisateurs, même si leurs machines en stockent des parties ; les utilisateurs n'ont une connaissance directe que de leur mise à disposition des ressources computationnelles et matérielles nécessaires à l'encryptage et la fragmentation des fichiers, ainsi qu'à leur stockage. La capacité CPU de leurs machines se mobilise, mais ils ne sauront pas à quelle tâche elle est précisément destinée à un moment donné. C'est ainsi que, dès la

---

<sup>279</sup> Au moyen des codes d'effacement, comme on a vu.

<sup>280</sup> Extrait de la Google Tech Talk Wuala par Dietrich Poell, transcrite par moi. On rendra compte plus tard dans le chapitre des différentes modifications – auxquelles se réfère la parenthèse dans l'extrait – dont ont fait l'objet, au fil du temps, des centres de serveurs utilisés tour à tour comme sauvegarde de sécurité, comme « nœuds stables » afin d'assurer la redondance pendant la phase d'amorçage (c'est-à-dire, jusqu'à ce que le nombre d'utilisateurs adoptant le système ne devienne suffisant à assurer sa stabilité), et auxquels la firme Wuala a progressivement réattribué une importance à la fois symbolique et très « matérielle » dans l'économie du dispositif.

<sup>281</sup> Entretien avec K. Gueden.

première version du dispositif, l'équipe Wuala estime nécessaire d'insérer dans les conditions d'utilisation du produit une clause très particulière, qui clarifie quelles ressources pourront être utilisées par le système et pour quel ensemble de tâches – tout en ne précisant pas quelle ressource servira quelle tâche, et comment, à un moment précis :

**General Terms and Conditions.** *Date of last revision: May 11, 2009.*

*(...) 5. Services of the User : The user acknowledges that Wuala may use processor, bandwidth and harddisk (or other storage media) of his computer for the purpose of storing, encrypting, caching and serving data that has been stored in Wuala by the user or any other users. The user can specify the extent to which local resources are used in the settings of the Wuala client software. The amount of resources the user is allowed to use in Wuala depends on the amount of local resources the user is contributing to Wuala. Resources are allocated and monitored in accordance with the Privacy Policy.*

Tableau 6. Extrait des conditions d'utilisation de Wuala, consultées à <http://www.wuala.com/en/about/terms>, version de mai 2009.

CONFIANCE ET FERMETURE DU CODE SOURCE : LA SÉCURITÉ « PAR OBSCURITÉ » EN DÉBAT

Depuis ses débuts en 2007, Wuala fait du niveau de sécurité élevé, promis par l'encryptage et la fragmentation des fichiers en local, une des pierres angulaires du service, au point que les informations fournies sur le site de la firme au sujet de la technologie développée par Wuala coïncident bientôt avec une page intitulée « Sécurité »<sup>282</sup>. Celle-ci met l'accent, selon les cas, sur des aspects de confidentialité (des données et des activités), de qualité de service (l'intégrité des données stockées), la permanence et la continuité du service (la distribution et la redondance des contenus stockés) : autant de termes qui renvoient à des acceptions de sécurité différentes. Lors de la Google Tech Talk d'octobre 2007, Dietrich met encore une fois la question de la sécurité au centre de son discours. Au milieu de considérations techniques poussées sur la structure cryptographique du dispositif, l'importance de la dimension locale des processus de sécurisation revient avec insistance et apparaît encore une fois comme un des points de force affichés de son système de stockage :

*« The system we use is 128 bit AES for encryption, and 2048 bit RSA for authentication; all data is encrypted, which means that files and metadata is encrypted, and all cryptographic operations are performed locally, which means on your machine, on your computer. For access control, we've come up with a simple but quite effective cryptographic tree structure, which is*

<sup>282</sup> Page « Sécurité » du site Wuala, <http://www.wuala.com/fr/learn/technology>.

*based on the definition of untrusted storage, which means that nothing is revealed as to the place where the access control structure is being stored, all operations are performed on the client and the storage is not involved in any cryptographic operations. Also it doesn't reveal anything to those who have access, so if two people have access, they have access to the folder but they don't know of each other<sup>283</sup>. »*

Plusieurs commentateurs et blogueurs notent également cette mise en relief de l'aspect sécurité de la part de Wuala. Certains d'entre eux semblent par ailleurs signaler entre les lignes que la sécurité de Wuala suppose, au moins à son stade de closed alpha (2007), un bon degré de confiance des utilisateurs dans la fiabilité de ces discours relatifs à l'encryptage :

*« Wuala is keen to talk up security, pitching its wares based on the fact that it employs encryption locally before files are sent to the Cloud and/or backed up via a peer-to-peer network in which other users also provide disk space on their own computers. The use of client-side encryption, says the company, means that even Wuala itself cannot access a user's files because the password<sup>284</sup> never leaves the device<sup>285</sup>. »*

La controverse sur la sécurité « en local » de Wuala est tout particulièrement marquée par une attaque assez véhémente de la part du cryptographe Colin Percival, un développeur et chercheur en informatique dont le blog, *Daemonic Dispatches*, est une des références sûres pour les experts du domaine<sup>286</sup>. En octobre 2007, l'expert publie sur son blog une critique acérée du dispositif, se basant sur une affirmation de Dietrich lors d'une interview :

*« In our system, everything is encrypted and the encryption is used by the CIA for top secret files ».*

Celle-ci sera ensuite promptement rectifiée par l'auteur de l'interview en

*« In our system, everything is encrypted and it is state-of-the-art encryption<sup>287</sup>, »*

---

<sup>283</sup> Extrait de la Google Tech Talk Wuala par Dietrich Poell, transcrite et partiellement soulignée par l'auteur.

<sup>284</sup> La clé d'encryptage.

<sup>285</sup> Mike Butcher (2008). « You Have Three Days to Check Out Wuala's 'Social Grid' Storage », *TechCrunch*, <http://techcrunch.com/2008/07/18/you-have-three-days-to-check-out-wualas-social-grid-storage/>

<sup>286</sup> Colin Percival (21 octobre 2007). « Wuala : Willful ignorance, or fraud ? », « Daemonic Dispatches » Blog, <http://www.daemonology.net/blog/2007-10-21-wuala-willful-ignorance.html>

<sup>287</sup> Allen Stern (2007). « Interview with Wuala CEO/Founder Dietrich Poell », *CenterNetworks*.

mais pas avant que Colin Percival ne soit parvenu à des conclusions aux implications assez destructrices pour la confiance de l'utilisateur lambda dans les aspects relatifs à l'encryptage et la sécurité du *closed alpha* de Wuala :

*« All told, I think Dietrich Poell falls more into the category of 'willfully ignorant' – he doesn't understand cryptography, and he apparently hasn't made any attempt to consult people who do. But whether he's deliberately lying about the security of Wuala or actually believes what he claims – that the encryption used by Wuala is used by the CIA for top secret files – doesn't really matter in the end: If you care about your data, don't trust him with it<sup>288</sup>. »*

Cette attaque sans détours vaut à Percival quelques interrogations de la part de ses lecteurs sur son objectivité :

*« a whole page calling wuala fraud is overkill. ive been a member for quite a while. have you ever even used wuala?<sup>289</sup> »*

Et, comme on pouvait s'y attendre, Dietrich contacte Colin en lui demandant de prendre en compte la rectification du passage de l'interview, et en fournissant des détails techniques additionnels sur la cryptographie qui ne sont pas présents sur le site de la firme. En particulier, le papier décrivant Cryptree, le système de cryptographie à l'œuvre dans Wuala (Grolimund *et al.*, 2006), fait pour la première fois son apparition dans un contexte non académique<sup>290</sup>, et un lien direct au papier figure depuis lors sur la page Sécurité du site Wuala<sup>291</sup>. Ce moment d'épreuve, que Dietrich décrira plus tard comme fondamental dans la stabilisation d'une légitimité pour le processus de sécurisation à l'œuvre dans Wuala<sup>292</sup>, est résolu en deux temps. D'abord, par un post de Percival : le cryptographe continue à souligner la grande complexité du système et les problèmes qu'une telle complexité peut soulever, tout particulièrement pour une technologie naissante,

*« in cryptography, just like every other area of computing, more lines of code inevitably means more bugs and more potential security flaws<sup>293</sup> »,*

---

<sup>288</sup> Percival, cit.

<sup>289</sup> Discussion de l'entrée de blog de Colin Percival, [http://www.daemonology.net/blog/2007-10-21-wuala-willful-ignorance.html#disqus\\_thread](http://www.daemonology.net/blog/2007-10-21-wuala-willful-ignorance.html#disqus_thread)

<sup>290</sup> Il avait été présenté au Symposium on Reliable Distributed Systems (SRDS) en 2006.

<sup>291</sup> <http://www.wuala.com/fr/learn/technology>

<sup>292</sup> Entretien avec D. Poell : *« We felt he was the one making unsubstantiated claims, but he was an expert in just the thing we were trying to sell, an effective encrypting system. It was very important that people were convinced, and it implied convincing him first. »*

<sup>293</sup> Colin Percival (26 octobre 2007). « Wuala Update. » « Daemonic Dispatches » Blog, <http://www.daemonology.net/blog/2007-10-26-wuala-update.html>

mais note aussi qu'il a reçu plusieurs retours montrant que la confiance des utilisateurs semble se stabiliser, à la suite de la communication plus ouverte et complète sur les détails techniques du fonctionnement du système que Wuala a mis en œuvre :

*« as Dietrich has pointed out, they're still in alpha testing, and have lots of time to get the security details right before Wuala enters widespread use. Dietrich certainly seems interested in fixing the problems (which is more than I can say of certain other people) and is providing information about how he is doing it, so we'll have to wait and see – hopefully the Wuala website will keep on being updated to describe how all of these issues are being addressed, so that Wuala's security can be assessed without needing to get into an email exchange<sup>294</sup>. »*

Une deuxième entrée de blog de Colin Percival, une année plus tard, entre dans les détails techniques pour expliquer comment le système de sécurisation de Wuala est en train de prendre forme : bien qu'exprimant toujours des réserves, le ton de l'expert en cryptographie est dans son ensemble rassurant pour les utilisateurs<sup>295</sup>.

Une de ses remarques porte pourtant une des premières traces d'une autre controverse, qui, surtout depuis l'ouverture de la version beta publique du logiciel, se révèle importante pour la crédibilité du système « fragments encryptés ». Il s'agit du débat sur la fermeture du code source du client P2P de Wuala : pour tout développeur ou utilisateur averti qui souhaiterait y voir plus clair dans la définition de sécurité proposée par Wuala, il n'est pas possible d'explorer le code, qui reste propriétaire et fermé.

*« The source code to the Wuala client is not available. While I have no reason to think that Dietrich or anyone working at Wuala has inserted any evil code into the Wuala client, at present there is no way to know. Somehow 'hey, download and run this program! It will back up all of your data over the internet, and it's perfectly safe because it encrypts everything first. Really it does! Trust me!' isn't a sales pitch which inspires a lot of confidence in me<sup>296</sup> »,*

conclut ironiquement Percival. De son côté, l'équipe de développeurs Wuala s'était confrontée, depuis les temps où Wuala s'appelait encore Kangoo, à la question de l'ouverture ou de la fermeture du code source. En 2005 et 2006, quand Dietrich et

---

<sup>294</sup> Id.

<sup>295</sup> Colin Percival (7 novembre 2008). « Wuala's Improved Security. » « Daemonic Dispatches » Blog, <http://www.daemonology.net/blog/2008-11-07-wuala-security.html>: « Over the following eight months, the people at Wuala made significant improvements (...) recognize that ultimately you're completely trusting Dietrich Poell and the Wuala staff with your computer. But (...) Wuala's security now looks fine. »

<sup>296</sup> Id.

Kurt sont encore des doctorants à l'ETH Zurich et écrivent les premiers bouts de code de Kangoo, ils sont immergés dans une culture académique de recherche ouverte et de développement de code en open source. Ils font usage, pour la première version du client Kangoo, de parties de code dérivées de plusieurs projets open source<sup>297</sup>, et des composantes du futur Wuala sont développées et brevetées en tant qu'open source – Wuala Webstart, logiciel pour le démarrage rapide de l'application, et Wuala Persistent Map, un prototype de la table de hachage distribuée qui sous-tendra Wuala, sont rendus disponibles sous licence LGPL (Lesser General Public License) sur le site SourceForge<sup>298</sup>.

Pourtant, quand Wuala sort du statut de projet universitaire pour devenir un « embryon » de service, en 2007, la décision est prise de fermer le code source de l'application, en *closed alpha* d'abord et en beta privé et public depuis, bien que le site de la firme continue à reconnaître et à lister les contributions open source encore présentes dans le code source du dispositif : « *We love open source technology and Wuala is based on a number of open source projects* », explique le site de Wuala<sup>299</sup>. A l'automne 2009, quand je m'enquiers de la cause de ce choix auprès de Dietrich, celui-ci m'explique qu'il s'agit, au moins jusqu'à ce que la firme précise son modèle d'affaires et recrute du personnel additionnel, d'une question de disponibilité de ressources, ainsi que de quantité et qualité de la « maintenance » et du dépannage éventuel que lui et ses collaborateurs peuvent consacrer à Wuala :

*« While having an open source client clearly has advantages, it requires a lot of resources to support and maintain the community around it. We currently don't have the necessary resources and thus open-sourcing the whole Wuala client is out of the question for the moment<sup>300</sup>. »*

C'est pourtant l'importance attribuée par l'équipe Wuala elle-même à la sécurité « augmentée » assurée par la fragmentation et l'encryptage des fichiers qui amène certains utilisateurs pionniers, de façon toujours plus pressante, à demander l'ouverture du code source. De façon intéressante, une des premières arènes où cette demande est formulée est, en mars 2010, le forum utilisateurs de Dropbox, un des principaux concurrents de Wuala qui implémente un modèle de stockage en cloud « classique<sup>301</sup> » ; Reginald O., un utilisateur qui paraît spécialement versé en techniques de cryptographie, en arrive à définir l'approche de fermeture du code

---

<sup>297</sup> Entre autres Apache Commons, Jakarta et Trove (la liste complète est disponible sur le site de Wuala à <http://wuala.com/fr/about/thirdpartycode>).

<sup>298</sup> A <http://sourceforge.net/projects/wualawebstart/> et <http://sourceforge.net/projects/quickbase/> respectivement. SourceForge est le plus large répertoire existant d'applications et logiciels open source.

<sup>299</sup> <http://wuala.com/fr/learn/technology>

<sup>300</sup> Entretien avec D. Poell.

<sup>301</sup> C'est à dire que la firme Dropbox possède un certain nombre de serveurs auxquels la totalité des fichiers et des données sauvegardés par les utilisateurs Dropbox sont envoyés. Voir <http://www.dropbox.com/>.

source de Wuala de la « sécurité par obscurité » :

*« Wuala has been known to at one point use AES-128 with ECB. ECB (electronic codebook) is bad because a would-be attacker, if watching the data, can recognize repeated blocks of data. While this has been fixed, (...) it is very much possible for employees of Wuala to know what files you have since meta data about the files are stored on the servers along with their hashes. It is very difficult to validate what Wuala is actually doing because their code and methods are closed source. Thus, you have to trust what they are saying is correct (rather than obtain independent review). (...) To me, this screams 'security by obscurity.' Their code, encryption, etc., are all proprietary with no review. For me, that causes me to distrust them<sup>302</sup>. »*

Au contraire de ce qui est arrivé dans le cas de l'entrée de blog controversée de Colin Percival, aucun des développeurs de Wuala n'intervient directement pour répondre à Reginald. Cependant, la même question est posée à nouveau, peu de temps après, sur le forum de Wuala, et les motifs de la fermeture du code source font, une fois de plus, objet de débat.

L'utilisateur jernst lance la discussion en mettant à nouveau en avant des questions de confiance. Il note qu'il s'agit là d'une question centrale, puisque la sécurité renforcée est un des aspects différenciants de Wuala, suite à un choix précis de ses développeurs. Mais il est plus difficile d'avoir confiance s'il n'y a pas de moyen de rentrer dans le code ou le vérifier :

*« One of the main strength of Wuala (for me anyway) is its security. However we currently have to trust Wuala on this topic. Having an open source client would have many advantages : peer review of the security and code, faster port to new platforms, faster development and bug fixes. »*

À défaut d'un code véritablement « ouvert » à la modification, par une licence libre, un autre utilisateur revendique, plus bas dans la même discussion, un code publiquement accessible et consultable :

*« Even when you don't release the source code under a GPL license it would still be great if you would at least publish the source code to allow people to check your promise of security. Anyone specifically choosing Wuala over other solutions cares about the encryption. Isn't it (this group) who really pays your bills? Isn't it (this) group with the willpower to bring their friends and family into using Wuala instead of DropBox, et al. ? »*

La conclusion de l'utilisateur Christian remarque, encore une fois, que c'est la capacité de Wuala à se différencier d'un point de vue technique et économique qui est en jeu :

---

<sup>302</sup> Forum Dropbox, « Wuala vs dropbox ». Mars 2010.  
<http://forums.dropbox.com/topic.php?id=40403>

*« It's not only trusting that you aren't evil but also trusting that you aren't making mistakes in the encryption process. Without having a higher security there isn't much reason to use wuala over something like dropbox<sup>303</sup>. »*

Sous la pression, Philipp, à ce moment l'un des membres plus récents du staff Wuala, donne une intéressante réponse aux partisans de l'open source :

*« Actually, we have no reason to lie to you here: it simply is not our business model, an open source community you have to nurture as well as get from it. (...) Open Source is ONE way to create confidence. But you trust your bank, for example, even though you can't go and look at their coffers. The difference between Wuala and Dropbox is that we have made sure that we do not even have the technical means to betray your trust. To stay with the banking metaphor: Dropbox is like a bank account. If the bank goes bust, you lose your money. Wuala is like a custody account, where you retain control of your data, even if we are compromised. »*

Plusieurs éléments sont présents dans la réponse du développeur Wuala. Tout d'abord, Philipp propose une redéfinition de la confiance que les utilisateurs cherchent à établir au moyen de l'ouverture du code, en suggérant une alternative qui évoque la « privacy par l'architecture » négociée par Faroo<sup>304</sup>. Comme l'avait fait Georg Schneider, le représentant de Wuala met l'accent sur les contraintes posées par le modèle technique de choix des développeurs. Un modèle technique se dessine où l'encryptage en local d'abord, et la fragmentation après – avant toute opération de partage ou téléchargement qui implique une circulation des données dans le réseau – fournissent, selon Wuala, la preuve que la firme « n'aurait même pas les moyens techniques » de trahir la confiance des utilisateurs. Dans le pire des cas – celui où le réseau serait compromis – un éventuel attaquant en retirerait seulement des fragments non exécutables, et sans signification indépendante.

Cet argument n'est pas une réponse directe à la question de l'ouverture du code, mais plutôt une négociation de ce qui est important pour l'entreprise et qui pourrait aussi l'être pour ses clients. Il vient renforcer une considération qui a trait au modèle économique de la firme, comme Dietrich l'avait souligné antérieurement. Une communauté de développeurs *open source* « auxiliaires » est, pour une firme avec des disponibilités de ressources financières et humaines très limitées, autant un investissement qu'une ressource : les retours, les critiques et les apports nécessitent de temps et d'attention afin d'être gérés, triés et incorporés au projet principal. La fermeture du code source est donc finalement présentée comme une décision pratique, que des développeurs « purs », baignant dans une culture de logiciel libre, n'auraient pas pris. Mais les entrepreneurs qu'ils sont devenus doivent prendre cette décision – peut-être temporairement, peut-être plus par nécessité que par choix – pour la durabilité de l'entreprise en considération de ses possibilités.

---

<sup>303</sup> Forum Wuala, « Opensource client ». 2010 et 2011.  
<https://forum.wuala.com/viewtopic.php?f=11&t=514>

<sup>304</sup> Voir le chapitre 4, Section 4.

## SECTION 2. STOCKAGE « SOCIAL » ET PRIVACY

« *In 2007, it was all starting to get social, »*

se remémore deux ans plus tard Dietrich Poell<sup>305</sup>. Cette référence à l'explosion des médias sociaux, qui, Facebook et Twitter en tête, entraînent à ce moment de façon de plus en plus pervasive dans le quotidien de millions d'internautes, situe les « premiers pas » de Wuala dans un contexte et une communauté de recherche et de développement qui, tout particulièrement depuis 2005, travaille à la mise au point de systèmes hybrides, au croisement entre réseaux P2P et réseaux sociaux<sup>306</sup>.

En 2007, Facebook existe depuis désormais trois ans. Des millions d'internautes y participent, et contribuent au succès massif de ces services basés sur le Web qui permettent aux individus de construire un profil public ou semi-public au sein d'un système fini, de définir une liste d'autres utilisateurs avec lesquels entretenir une relation, et de voir/parcourir la liste de leurs connexions et celles faites par d'autres au sein du système (Boyd & Ellison, 2007). En parallèle à leur croissance spectaculaire, ces réseaux sociaux suscitent des vives discussions et controverses, à la fois parmi les experts et le grand public, au vu des usages que leurs administrateurs font des données privées des utilisateurs, parfois en permettant à des applications extérieures d'y accéder, et parfois à des fins commerciales directes (Boyd, 2008). Des voix se lèvent, à la fois parmi les chercheurs en informatique et en droit des nouvelles technologies (Acquisti & Gross, 2006 ; Elkin-Koren, 2006), pour souligner que peu d'utilisateurs ont conscience du fait qu'en utilisant ces applications, ils laissent ouvert l'accès, par des publics non ou mal définis, à certaines informations privées conservées sur les serveurs des entreprises qui proposent les services – comportement qui est défini comme risqué par plusieurs commentateurs<sup>307</sup>.

Dans ce contexte, plusieurs développeurs, parmi lesquels figure l'équipe Wuala, travaillent au croisement de deux grandes « tendances d'utilisation ». D'un côté, ils identifient dans une architecture de réseau de type *peer-to-peer* une façon possible d'aborder sous un angle novateur les questions de *privacy* des données personnelles – par leur relocalisation et « réappropriation » dans les terminaux des utilisateurs, qui pourraient héberger leur propre profil et les informations qu'il contient<sup>308</sup>. De l'autre, ils voient dans la modélisation au niveau architectural des phénomènes qui assurent alors le succès retentissant des réseaux sociaux – tels que les liens d'amitié, la formation de communautés et groupes, l'attribution de différents degrés de

---

<sup>305</sup> Entretien avec D. Poell.

<sup>306</sup> Cette section reprend mes articles [Musiani, 2010a] et [Musiani, 2010b].

<sup>307</sup> Par exemple, Michael Arrington (2007). « Facebook Privacy Issue Won't Die ». TechCrunch, <http://techcrunch.com/2007/11/26/facebook-privacy-issue-wont-die/>

<sup>308</sup> Voir aussi [Moglen, 2010] et [Aigrain, 2010, 2011].

confiance – une façon de remédier à certaines fragilités des réseaux P2P classiques : afin de rendre filtrages, blocages et identifications plus difficiles, ceux-ci maintiennent autant que possible l'anonymat des utilisateurs même quand la connaissance réciproque de l'identité entre nœuds communicants aurait rendu le réseau plus robuste (Figueiredo et al., 2008).

Dietrich et son équipe travaillent donc sur quelque chose de « *fashionable* », comme remarque ironiquement Kurt<sup>309</sup>. L'idée de faire de Wuala un « *social P2P storage*<sup>310</sup> » remontait cependant aux premières phases du projet Kangoo, qui précède, bien que de peu, la véritable explosion des réseaux sociaux en Europe. Le mémo « Kangoo », que Dietrich avait soumis au VentureLab en 2006, contient déjà l'idée de continuité entre la fonction de stockage et celle de partage, et souligne l'aptitude de la technologie P2P à soutenir cette continuité:

*« A user can securely store his private files online, so that he can conveniently access them from anywhere at any time. More importantly, he can also use it to share his files such as photos, digital videos, music, or documents with his friends (...) The openness and distribution of the system allows for a multitude of possibilities. Services such as Yahoo Photo, Flickr, YouTube, and YouSendIt impressively demonstrate the need for free online storage. Thanks to the massive sales of digital photo and video cameras and the continuing spread of software production tools for images, videos and music, this need will constantly grow. More than 100 million Internet users are already sharing their digital photos online today, and a rapidly growing number of people are sharing their video clips online<sup>311</sup>. »*

Les développeurs de Wuala travaillent donc sur une architecture qui puisse mêler une structure distribuée, avec le déroulement de certaines opérations de sécurisation et conservation de données en local, à des dynamiques propres au réseautage social, dans lequel des liens directs – directs dans les « couches inférieures » aussi bien qu'au niveau de l'interface – peuvent être établis entre les utilisateurs pour les opérations de partage. Les infrastructures de réseautage social étant parfaitement aptes à découvrir et établir des liens *sociaux*, mais très peu aptes à permettre des connexions entre un usager et ses pairs au moyen de liens *de réseau* (liens « matériels », au niveau de l'architecture), le défi devient pour les développeurs d'envisager une architecture novatrice, capable d'intégrer le réseautage aux deux niveaux de l'interface et de l'application<sup>312</sup>.

Encore une fois, c'est une conception de la confidentialité qu'on a déjà appelé

---

<sup>309</sup> Entretien avec K. Gueden.

<sup>310</sup> Frederic Lardinois (2008). « Wuala's Social Storage comes to the Web ». ReadWriteWeb, [http://www.readriteweb.com/archives/wualas\\_social\\_storage\\_web\\_interface.php](http://www.readriteweb.com/archives/wualas_social_storage_web_interface.php)

<sup>311</sup> Extrait du Memo « Kangoo » présenté par Dietrich Poell à venturelab.ch en mai 2006.

<sup>312</sup> Voir aussi [Figueiredo et al., 2008].

« privacy par la technique » qui se dessine dans la démarche de l'équipe Wuala et que l'entreprise souhaite proposer à ses clients. Cette conception de la *privacy* prend forme avec les contraintes et les opportunités liées au traitement et à l'emplacement des données, selon les moments et les opérations se déroulant au sein du système ; en particulier, avec le rapport entre la confidentialité des données (personnelles, mais aussi les contenus stockés dans le cloud P2P) et le système d'allocation de ressources sur lequel se base le dispositif. Elle s'explique tout particulièrement dans les discussions et les évolutions du rôle, dans l'économie du système, de deux « objets » : le mot de passe qui identifie l'utilisateur vis-à-vis du réseau, et la clé d'amitié qui sert à deux utilisateurs pour échanger des contenus entre eux.

#### MOT DE PASSE ET RESPONSABILITÉ DE L'USAGER

Dans les intentions de Dietrich, le rôle du mot de passe sélectionné par l'utilisateur pour le système Wuala devait être quelque chose qui

*« would strike the user as soon as he had access to the system the very first time<sup>313</sup> ».*

En effet, le formulaire virtuel qui s'ouvre lors de la première souscription au service peut surprendre l'utilisateur : il informe que

*« Nous ne connaissons pas votre mot de passe car il ne quitte jamais votre ordinateur. S'il vous plaît, n'oubliez pas votre mot de passe et utilisez, le cas échéant, l'indice de mot de passe<sup>314</sup>. »*

Le statut du mot de passe est donc négocié, au delà de sa signification habituelle d'identifiant unique vis-à-vis du système, pour définir, préciser et « légitimer » le processus d'encryptage et décryptage en local des données du système Wuala. Sous cet aspect se concentrent la promesse de sécurité et de *privacy* de Wuala et la confiance de l'utilisateur, puisqu'il devient le symbole et la représentation graphique de la dimension « locale » du processus d'encryptage, du fait qu'il ne quitte pas l'ordinateur de l'utilisateur qui l'a créé. Les opérations, pour la plupart gérées de façon automatique, liées à la protection des données personnelles sont donc hébergées sur la propre machine de l'utilisateur. Celui-ci voit sa place au sein de l'architecture de l'application modifiée : nœud dans un réseau de nœuds équivalents, il devient lui-même un serveur, plutôt qu'un point de départ et d'arrivée pour des opérations qui sont par ailleurs conduites sur une autre machine.

En attribuant ce statut au mot de passe, les développeurs de Wuala proposent aussi

---

<sup>313</sup> Entretien avec D. Poell.

<sup>314</sup> Voir aussi FAQ de Wuala, "Security and Privacy", à [http://www.wuala.com/fr/support/faq/c/20#a\\_20](http://www.wuala.com/fr/support/faq/c/20#a_20).

un rééquilibrage entre les droits que l'utilisateur exerce sur ses données et les droits acquis par le fournisseur du service sur ces mêmes données, au moyen d'un compromis. Puisque le mot de passe demeure chez l'utilisateur et n'est pas envoyé aux serveurs de la firme, celle-ci ne peut pas récupérer le mot de passe si besoin : l'utilisateur ne voit donc pas seulement sa *privacy* renforcée, mais en même temps et pour les mêmes raisons, la responsabilité de ses actions dans le système se voit augmentée – tandis que le fournisseur renonce à une partie de son contrôle sur les contenus qui circulent grâce au service qu'il gère. Les conditions d'utilisation de Wuala, plus spécifiquement, sa politique de confidentialité, rendent ce compromis officiel :

*Wuala encrypts and decrypts all files locally. Your password never leaves your computer, hence, we do not know your password. This protects your privacy, but also means that apart from sending you password hint (specified when creating your account or check your account settings) we cannot help you recovering it. (If you forget your password, you can request your password hint. If your password hint does not help you, we cannot do anything about it. Since your password never leaves your computer, we do not know your password. We cannot recover your password<sup>315</sup>.*

La signification de cette « renonciation », m'explique Dietrich, est double : d'un côté, l'équipe Wuala souhaitait rendre évident, presque « traduire » dans un objet spécifique, auquel l'utilisateur puisse aisément faire référence, le processus « obscur » et peu familier de l'encryptage en local, qui, comme on l'a vu, suscite des controverses et des perplexités<sup>316</sup>. De l'autre côté, qui est souligné par Philipp au cours du débat sur la fermeture du code source analysé précédemment, il est aussi question du modèle économique de Wuala : plus l'entreprise connaît d'éléments par rapport à ses utilisateurs, plus elle est obligée de soumettre ceux-ci à une surveillance et un contrôle périodiques, ce qui nécessite, encore une fois, un investissement de ressources et de temps :

*« If we can know what is in your account, starting with your password, we have heightened obligations to police the content and to make sure nobody can eavesdrop on the traffic<sup>317</sup> ».*

---

<sup>315</sup> Wuala Privacy Policy, <http://www.wuala.com/en/about/privacy>

<sup>316</sup> Entretien avec Dietrich Poell.

<sup>317</sup> Forum Wuala, « Opensource client ». 2010 et 2011. <https://forum.wuala.com/viewtopic.php?f=11&t=514>

## LIENS DIRECTS ET « CLÉ D'AMITIÉ »

Le deuxième point critique auquel l'équipe se trouve confrontée au cours de sa démarche de développement est le moment où les contenus que l'utilisateur introduit dans le système, une fois encryptés et fragmentés, quittent la dimension locale du client P2P installé sur l'ordinateur de chaque utilisateur, pour être dispersés et stockés dans le réseau P2P – l'ensemble des ordinateurs des usagers – et les centres de données Wuala. Comme le souligne Kurt, le défi qui se posait à l'équipe de développeurs à ce stade était double : non seulement assurer qu'à tout moment l'extraction d'un fichier est possible, mais aussi – ce qui nous intéresse plus particulièrement ici – qu'aucun usager du système n'a accès à des contenus dont il n'est pas le destinataire prévu<sup>318</sup> : en somme, il s'agit pour les développeurs d'obtenir un équilibre entre l'agilité du dispositif et la confidentialité des données qui circulent.

A la démarche qu'on a analysé précédemment – se basant sur les trois aspects d'encryptage, fragmentation et redondance – les développeurs doivent superposer une « couche » qui puisse sécuriser les échanges et les partages entre utilisateurs. La manière de façonner cette couche est pendant longtemps sujet de débat. C'est seulement dans un deuxième temps que – peut-être, note Dietrich, l'idée a pris forme avec le bruit médiatique autour des « serveurs tout-puissants » de Facebook<sup>319</sup> – l'équipe décide de travailler sur un système pouvant permettre des échanges directs aux deux niveaux de l'interface et de l'application.

Le choix, qui a déjà été opéré, de baser le système sur une architecture P2P leur facilite la tâche, puisque la mise en relation directe de deux machines échangeant des paquets de données aurait une contrepartie intuitive dans le geste d'échange ou de partage effectué par les utilisateurs au moyen de l'outil social :

*« The distributed system that made programming more complicated in other areas, for once, could enable us to build a system in which the experience of the user would mirror what was happening below. Whenever a user was publishing a file, the encryption key of that file would be revealed to the public; but similarly, when the user wanted to share with other specific users, the key of these files would be securely revealed only to the friends and groups he shared the file with. Among us [les développeurs], at some point we called it the friendship key, and it was something that stayed afterwards<sup>320</sup>. »*

Pour les opérations de partage, la « clé d'amitié » entre donc en jeu, une autorisation d'échange qui doit être connue seulement de celui qui l'attribue et de celui qui la

---

<sup>318</sup> Entretien avec K. Gueden.

<sup>319</sup> Entretien avec D. Poell.

<sup>320</sup> Id. Je souligne.

reçoit, et qui permet de reconstituer le fichier. L'introduction d'un mécanisme d'« ajout d'amis », propre des réseaux sociaux, et le rapport direct entre le lien d'amitié établi et l'accès aux données de l'ami – mise en relation dont sont éliminés les intermédiaires – doivent renforcer ultérieurement la dimension de confidentialité dans les activités de stockage et partage. Au cours de la Google Tech Talk, Dietrich dépeint la « clé » comme étant le principe de base des échanges dans Wuala, et ne résiste pas à la tentation d'insérer, en conclusion de son discours à prévalence technique, une référence à l'actualité qui ne manque pas d'humour<sup>321</sup> :

*« When they first got friends, Alice and Bob have exchanged the friendship key, now Alice encrypts the file key with the friendship key – this is a simplification of what actually happens, but it serves for this example – and exchanges it with Bob, who proceeds to download the file. No strings attached, and no servers attached, either<sup>322</sup>. »*

Si l'interface de la version alpha de Wuala est développée pour ressembler beaucoup à tout système de fichiers, avec ses dossiers aux couleurs différentes selon le niveau plus ou moins élevé de confidentialité dont ils sont investis<sup>323</sup>, les fonctionnalités « sociales » du dispositif se voient accentuées, lors de l'introduction de la version beta, par son interface utilisateur. Celle-ci est désormais construite de manière à suggérer une continuité avec les réseaux sociaux qui deviennent de plus en plus populaires : le vocabulaire de « profil », « partage », « amis » et « groupes » est repris, en prévoyant que chaque usager puisse attribuer différents degrés de confiance à ses amis (ou à l'ensemble du réseau pour les contenus publics), pour chacun des fichiers qu'il stocke dans Wuala<sup>324</sup>.



<sup>321</sup> L'expression « *No strings attached* », qui signifie « sans conditions (déplaisantes) », veut dire littéralement « sans cordes attachées » : en faisant référence au débat médiatique en cours sur les risques que les serveurs « tout-puissants » de Facebook représentent pour la privacy, Dietrich reprend cette formule pour suggérer que le service Wuala est « sans serveurs attachés ».

<sup>322</sup> Extrait de la Google Tech Talk par Dietrich Poell, transcrite par moi. Comme on verra par la suite, l'affirmation « *no servers attached* » est à son tour une simplification.

<sup>323</sup> Wuala Privacy Policy, <http://www.wuala.com/en/about/privacy>

<sup>324</sup> Entretien avec D. Poell.

*Image 5. Capture d'écran de la partie « réseau social » de l'interface Wuala vue par l'utilisateur, stade de closed alpha sur système d'exploitation Linux. Image par Daniel Bartholomew sur LinuxJournal.com.*

Les notions de réseau social et de clé d'amitié sont donc comprises ici non seulement comme une « couche marketing », ou comme une manière de structurer l'organisation pratique du stockage, mais elles contribuent aussi à changer le dispositif dès le niveau de l'architecture. Elles proposent une définition d'organisation sociale des utilisateurs, qui a trait à la confiance réciproque et à l'ouverture/mise à disposition d'une partie de ses ressources informatiques – ainsi qu'à une reconfiguration de la chaîne des intermédiaires dans l'acte d'échange et de partage.

#### CONFIDENTIALITÉ DES DONNÉES ET ALLOCATION DE RESSOURCES

Le troisième aspect qui marquera la définition de la *privacy* à l'œuvre dans Wuala survient lors de la mise au point de la politique de confidentialité du système. En effet, l'équipe Wuala y introduit, suite au démarrage de la version beta, une section précisant les conditions d'allocation, à telle ou telle autre opération de stockage et partage, des ressources des ordinateurs participants au système.

Comme on a pu voir, dès le lancement de la première version du dispositif, le statut très particulier des ressources utilisées par le système avait rendu nécessaire de préciser dans les conditions d'utilisation un certain nombre d'aspects : le rôle des ordinateurs des usagers, les types de ressources qui pouvaient être utilisées par Wuala, leurs fins, et comment l'utilisateur pouvait établir, et communiquer au système, la quantité maximale de ses ressources locales mises à disposition du réseau-système de stockage<sup>325</sup>. Mais cette première version des conditions d'utilisation n'explicite pas l'articulation entre la mise à disposition de ressources de la part des utilisateurs et la façon dont les ressources sont affectées ou attribuées à des opérations différentes par le système.

Pourtant, comme en viennent à le conclure Dietrich, Kurt et leur équipe, la façon dont s'articulent ces deux aspects a des implications importantes pour la confidentialité des données circulant dans le système (à la fois les informations personnelles des utilisateurs et les contenus stockés). L'« accaparement » et l'allocation des ressources de la part du système Wuala ou de la firme Caleido

---

<sup>325</sup> Voir les conditions d'utilisation de Wuala dans un précédent tableau : « The user acknowledges that Wuala may use processor, bandwidth and harddisk (or other storage media) of his computer for the purpose of storing, encrypting, caching and serving data that has been stored in Wuala by the user or any other users. The user can specify the extent to which local resources are used in the settings of the Wuala client software. » [Extrait des conditions d'utilisation de Wuala, <http://www.wuala.com/en/about/terms>].

impliquent pour les utilisateurs une surveillance possible ou un monitoring de ces ressources, de façon qui peut être plus ou moins automatique, invasive, poussée. Le problème est notamment posé par un utilisateur sur le forum de Wuala en avril 2009<sup>326</sup>, et à la suite de cette discussion, deux modifications sont apportées aux conditions d'utilisation un mois plus tard : si les termes généraux indiquent désormais que

*« Resources are allocated and monitored in accordance with the Privacy Policy<sup>327</sup> »,*

la politique de confidentialité elle-même précise le degré d'automation et d'omniprésence du système d'allocation et monitoring des ressources :

*« In order to ensure a fair allocation of resources within Wuala, various data about the computers participating in the Wuala Network is collected. This data includes their IP addresses, disposability and the amount of resources they are contributing (e.g. bandwidth, memory). (...) Caleido keeps track of how much storage space (e.g. bandwidth, memory) you have used and earned within Wuala. (...) Caleido collects statistical information for the purposes of monitoring, debugging and improving Wuala. This includes automatically generated problem, performance, network analysis and general usage reports, as well as logs of the connections and queries made to Caleido's servers (including the involved IP addresses), as well as analytical data about the usage of the Wuala website. However, none of this data contains information from your private or shared files<sup>328</sup>. »*

Le fonctionnement du système d'allocation comporte donc effectivement le rassemblement de certaines informations concernant les ressources matérielles, de mémoire, de calcul, mises à disposition par les ordinateurs participants. Cependant, ces informations sont traitées automatiquement et de façon agrégée, à des fins statistiques. La récolte d'informations, affirme la version révisée du document, a pour but le calcul automatique de la quantité d'espace de stockage que chaque utilisateur met à disposition (et, comme on le verra dans la prochaine section, de la mesure dans laquelle il peut réclamer du « cloud P2P », c'est à dire un espace de stockage équivalent dans le réseau d'ordinateurs participants).

Les développeurs font aussi appel à l'esprit collaboratif des utilisateurs, en sollicitant des retours sur le forum et sur d'autres plateformes, et en soulignant l'utilité que ces informations auront pour l'amélioration et le développement ultérieur du dispositif. Mais surtout, en reformulant l'enjeu essentiel de ce système

---

<sup>326</sup> Wuala Forum, discussion « Put my mind at ease (privacy concern) ». <https://forum.wuala.com/viewtopic.php?f=34&t=435>

<sup>327</sup> Extrait des conditions d'utilisation de Wuala, consultées à <http://www.wuala.com/en/about/terms>, version de mai 2009.

<sup>328</sup> Wuala Privacy Policy, <http://www.wuala.com/en/about/privacy>

d'allocation – la stabilité et durabilité de Wuala – le document contribue à séparer cet enjeu de la question de la confidentialité des données personnelles et des contenus stockés. Les informations rassemblées sur l'équipement de stockage (c'est à dire, les ressources locales des utilisateurs, que chacun d'entre eux met à disposition) n'impliquent pas une intrusion dans les contenus stockés eux-mêmes. Ceux-ci restent, en principe, sauvegardés par le système mis en place au moyen de l'encryptage et de la fragmentation des contenus ; de la sauvegarde du mot de passe en local ; des partages directs, basés sur l'échange des clés d'amitié.

### SECTION 3. ESPACE DISQUE CONTRE ESPACE EN LIGNE : LA RECHERCHE D'UN MODÈLE SOCIO-ÉCONOMIQUE DURABLE POUR LE STOCKAGE DISTRIBUÉ

Lors de ma première série d'entretiens avec les membres de l'équipe Wuala, deux années après la sortie du *closed alpha* – à l'automne 2009 – mon impression est que ceux-ci demeurent assez perplexes quant aux modèles économiques pouvant s'adapter au système de stockage distribué qu'ils ont développé et qu'ils sont en train de perfectionner. Dietrich envisage soit la possibilité de faire acheter aux utilisateurs/clients de l'espace de stockage dans les centres de données Wuala pendant la phase d'amorçage, soit d'installer un système d'annonces publicitaires, bien que pour le moment – déclare-t-il – l'équipe donne la priorité au développement de la technologie, pendant que Caleido vit sur les revenus de AddressBook et d'autres produits et services « classiques » :

*« At the moment, we're experimenting with ads. However, there are also other possible business models such as photo finishing, buying additional storage (you can trade if you want, if you don't, you could buy some additional online storage for a really low price), or also other forms of sponsoring that we're working on. Anyway, currently we are self-funded and not looking for money on this side. We have put some of our savings into Wuala and I can put the revenue from my former company, which is still running very well, into this new company. I'm afraid we have no brilliant ideas on how to make this a profitable business yet, but this hasn't nearly been our focus, either<sup>329</sup>. »*

De son côté, Kurt souligne que la technologie de *cloud* P2P sous-tendant Wuala est assez complexe à implémenter, et mérite que les efforts de l'équipe s'y concentrent de façon prioritaire. L'objectif est de mettre sur le marché un produit déjà doté d'une certaine stabilité et fiabilité, et en des temps relativement rapides :

*« Our focus was on the technology and the users, rather, on users and the technology. Wuala is a complex technology with many potential applications, and one of the biggest challenges was to focus. This is especially important for something that takes years to complete, while the earth turns around and things change. For us, it was very important to release an early closed alpha version and invite key users who could invite their friends, so that we could listen to feedback and build the application together with our users<sup>330</sup>. »*

Plus d'une année après – début 2011 – je pose les mêmes questions aux développeurs de Wuala. Peut-être est-ce dû à « l'attention à la technologie et aux usagers » dont parle Kurt, mais à côté de mécanismes de revenus supplémentaires comme la vente d'espaces additionnels, le véritable cœur du modèle technique de

---

<sup>329</sup> Entretien avec Dietrich Poell, septembre 2009.

<sup>330</sup> Entretien avec Kurt Gueden, septembre 2009.

distribution à l'œuvre dans Wuala est devenu aussi bien son modèle socio-économique : un utilisateur peut échanger l'espace inutilisé du disque dur de son propre ordinateur contre de l'espace de stockage sur le « nuage d'utilisateurs » de Wuala, pourvu qu'il reste en ligne un certain temps par jour<sup>331</sup>.

Le principe à la base de ce mécanisme de troc entre espace disque et espace en ligne, envisagé par Dietrich et son équipe, n'est pas de donner une plus grande capacité de stockage mais plutôt d'améliorer la qualité de l'espace de stockage dont l'utilisateur est normalement doté sur son disque dur ; il y gagne la possibilité de transformer son disque dur local en disque dur en ligne, potentiellement accessible en mobilité, et cela même si son propre ordinateur est éteint. Bien que ce modèle ne fournisse pas des revenus directs, de son côté, l'entreprise s'assure ainsi une certaine durabilité dans le temps des ressources matérielles et computationnelles à sa disposition, la stabilisation d'un ensemble d'usagers « pro », et la possibilité de s'appuyer sur ce noyau de ressources dont la présence est plus « sûre » pour proposer des solutions de stockage *ad hoc* à des acteurs institutionnels et du secteur privé.

#### AU COMMENCEMENT ÉTAIT L'APPLICATION BUREAU... QUI RENCONTRA LE WEB

Avant le démarrage du *closed alpha*, une discussion s'engage au sein de l'équipe sur la « porte d'entrée » de l'utilisateur dans le système Wuala. Deux alternatives sont considérées : une application bureau (ce qui nécessiterait une installation, de la part de l'utilisateur, d'un logiciel client sur sa machine) et la possibilité de démarrer Wuala directement depuis le web, au moyen d'une application Java<sup>332</sup>, sans que le téléchargement ou l'installation soient nécessaires, afin de faciliter l'accès aux fichiers « en mobilité », sur plusieurs machines, mais conservant la dimension « locale » des opérations d'encryptage et décryptage.

Par ailleurs, Wuala ne sera pas un service Web basé sur un navigateur, car cela comporterait une rupture totale avec la partie « en local » du modèle, avec la conception particulière de sécurité et confidentialité que celle-ci comporte – et donc avec ce qui différencie le système de stockage par rapport à ses concurrents :

« Vous vous demandez peut-être, pourquoi Wuala n'a pas d'application web basée sur un navigateur comme tous les autres services de ce type ? La réponse est simple [...] pour qu'une application web puisse fonctionner, un serveur web devrait envoyer vos données dans un format non crypté vers votre navigateur - certes possible via une connexion HTTPS sécurisée, mais les données seront tout de même complètes - car aujourd'hui, les navigateurs

---

<sup>331</sup> Entretien avec Dietrich Poell, janvier 2011.

<sup>332</sup> Le web start de Java (une plateforme permettant d'exécuter des programmes écrits en langage de programmation Java) permet de stocker rapidement une application démarrée depuis le Web sur le disque local de l'utilisateur. Il s'agit donc, pour simplifier, d'une « voie moyenne » entre un client installé en local et une application Web pure.

ne sont pas capables de faire le chiffrement et le déchiffrement des données en local. Cela signifie que ce serveur web aurait besoin de connaître la clé pour décrypter vos fichiers. En opposition avec d'autres services, Wuala crypte et décrypte tous les fichiers directement sur votre ordinateur. [...] Par conséquent, notre serveur web ne connaît pas la clé de vos fichiers, et ne peut donc pas les décrypter. En bref : toutes les applications par navigateur Web ont accès à vos données. Puisque le mécanisme de sécurité de Wuala ne permet pas l'accès à vos fichiers, vos données ne sont pas accessibles par une application web<sup>333</sup>. »

Bien qu'il soit clair à ses yeux que le client, porte d'entrée dans le réseau P2P, est fondamental d'un point de vue à la fois technique et « symbolique », Dietrich se montre initialement assez perplexe quant à la possibilité que Wuala soit *exclusivement* une application bureau, car l'intégration à d'autres services serait moins immédiate ; par conséquent, l'utilisateur devrait investir plus de temps et de ressources dans son « premier essai » de Wuala, et risque peut-être de se désister à l'idée de devoir installer le client en local<sup>334</sup>.

*« With the P2P idea behind it, Wuala was clearly born as a desktop application (...) simple, well integrated into the operating system, network drive, drag & drop, upload in the background... There are a lot of advantages for such an application if it is on the desktop. However, soon we started working on a in-browser and web integration, so that people would be able to access data also through the web. Also, we think it is important in the launch phase that people can easily link and relate to Wuala and make it part of their everyday habits, so there may also be a Facebook integration<sup>335</sup>. »*

La discussion qui s'engage parmi les utilisateurs lors de la sortie du *closed alpha*, qui ne comporte pour le moment que l'application bureau, ne contribue pas à établir une préférence claire parmi les utilisateurs, dont les développeurs pourraient tenir compte dans les étapes suivantes. Selon les priorités des différents utilisateurs, soit le téléchargement du client, ou même de l'application Java pour le démarrage direct depuis le Web, est vu comme un poids excessif par rapport à l'agilité que pourrait offrir une interface web ; soit il est considéré comme une composante « incontournable » de la promesse de sécurité offerte par l'encryptage des données en local, l'aspect essentiel pour la différenciation du produit par rapport à ses nombreux concurrents :

*« i don't like how you have to download a desktop application or a java app in order to view the files that you have on wuala. i would like it much better if there was a web interface... the app is almost 7.5MB! that's crazy! »*

---

<sup>333</sup> Page « Web Access » du site Wuala à <http://www.wuala.com/fr/launch/>

<sup>334</sup> On se rappellera la controverse autour du téléchargement du client P2P dans le cas de Faroo.

<sup>335</sup> Entretien avec D. Poell, septembre 2009. L'intégration Facebook ne sera pas implémentée par la suite (au moins jusqu'au moment où on écrit, trois ans plus tard).

*« You worry about your files being safe and you want a WEB interface? NO! I want something LOCAL that I KNOW is encrypting my files BEFORE it leaves my computer. Wuala does just that. If any others do that, forgive my ignorance, but I can actually see the files (fragments) that Wuala makes... 7.5 MB is seriously surprises you? If you try to restrict the size of your program, you're not going to be able to add all the features you need to stand out from the competition, and thus make your product ROCK! Which Wuala does quite nicely. I have my hands in Wuala, IDrive, Dropbox, Syncplicity, Adrive, MediaFire, MyOtherDrive, Skydrive, XDrive (defunct), and Mybloop (defunct). And I have to say that Wuala tops them all, followed closely by Syncplicity. »<sup>336</sup>*

C'est finalement<sup>337</sup> l'importance de l'encryptage en local, réaffirmée par plusieurs utilisateurs, qui amène Dietrich et l'équipe Wuala à pencher vers une solution à deux faces, qui fait co-exister le client et le démarrage depuis le Web au moyen de Java – et à rejeter cependant une solution « purement Web ». D'un point de vue technique, l'application bureau réaffirme les aspects « locaux » de la démarche de stockage et partage ; du point de vue du positionnement économique, elle veut réaffirmer aux yeux des usagers l'avantage concurrentiel du dispositif et l'« originalité » du dispositif.

Le choix d'avoir une double « porte d'entrée » au système, mêlant client et démarrage Web, détermine par ailleurs, pour la première fois, un déplacement qui influencera de manière significative, par la suite, l'évolution du dispositif. De « stockage social en P2P », le service sera désormais appelé « cloud P2P » : les aspects de réseautage social co-existent maintenant avec une idée de plateforme, un système plus souple, accessible en mobilité – ainsi que plus évanescent et moins clair à appréhender pour l'utilisateur<sup>338</sup>. Le principe du « faire le pont entre le bureau et le Web » [*bridging desktop and web*] devient la pierre angulaire du modèle économique de Wuala : le troc d'espace sur le disque dur de l'utilisateur contre de l'espace de stockage en ligne.

## TROC ET « QUALITÉ DE SERVICE »

Comme on l'a déjà évoqué<sup>339</sup>, l'option de stockage partagé n'est pas proposée, au

---

<sup>336</sup> Discussion sur le site MakeUseOf.com : <http://www.makeuseof.com/tag/what-do-you-get-if-you-combine-dropbox-and-mozy-wuala/>

<sup>337</sup> A l'été 2010.

<sup>338</sup> Ce qui, comme on verra dans la dernière section, sera source de nombreuses controverses quant au statut des utilisateurs et de leurs machines dans le dispositif.

<sup>339</sup> Section 2 de ce chapitre.

début, en tant que choix obligé pour l'utilisateur : au cours de ses premiers mois d'existence, Wuala est en effet programmé pour donner la possibilité aux utilisateurs ayant besoin de moins d'un gigaoctet de mémoire pour leur stockage de ne pas l'obtenir du réseau d'ordinateurs connectés, mais des serveurs de la firme (« *You start with 1 GB provided by us*<sup>340</sup> »), comme offre de base gratuite. Cette option fait partie des mesures de « recrutement » et de stabilisation d'une base d'utilisateurs, que l'équipe estime opportun de mettre en œuvre pendant la phase d'amorçage<sup>341</sup>, mais ne rencontre qu'un intérêt très mitigé de la part des usagers de la première heure. Ce qui, comme le souligne Dietrich, était peut-être à prévoir – les utilisateurs pionniers de Wuala s'intéressent spécifiquement au système de stockage partagé, et forment un groupe plus concerné que l'utilisateur lambda par les aspects techniques de l'infrastructure P2P qui le rendent possible :

*« people who tried Wuala were interested in the P2P storage mechanism, and that involved the storage trading*<sup>342</sup>. »

Au delà de l'offre de base, une alternative est proposée à l'utilisateur pour étendre cette capacité. Il peut soit acheter des gigaoctets de mémoire supplémentaires, qui lui seront attribués sur les serveurs Wuala, soit mettre à la disposition de Wuala une partie de son disque dur : il pourra ainsi troquer de l'espace inutilisé du disque de son ordinateur contre un stockage équivalent sur la plateforme d'utilisateurs Wuala.

La façon de quantifier les ressources dont chaque utilisateur peut disposer dans le système Wuala est établie de la manière suivante : elle dépend de la quantité d'espace que chaque utilisateur fournit en local, multiplié par le pourcentage de temps qu'il passe en ligne :

*« The amount of resources the user is allowed to use in Wuala depends on the amount of local resources the user is contributing to Wuala*<sup>343</sup>. »

Si un utilisateur met à disposition 100 Go d'espace de son ordinateur, il pourra obtenir jusqu'à 100 Go d'espace en ligne en échange, qui variera selon le temps qu'il passe en ligne : s'il reste connecté 70% du temps, par exemple, il obtiendra 70 Go d'espace en ligne. Comme souligne Dietrich, l'idée à la base de ce mécanisme n'est pas de fournir davantage d'espace de stockage à chaque utilisateur, mais d'en améliorer la « qualité » ; l'espace de stockage en local est converti en espace de

---

<sup>340</sup> Site Wuala, page « Trading Storage », <http://www.wuala.com/fr/storage/trade/> (vue pour la dernière fois en août 2011, elle redirige actuellement à la page « Technologie » du site Wuala).

<sup>341</sup> Dans cette même perspective de recrutement et d'intéressement, l'offre de base au stade de *closed alpha* peut aussi être augmentée sans souscrire à l'option de stockage partagé, en invitant des amis. Les conditions d'utilisation expliquent : « *Caleido provides the user with a basic amount of online storage (e.g. 1 GB) for free. Successfully inviting new users or trading local resources can increase the user's available storage. Caleido can limit the amount of storage that can be obtained by inviting new users. Currently available storage and storage usage are displayed in the Wuala client.* »

<sup>342</sup> Entretien avec D. Poell, janvier 2011.

<sup>343</sup> Extrait des conditions d'utilisation de Wuala, consultées à <http://www.wuala.com/en/about/terms>, version de mai 2009.

stockage en ligne, d'accessibilité plus facile surtout dans des situations en mobilité :

*« The idea behind storage trading came up because we aimed at improving the quality of users' storage: it was not sustainable to propose a deal that increased the available space, but it became possible to build on local space to offer a global and mobile solution, accessible when the user's very own computer was shut off. The value of the whole trading mechanism became about the different quality of the storage it was able to provide, rather than the quantity<sup>344</sup>. »*

Une condition est également posée aux utilisateurs, pour qu'ils puissent faire partie du programme de troc : ils doivent être en ligne au moins quatre heures par jour (Image 6). Encore une fois, il s'agit d'un choix à la fois de stabilisation technique et d'optimisation économique, qui vise à augmenter la stabilité et la persistance d'un « noyau dur » d'utilisateurs prêts à contribuer un minimum de ressources, de manière constante dans le temps, au système distribué :

*« The reason for this is that it just doesn't make sense to store fragments in a computer that is online just, say, half an hour a day<sup>345</sup>. »*

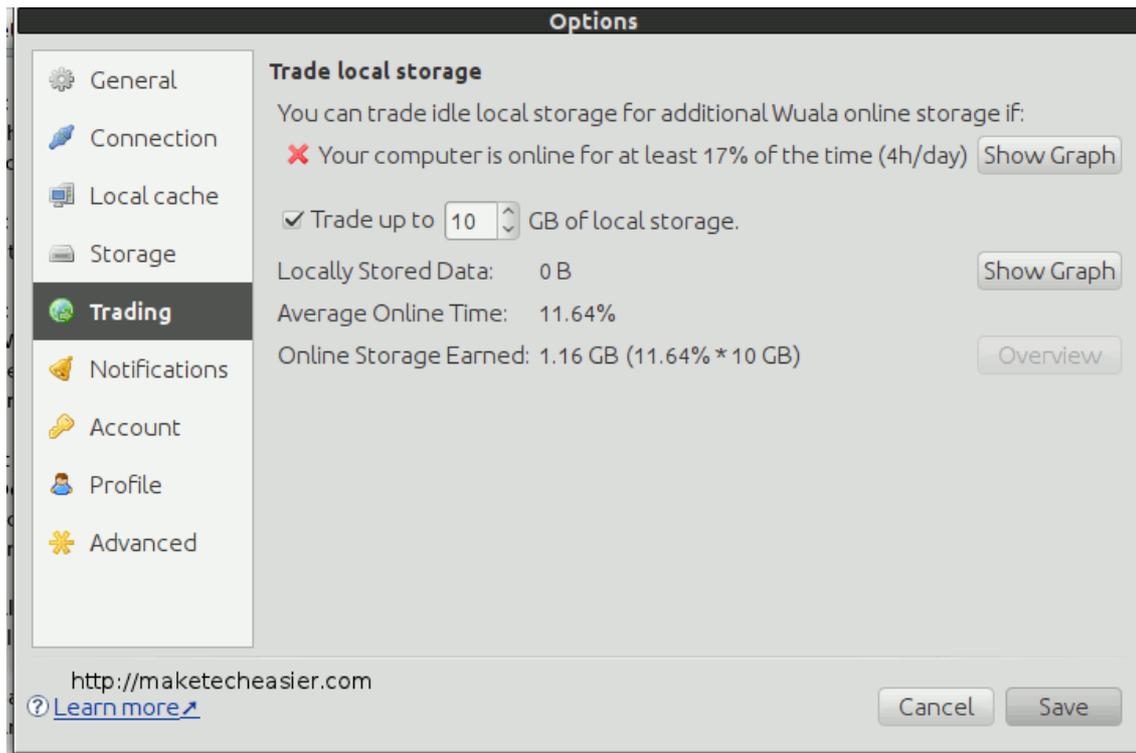


Image 6. Capture d'écran du programme de réglage des paramètres pour le

<sup>344</sup> Entretien avec D. Poell, janvier 2011.

<sup>345</sup> Entretien avec K. Gueden, février 2011.

*stockage partagé dans Wuala, vu par l'utilisateur. Image par Damien Oh sur maketecheasier.com.*

La mise en place de ce mécanisme implique également, par la suite, l'implémentation d'une fonctionnalité qui tient compte des fluctuations, au fil du temps, dans la durée moyenne des connexions des utilisateurs. Dès que la condition de rester un minimum de quatre heures par jour en ligne est satisfaite, plusieurs d'entre eux se posent en effet la question de ce qui se passe s'ils diminuent leur temps de connexion – et donc, de ce qui arrivera à leurs données si le stockage qu'ils ont acquis décroît ensuite<sup>346</sup>.

Un système de « rééquilibrage » est donc mis en place, selon lequel la quantité de stockage en ligne diminue si le temps de connexion moyen d'un utilisateur décroît, ou si l'utilisateur baisse la limite de la quantité d'espace que son client Wuala a la permission de troquer. Si ce faisant, l'utilisateur arrive à avoir moins de stockage en ligne que la quantité de données qu'il a stockées sur Wuala, il pourra toujours accéder à ses données mais ne pourra pas stocker des fichiers additionnels jusqu'à ce que son temps de connexion augmente à nouveau ou que, informé du changement de situation par Wuala, il n'en supprime une partie lui-même<sup>347</sup>.

L'adhésion au mécanisme de stockage partagé et le fait de rester un certain temps en ligne représentent les deux critères qui permettent de qualifier un utilisateur comme « Pro » au sein du système. La définition d'utilisateur « Pro » reste dynamique au cours des différentes versions du dispositif, mais doit permettre, par exemple, de désélectionner les mécanismes de revenus « parallèles » de Wuala, tels que les annonces publicitaires :

*« By trading local resources (...) the user gains the so called 'Pro' status. The 'Pro' status gives the user access to additional features (e.g. being able to disable the advertisements in the settings of the Wuala client software). Caleido can remove, extend or otherwise change the features associated with the 'Pro' status at any time. The features essential to the use of Wuala will always be available to all users at no cost, whether they have the 'Pro' status or not<sup>348</sup>. »*

Un deuxième critère n'est pas, au début, explicitement noté dans les conditions d'utilisation, mais il devient au fil du temps tout aussi important pour le bon fonctionnement technique et la durabilité du système : le type de connexion Internet des utilisateurs. En effet, plusieurs d'entre eux se voient refuser par leur client

---

<sup>346</sup> Discussion du mécanisme « Trading Storage » sur le site Wuala, <http://www.wuala.com/fr/support/faq/c/16> (cette page n'est plus en ligne, vue pour la dernière fois en juillet 2011).

<sup>347</sup> Entretien avec K. Gueden, février 2011.

<sup>348</sup> Extrait des conditions d'utilisation de Wuala, consultées à <http://www.wuala.com/en/about/terms>, version de mai 2009.

Wuala l'adhésion au programme de troc parce que leur « connexion Internet ne permet pas de troquer du stockage<sup>349</sup> ». Le type de connexion Internet se qualifie donc comme un enjeu de première importance pour que l'utilisateur puisse pleinement bénéficier du mécanisme de troc, car il peut même arriver à constituer une véritable barrière à l'entrée du système, au cas où le client Wuala se heurte à un pare-feu particulièrement restrictif, qui n'accepte pas les connexions entrantes sur le port utilisé (et affiché) par le logiciel<sup>350</sup>.

Les nombreux problèmes rencontrés par les utilisateurs à ce sujet amènent finalement l'équipe Wuala à définir plus précisément la « qualité de service » que l'utilisateur obtient avec le mécanisme de troc. Des données relatives à la connexion Internet font désormais part de cette définition : l'accessibilité de l'ordinateur d'un utilisateur par les autres ordinateurs qui composent le système ; la qualité de la connexion de cet ordinateur à l'Internet ; et la capacité de celle-ci à permettre la circulation de certains types de trafic (Tableau 7).

Vous pouvez recevoir des données uniquement quand votre ordinateur est accessible, c'est à dire uniquement si d'autres ordinateurs peuvent se connecter à vous. [...]

4. Connexion non appropriée pour le troc : Wuala ne peut envoyer et recevoir de paquets UDP<sup>351</sup>. Il y a de fortes chances que vous soyez derrière un pare-feu restrictif qui bloque le trafic UDP.

Vous pouvez atteindre l'accessibilité complète quand la redirection de port est activée sur votre routeur. Vous devez configurer votre routeur de façon à rediriger les paquets UDP qui arrivent au port Wuala vers l'ordinateur sur lequel Wuala est installé.

*Tableau 7. Le type de connexion Internet comme barrière à l'entrée du mécanisme de troc [Section « Trading Storage » du site Wuala].*

## DERRIÈRE LE NŒUD, L'USAGER : MOTIVATION ET RESPECT DES RESSOURCES

La stabilisation du mécanisme de troc en tant que modèle économique durable pour le stockage distribué passe également par la mise en place d'un système d'incitation des usagers au partage des ressources. Les préoccupations des développeurs de Wuala rejoignent ici un des enjeux classiques des réseaux distribués : créer un système qui, au moyen de mécanismes de réputation, de motivation (*incentives*) et d'attribution répartie de responsabilité, assure une gestion de ressources équitable entre les pairs, à la fois celles que chaque utilisateur peut réclamer du système et

<sup>349</sup> Discussion du mécanisme « Trading Storage » sur le site Wuala.

<sup>350</sup> Id.

<sup>351</sup> Le *User Datagram Protocol* (UDP) est l'un des protocoles de télécommunication les plus répandus de l'Internet ; sa fonction est de permettre la transmission de données de manière très simple entre deux entités, chacune définie par une adresse IP et un numéro de port.

celles qu'il apporte à son fonctionnement (Oram, 2001 ; Taylor & Harrison, 2009). Dans le cas de Wuala, les développeurs doivent prendre en considération pour la mise au point du mécanisme trois types de ressources qui sont directement ou indirectement « partagées » – il s'agit de l'espace disque local, du temps passé en ligne, et de la bande passante :

*« Speaking about incentives and fairness is important when you deal with any system in which users have to contribute their resources. What you want to do is to prevent free-riding, that is to say you don't want to have nodes that are consuming data but not contributing much. Now if we look at the resources in our system, we have local disk space, then we have online time and also upload bandwidth<sup>352</sup>. »*

L'équipe de développeurs décide de travailler sur les deux premiers types de ressources comme s'ils ne constituaient qu'un seul et même aspect, puisque dans l'économie du dispositif, un ordinateur qui reste en ligne pendant de longues sessions – voire en permanence – est particulièrement précieux puisqu'il contribue le plus à la décroissance du facteur de redondance nécessaire au fonctionnement du système. Le principe d'adhésion au mécanisme de troc, selon lequel l'utilisateur gagne autant de stockage qu'il en met à disposition sur son disque multiplié par le temps qu'il passe en ligne, est donc aussi le résultat de la volonté des développeurs de motiver les utilisateurs à rester en ligne de façon la plus constante possible, et pour des longues sessions<sup>353</sup>. Bien que ce principe ne reflète pas exactement la réalité technique plus complexe que sous-tend le mécanisme, les développeurs de Wuala souhaitent que sa formulation concise et immédiate « accroche » les utilisateurs à l'esprit du système – tandis que de leur côté, ils se confrontent à cette même réalité technique complexe pour essayer de quantifier et mesurer l'apport de chaque nœud, et le mobiliser pour le bien du système dans son ensemble :

*« Thus we came up with that very simple formula (...) In reality this is more complex, because the necessary redundancy doesn't exclusively nor necessarily depend on the time you are online, but we wanted to have a simple formula the users could connect to and understand the system. Actually (the) tricky part here was not the formula, but for us how to measure those variables. There are a number of different mechanisms that can check and double-check if the reports of every pair are correct, so that you can find out if a node in the system still stores that particular fragment. We are still experimenting with them, but making progress<sup>354</sup>. »*

---

<sup>352</sup> Entretien avec K. Gueden, février 2011.

<sup>353</sup> Une incitation supplémentaire, qui sera introduite dans la version beta, est la possibilité de gagner du stockage à partir de plusieurs terminaux qui seraient en ligne à des moments et pour des durées différentes : « *So if you have a laptop that is online only a few hours a day, but you also have a desktop computer at home which runs a lot more, the online time is added to your user account.* » (Extrait de l'entretien de janvier 2011 avec D. Poell).

<sup>354</sup> Entretien avec K. Gueden, février 2011.

Il est également important pour la durabilité du dispositif que les utilisateurs allouent une quantité suffisante de bande passante aux téléchargements exécutés au moyen du client Wuala. Les développeurs choisissent de récompenser la mise à disposition de bande passante de la part de chaque utilisateur en lui attribuant une plus grande rapidité de téléchargement de contenus à partir du cloud P2P. Du point de vue de l'utilisateur, la motivation à la base de l'allocation de bande passante au service se veut donc assez clairement identifiable et compréhensible :

*« the more upload bandwidth you provide, the faster your download speed will be<sup>355</sup>. »*

La partie la plus problématique pour l'équipe se révèle être, encore une fois, l'implémentation technique de ce principe d'allocation de ressources – ou plus précisément, du mécanisme qui devra monitorer et éviter les « passagers clandestins » ou *free-riders* (Adar & Huberman, 2000 ; Shirky, 2001), et faire que le système Wuala ne soit pas formé de nœuds minimisant chacun de leur côté leur contribution de bande passante au système, le transformant en un réseau de « presque îles ». Le choix de la réciprocité, « un prêté pour un rendu » (« *tit-for-tat* ») est laissé de côté au vu de l'asymétrie des intérêts des utilisateurs/nœuds : c'est à dire, il n'arrive presque jamais que deux nœuds réclament une ressource l'un de l'autre à un même moment. Par ailleurs, l'idée de faire surveiller les transactions par les logiciels clients des utilisateurs eux-mêmes est aussi éliminée, au vu de la relative facilité avec laquelle un utilisateur averti pourrait en truquer les véritables performances :

*« Now again, the tricky part here is how you achieve that, because if you look at any typical transaction, then you will see there is an asymmetric interest. Let's say that a client node wants something from a storage node, but the storage node does not want anything from that client at that time. Because of this asymmetric interest, you just cannot use tit-for-tat. This has been very efficient in BitTorrent, but for this system we cannot use it. The other idea would just be that the system believes the software client, so the software client would track the different contribution of users and then report it to another one whenever there is a transaction. But of course this is a bad idea because the software client could easily be hacked in this circumstance. An example of this has been Kazaa<sup>356</sup>. »*

La solution qui est finalement adoptée par Wuala est un mécanisme de motivation et de réputation qui tire profit de l'aspect distribué du système pour limiter les effets néfastes de fausses informations éventuellement envoyées par les clients. Il est aussi important pour les développeurs que le système soit capable de passer à l'échelle avec une grande quantité de petites transactions.

---

<sup>355</sup> Entretien avec K. Gueden, septembre 2009.

<sup>356</sup> Id.

La pierre angulaire de ce système de réputation devient donc la capacité de chacun des nœuds du réseau à exécuter de manière automatique un grand nombre de transactions et observations, de garder trace de la bande passante qu'ils ont reçue pendant ces transactions, et de partager cette information avec un nombre limité et prédéterminé de nœuds voisins, plutôt qu'avec l'ensemble du système. Cela limite l'influence des informations qui ne seraient pas fiables :

*if you are an attacker, you cannot influence the whole network, but only your immediate neighbour. Cheating can be only within a very small band<sup>357</sup>,*

et permet une mise à jour progressive de la « réputation » des différents nœuds de stockage – ce qui se traduit, à son tour, par une allocation de bande passante en proportion à la réputation.

L'importance que le mécanisme de troc assume dans le dispositif rend ses éventuelles défaillances très importantes aux yeux des utilisateurs. C'est la petite équipe de soutien technique de Wuala qui se trouve, dans ces cas, particulièrement sollicitée. Son responsable, Hans, en arrive à définir son propre travail, pendant la période très délicate entre le *closed alpha* et le *public beta* (juillet 2007 – octobre 2008), comme consistant surtout en

*« public relations with people showing up and telling us about problems<sup>358</sup> »*

au sujet du mécanisme de répartition des ressources. Qu'il s'agisse de problèmes dans le calcul de l'espace en ligne attribué à la suite du troc, ou dans le déroulement des opérations de téléchargement des fragments dans le réseau P2P, ces questions démontrent au jour le jour la complexité du système et l'engagement de la communauté dans son progrès :

*« There has been for example a problem in space calculation. After deleting files, this guy was not getting back the storage space he was expecting. There is a way for Wuala to recalculate the space, but the query has to be done on a regular basis as there is a known bug in the quota calculation, that sometimes prevents that the space is given free again. (...) This other guy posts a message in which he says that he has now a lot of gigabytes of storage as he is online a lot and trades a lot, and (...) the upload seems to work, I mean it was showing the upload speed and the green status of the download, but it was always displaying 0% as percentage of upload, the passing of hours was not changing anything about that. First we thought this was a displaying problem, as the upload appeared active. But it wasn't that way and in this case, you do not really know, as the one who has to fix the problem, what are even the information that can be useful to you and the ones you can just skip. In the end the guy wrote back after a lot of back and*

---

<sup>357</sup> Entretien avec D. Poell, janvier 2011.

<sup>358</sup> Entretien avec Hans, responsable de l'équipe de support technique de Wuala en 2009 et 2010, 20 octobre 2009.

*forth and said that after three days, the upload rate was finally showing and was in progress<sup>359</sup>. »*

Au delà de leurs effets sur la performance technique de l'outil, les problèmes et défaillances du mécanisme de troc sont parfois associés par certains utilisateurs à une idée de « manque de respect », de la part d'autres utilisateurs ou des développeurs eux-mêmes, pour les ressources qu'ils mettent en commun. C'est ce qui pousse l'équipe à inclure dans le site Web de la firme, à l'intention de l'utilisateur, des indications précises quant à la quantité des ressources à mobiliser dans le système (Image 7). Qu'il s'agisse de mémoire sur disque ou de bande passante, ces indications sont destinées à obtenir un compromis entre le bon fonctionnement du dispositif et une charge du système Wuala sur la machine de l'utilisateur qui ne soit pas trop importante – c'est à dire, un client P2P qui ne soit pas trop « lourd » pour le terminal :

*« Comment puis-je limiter la quantité de mémoire que Wuala utilise ?*

*Sous Linux, vous pouvez mettre une limite de mémoire en adaptant le script de démarrage. (...) Sous Windows, vous pouvez mettre la limite de mémoire en éditant le fichier Wuala.vmoptions, qui peut être trouvé dans les directives d'installation. Plus Wuala dispose de mémoire, moins il est probable qu'il se bloque en téléchargeant des fichiers lourds ou en dimensionnant de grandes images.*

*Quelle quantité de bande passante Wuala consomme-t-il ?*

*Vous pouvez configurer la quantité maximum de bande passante que Wuala utilisera dans les options sous « Connexion » (...) Pour une meilleure performance, nous recommandons de configurer la bande passante d'émission et de réception à environ 10% en dessous de la bande passante de votre connexion internet. Ceci permet un haut débit sans entraîner un long temps de latence<sup>360</sup>.* »

---

<sup>359</sup> Id.

<sup>360</sup> FAQ de Wuala, « Questions Techniques » (<http://www.wuala.com/fr/support/faq/c/19>). Je souligne.

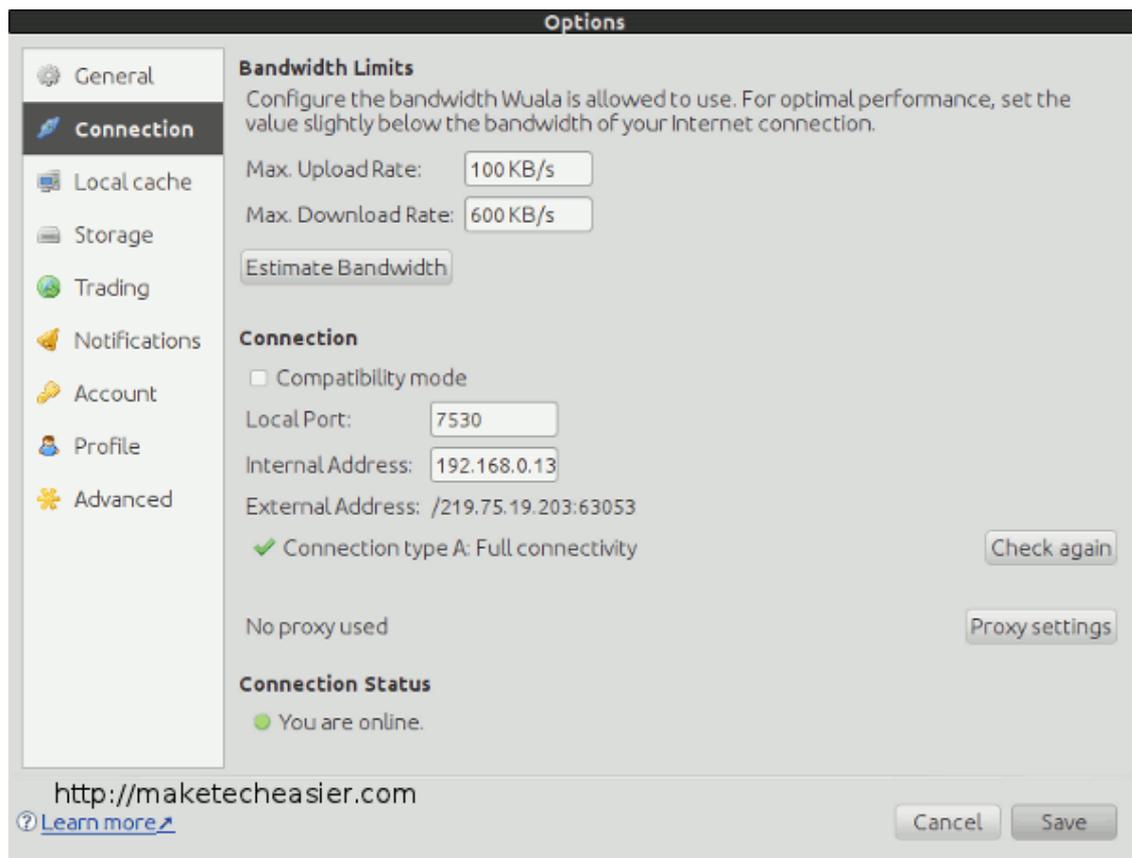


Image 7. Capture d'écran du réglage de la mise en commun de ressources pour le stockage partagé dans Wuala, vu par l'utilisateur. Image par Damien Oh sur maketecheasier.com.

#### « OUT THERE » OU « IN THERE » ? LE NUAGE D'USAGERS ET LA CULTURE DU PARTAGE

Si l'entretien du système dans ses premières et délicates phases de développement se présente depuis le début comme un aspect à ne pas négliger, l'enjeu principal pour le succès du mécanisme de troc semble pourtant être la manière dont les utilisateurs s'en saisiront et se l'approprieront – comment ils réagiront au fait de stocker ses données dans un environnement qui est en même temps local et global, concentré et dispersé. Le « nuage d'utilisateurs » sera-t-il perçu comme une inconnue encore plus grande qu'une plateforme d'entreprise, ou comme une attribution de confiance à sa communauté ? Comment les utilisateurs se saisiront-ils du système distribué ? Le principe de partage sera-t-il reçu majoritairement comme une opportunité ou comme une complication ou incertitude additionnelle ? Telles sont les questions qui inquiètent quelque peu l'équipe Wuala au fur et à mesure que les avis sur leur produit se multiplient à travers le Net, et que celui-ci est examiné, comparé, testé. Dietrich et son équipe estiment que le succès du dispositif dans son ensemble, ainsi que la survie même de l'entreprise à moyen terme, se joueront essentiellement autour de la question du troc. Il remarque :

*« not only was it 'our own thing', it was also the thing that was more likely to raise a lot of questions<sup>361</sup>. »*

Les retours de la part des utilisateurs de la première heure semblent effectivement se concentrer sur la manière dont la distribution du modèle au niveau technique a des effets sur le niveau d'implication, participation et « collaboration » des usagers au dispositif, et reconfigure ainsi leur statut de clients<sup>362</sup>. Certains apprécient la possibilité de troquer l'espace disque pour l'espace en ligne, et souscrivent très volontiers à l'option qui constitue à leurs yeux la vraie nouveauté du dispositif, capable de déterminer un avantage économique :

*« You then choose the files you want to backup to the Wuala service and it copies them out to the cloud. Sounds pretty normal right? Here is where Wuala stands out from the rest. You can choose to share several gigabytes of your local hard drive to the cloud and earn more storage for yourself<sup>363</sup>. »*

*« If you need a lot of online storage space, and at the same time, you have plenty of free space in your local hard drive, you can trade your local space in exchange for online space<sup>364</sup>. »*

*« Overall, for those of us with home file servers that have unused space, this might be a great way to get cloud storage for fairly cheap<sup>365</sup>. »*

D'autres considèrent la situation en termes de coûts et de bénéfices non seulement pour les utilisateurs, mais pour l'entreprise elle-même, et notent que des questions techniques telles l'allocation de bande passante et d'espace disque dans le système distribué se traduisent en modifications très « visibles » pour l'expérience utilisateur :

*« [T]he technology behind Wuala is truly amazing. By using a 'grid' algorithm, Wuala can take advantage of unused disk space across its network of users in addition to data-centre storage. This keeps Wuala's costs down, because there is no need to buy space in data-centres and it also makes their service incredibly attractive to newcomers [...] Another benefit of the Wuala P2P 'grid' is the virtually unlimited bandwidth for both upload and download – basically the same benefits bit-torrent downloads have over traditional CDNs, with up to 100 peers in parallel for each file. I've noticed a sizeable increase in speed over Microsoft's LiveMesh, Dropbox and Mozy.*

---

<sup>361</sup> Entretien avec D. Poell, janvier 2011.

<sup>362</sup> Entretien avec K. Gueden, février 2011.

<sup>363</sup> Homeserver (2010). « New Online Storage Option with Wuala by StockFast ». *Homeservershow.com*.

<sup>364</sup> Damien Oh (2011). « Wuala : An Alternative to Dropbox with Security in Mind ». *MakeTechEasier*, <http://maketecheasier.com/wuala-an-alternative-to-dropbox-with-security-in-mind/2011/08/15>

<sup>365</sup> Matt Smollinger (2011). « Wuala by StockFast Reviewed ». *SmallCloudBuilder*.

*I've had sustained upload speed of over 900Kb/s, the highest I've personally seen in an online file storage service<sup>366</sup>. »*

D'autres encore relient les caractéristiques techniques de Wuala avec les tendances courantes du marché de l'industrie informatique, à la fois en termes des contraintes posées par les fournisseurs d'accès Internet sur la quantité de bande passante octroyée à chaque internaute, et des coûts décroissants du matériel informatique nécessaire au stockage présent sur les terminaux domestiques :

*« Trading also gets interesting as ISPs continue to try to impose harsher bandwidth caps. This makes it very important that you properly manage Wuala's bandwidth limiters, especially in Canada where many providers have been forced by Bell to drop their bandwidth caps to 25 GB per month<sup>367</sup>. »*

*« Given how cheap hard disks have become, this seems like a fair trade-off. Users who share their local hard disk space can also turn off advertising on Wuala<sup>368</sup>. »*

Enfin, certains commentaires semblent indiquer comme raison principale qui les incite à choisir le mécanisme de troc la dynamique de partage et de réseautage préconisée plus ou moins implicitement par le modèle. Le système Wuala permet de mettre à profit des ressources qui ne seraient en tout cas pas utilisées activement, et pour un utilisateur, cela représente une valeur intrinsèque :

*« I think the trading/p2p feature of wuala is the greatest feature. I have a nix box that is on all the time. For me, I get warm and fuzzy all over knowing that I am sharing my extra 50GB of hard drive space with the rest of the world. My computer is running anyway and I have the harddrive space. When you add a QOS rule to put wuala on a lower priority, you never even see a change in network speed<sup>369</sup>. »*

Avec l'augmentation d'essais et de tests effectués par des utilisateurs avertis, le caractère controversé des choix de distribution et de partage faits par Wuala commence par ailleurs à prendre forme dans nombre d'autres interventions. Les doutes sur la faisabilité, la durabilité et la "prise en main" du modèle Wuala se concentrent, suivant les mots des utilisateurs, autour de deux grandes questions.

---

<sup>366</sup> Stefan Neagu (2010). What do you get if you combine Dropbox and Mozy ? Wuala ! », MakeUseOf.com, <http://www.makeuseof.com/tag/what-do-you-get-if-you-combine-dropbox-and-mozy-wuala/>

<sup>367</sup> Smollinger, cit.

<sup>368</sup> Frederic Lardinois (2008). « Wuala Launches P2P Cloud Storage Solution ». *ReadWriteWeb*, [http://www.readriteweb.com/archives/wuala\\_launches\\_p2p\\_cloud\\_stora.php](http://www.readriteweb.com/archives/wuala_launches_p2p_cloud_stora.php)

<sup>369</sup> Discussion de l'article de Stefan Neagu sur MakeUseOf.com, <http://www.makeuseof.com/tag/what-do-you-get-if-you-combine-dropbox-and-mozy-wuala/>

La première concerne la stabilité socio-technique du dispositif : la mesure dans laquelle un nombre suffisant d'utilisateurs apportera suffisamment de ressources, de façon suffisamment constante, pour que les contenus stockés sur le système soient présents de façon durable et fiable.

*« I gather that more than 100 million files are reported to be stored on the Wuala grid ? So there's no need to worry that the service might go down in the foreseeable future ?<sup>370</sup> »*

*« So wait, my ability to access my files online depends on other users keeping them on their hard drives and staying connected to the net? Yeah, sign me up...<sup>371</sup> »*

Un utilisateur s'inquiète aussi de ce qui peut arriver aux données stockées au cas où leur propriétaire change radicalement la quantité de temps qu'il passe en ligne, jusqu'à descendre de façon régulière sous les quatre heures de connexion par jour qui constituent la condition d'entrée dans le mécanisme de troc :

*« Since you have to be online a certain amount of time, there are ramifications if you go offline. Say you go on vacation and shut down your computer(s) for a few weeks. You will probably lose most, if not all of your traded storage<sup>372</sup>. »*

La deuxième question, particulièrement soulignée par certains usagers, a trait aux frontières « nuageuses », au sens propre et figuré, de l'accès aux données personnelles ou stockées par d'autres utilisateurs. D'après certains, les répercussions de la distribution et la répartition des données et des ressources, inscrites dans l'outil, sur la confidentialité des contenus et la sécurité des équipements, sont peu claires :

*« Linux client and meant to be more secure because encryption before sending off, but if files are being stored on other peoples computers as opposed to on there servers what stops ppl using brute force to hack there [sic] way in?<sup>373</sup> »*

*« Even with the strong encryption Wuala uses, though, the fact that some information is going to be stored on machines outside of even Wuala's control is not going to sit well with a lot of people. While most private users can probably live with these risks (which are inherent in any cloud storage*

---

<sup>370</sup> Id.

<sup>371</sup> Discussion de l'article de David Chartier (2008). « First look : Wuala online P2P backup service makes us nervous », forum de *ArsTechnica*, <http://arstechnica.com/civis/viewtopic.php?f=2&t=7507&p=113019#p113019>

<sup>372</sup> Discussion de l'article de Matt Smollinger sur SmallCloudBuilder.

<sup>373</sup> Discussion de l'article de Damien Oh sur MakeTechEasier, <http://maketecheasier.com/wuala-an-alternative-to-dropbox-with-security-in-mind/2011/08/15>

*solution), businesses will probably stay very far away from Wuala<sup>374</sup>. »*

Certains commentateurs expriment à leur tour leur perplexité, non sur le modèle technique et économique de Wuala lui-même, mais sur la façon dont il pourra être perçu par l'utilisateur lambda qui ne verrait pas un intérêt intrinsèque aux détails des explications fournies par les développeurs. David Chartier de *Ars Technica*, un des magazines Web les plus réputés pour l'analyse des tendances de l'Internet, se montre sceptique sur le fait que la « culture de partage », dont Wuala est selon lui un exemple, est suffisante pour qu'un nombre important d'utilisateurs se laissent convaincre par le modèle d'éclatement des données et des ressources proposé par le service :

*« On the very significant flip side of that coin, however, are the security concerns Wuala evokes by not only storing (portions of) files on offsite servers, but on random users' computers across the world. Casual users may not care if their hard-won collection of Hilary Duff pictures falls into the wrong hands, but SOHO businesses and anyone larger can't be blamed for not touching Wuala with a ten-foot pole (...) we're skeptical as to how far Wuala will get. It's true that the onslaught of social media and mashup web sites is spreading a culture of sharing data and (some) personal information. But the fact that Wuala stores portions of files across all of its users' computers—no matter how encrypted—will likely be too much of a turnoff for the majority of potential users<sup>375</sup>. »*

Au delà des « pionniers » *digital-literate* de la première heure, qui auraient l'envie et le temps de rentrer dans les détails de l'encryptage et de la fragmentation – conclut Chartier – le partage et la mise en commun des ressources risquent d'avoir un effet boomerang : non seulement parce que un utilisateur aurait l'impression d'avoir ses propres données reparties « de façon aléatoire » sur les machines d'autres utilisateurs, mais aussi parce que ce même utilisateur verrait s'accumuler sur sa propre machine des données de provenance incertaine, et le concept de sécurité offert par Wuala ne suffirait pas à compenser cette situation d'incertitude.

#### ENTRE FRAGMENTS HÉBERGÉS ET SURVEILLANCE COLLABORATIVE : LA « LÉGALITÉ » DU CLOUD P2P

La question posée par Chartier – l'hébergement de données créées par d'autres utilisateurs sur son propre ordinateur – se trouve bientôt débattue en termes de « légalité » de ces contenus, et en conséquence, de celle de l'outil qui permet de les stocker et de les partager – ainsi que de la possible responsabilité de ses utilisateurs

---

<sup>374</sup> Discussion de l'article de Frederic Lardinois sur *ReadWriteWeb*, [http://www.readwriteweb.com/archives/wuala\\_launches\\_p2p\\_cloud\\_stora.php](http://www.readwriteweb.com/archives/wuala_launches_p2p_cloud_stora.php)

<sup>375</sup> David Chartier (2008). « First look : Wuala online P2P backup service makes us nervous », *Ars Technica*, <http://arstechnica.com/old/content/2008/08/first-look-wuala-online-p2p-backup-service-makes-us-nervous.ars>

en cas de violations du droit. A l'automne 2007, le moment où Wuala démarre en *closed alpha*, les débats au croisement des questions de propriété intellectuelle, de confidentialité des données personnelles et de partage de fichiers au moyen de réseaux P2P sont particulièrement animés en Europe : un tribunal allemand vient de refuser la demande de certains groupes de l'industrie musicale, que les FAI identifient les comptes suspects d'être utilisés pour violation du copyright par le partage de fichiers en P2P<sup>376</sup> ; la Société civile des producteurs de phonogrammes en France (SCPP) assigne trois éditeurs de logiciels P2P en justice pour ne pas avoir entrepris les actions nécessaires à empêcher les utilisateurs d'accéder à des contenus illégaux<sup>377</sup> ; un avocat-général de la Cour européenne de justice prend officiellement position dans le cas controversé *Promusicae vs. Telefónica de España* sur la divulgation d'informations de trafic à des fins de poursuite de violations de la propriété intellectuelle<sup>378</sup>.

Malgré l'étiquette novatrice mais « brûlante » de stockage en « P2P social » qui accompagne Wuala, de leur côté, les développeurs – fidèles à leur principe de « priorité à la technologie » – ont tout d'abord rapidement écarté le problème avec l'introduction d'une lapidaire FAQ dans le site Wuala

« *Is the service legal ?* »  
« *Yes, Wuala is legal.* »<sup>379</sup>,

réponse qui se révélera bientôt insuffisante pour plusieurs utilisateurs, se demandant et demandant donc à en savoir plus sur le positionnement du dispositif par rapport aux questions juridiques problématiques autour du P2P. Qu'est-ce que fait et de quoi est faite la « légalité » de Wuala ?

Ce sont, une fois de plus, les utilisateurs les plus « intéressés » qui influencent le projet et contribuent à le transformer – une histoire qui peut être démêlée et racontée grâce aux « posts » et aux discussions sur le forum Wuala. Un utilisateur souligne que les fonctionnalités mêmes de Wuala qui fournissent confidentialité et sécurité – l'encryptage des contenus en tout premier lieu – en font un système de P2P privé qui, comme nombre de *darknet* (Wood, 2010), faciliterait la prolifération d'activités d'échange illégales. De façon intéressante, d'après cet utilisateur, la question n'est

---

<sup>376</sup> Pinsent Mason (2007). « Music industry rebuffed across Europe on file-sharing identifications », *Out-Law.com*, <http://www.out-law.com/page-8353>

<sup>377</sup> Société civile des producteurs de phonogrammes en France (SCPP), « La lutte anti-piraterie », <http://www.scpp.fr/SCPP/Home/LUTTEANTIPIRATERIE/tabid/70/Default.aspx> (dernière mise à jour mai 2011).

<sup>378</sup> Cour de justice de l'Union européenne (2007). Conclusions de l'Avocat Général M<sup>me</sup> Juliane Kokott présentées le 18 juillet 2007, Affaire C-275/06 - Productores de Música de España (Promusicae) contre Telefónica de España SAU, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:62006CC0275:FR:HTML>

<sup>379</sup> FAQ de Wuala, <http://www.wuala.com/fr/support/faq/c/16> (cette page n'est plus en ligne, vue pour la dernière fois en juillet 2011).

pas tellement que cette multiplication de partages illégaux ait effectivement lieu – mais que au vu de ses caractéristiques, le dispositif puisse être interdit à la suite d'éventuelles décisions juridiques restrictives concernant d'autres outils P2P, notamment de partage de fichiers :

*« There is also the question of how people will use Wuala - after all, the service provides almost unlimited and strongly encrypted storage. As Wuala has little control over what is shared on the network, Wuala might turn out to be a haven for rather unsavory activities<sup>380</sup>. »*

Un contributeur d'outre-Atlantique remarque que la définition de la légalité du dispositif pourrait, du point de vue de l'utilisateur, se baser sur le fait que, *de facto*, aucun fichier n'est stocké dans sa totalité sur la machine d'aucun utilisateur. Cela pourrait suffire à garantir que celui-ci ne stocke pas du matériel illégal, à la différence d'autres applications pour lesquelles l'ouverture au partage est en soi une indication qu'on possède le fichier dans sa totalité. L'utilisateur pose par ailleurs la question du contrôle sur sa machine et explique que, indépendamment de ses propres activités, permettre le troc signifie accepter volontairement de renoncer à une partie de ce contrôle, et accepter ses éventuelles implications légales :

*« This brings us to the legality of group sharing and allowing Wuala to use your computer as part of its cloud. Wuala emphasizes that no one single file is ever stored on your computer. Instead, it uses (...) encrypted file fragments. These file fragments are generated on the source machine, and Wuala's servers spread them out redundantly across several sources. This ensures your file is always accessible, and ensures no one computer ever has more than a couple of bits or bytes of your data. Subsequently, traders don't have to worry they are storing illegal data. This is a gray area in the US, especially when it comes to child pornography. Wuala strictly forbids trafficking of such files, but Wuala also can't really police their network 100%. It's certainly not as bad as sharing an illegal file over BitTorrent, since you most likely have the full file in that scenario. Understand that trading means you give away control of part of your drive, and accept any legality in doing so<sup>381</sup>. »*

L'enjeu au centre de cette question de la « légalité » est par ailleurs réaffirmé lors d'une discussion sur le forum utilisateurs du principal concurrent « centralisé » de Wuala, Dropbox : la responsabilité de ce qui circule dans le système peut être moins aisément et moins directement attribuée au fournisseur de service, puisque celui-ci s'est lui-même mis dans la position d'en savoir le moins possible sur ces contenus – jusqu'à ne pas posséder les moyens techniques de les monitorer. Mais est-ce que les utilisateurs peuvent, pour autant, être tenus pour responsables ?

---

<sup>380</sup> Discussion de l'article de Frederic Lardinois sur *ReadWriteWeb*, cité.

<sup>381</sup> Discussion de l'article de Matt Smollinger sur *SmallCloudBuilder*, cité.

La discussion s'engage sur certains cas récemment surgis aux Etats-Unis qui semblent suggérer cette possibilité, conduisant au décryptage « forcé » des données et à la responsabilité subséquente au « simple » stockage de contenus, indépendamment du fait qu'ils soient cryptés ou non, et que les utilisateurs sachent ou pas ce qui est effectivement stocké. Pour certains d'entre eux, une telle éventualité semble être une condition suffisante à se retirer du système :

*« With Dropbox (and just about all other cloud storage systems), your data is stored on either servers that are controlled by the company itself or on trusted third-party mediums (such as Amazon S3, which Dropbox uses). With Wuala, your data is encrypted, broken into chunks, and stored on not only servers owned and controlled by Wuala, but also the computers of Wuala users. That last part (...) creates (...) a very real legal problem. (...) The reason has to deal with the acceptance, storage, and transmission of illegal content. Now, I'm not talking about all the RIAA/MPAA/Copyright crap. Rather, I'm talking about more serious things such as child pornography or classified information. Although the data is encrypted and you may only have bits and pieces of it, if the file in question is determined to be illegal and you are identified as having the file, you could legally be in trouble. »*

*« Since users only store random, incongruous chunks of encrypted data without knowing the decryption key, how could the user EVER be held accountable if the data turned out to be illegal? Also, since the data is incongruous, if hacking was possible and the end user COULD decrypt the blocks, they wouldn't even be holding a full file. »*

*« The real concern that I have is that the social storage aspect of storing files, even encrypted, can be potentially broken. Sure, there are only bits and pieces of files (fragments), but all it takes is the right person with the right fragments to be able to determine what the file may actually contain (even if it's only bits and pieces of it). this can create a legal issue, even if one hasn't been issued yet. This is becoming even more true in the US, where you have cases being made today by the US DoJ about encryption, the legalities of forcing someone to decrypt data (removing plausible deniability), and whether or not the mere storage of such data, even encrypted, can constitute involvement. These are all real concerns that I, for one, am not willing to take to chance [sic]<sup>382</sup>. »*

De leur côté, les développeurs de Wuala sont conscients qu'un positionnement « officiel » de leur part est nécessaire par rapport aux questions qui entourent l'« étiquette P2P » de Wuala, et aux inconnues juridiques posées par le système de troc et la mise en commun des ressources. En particulier, il s'agit pour eux de clarifier si et comment les caractéristiques de leur système se rapprochent des dynamiques qui posent problème dans le cas de systèmes classiques de partage de fichiers en P2P. La réponse qu'ils proposent à la communauté d'utilisateurs comporte deux volets, l'un centré sur l'utilisateur et sa machine, l'autre concernant le réseau dans son ensemble, et le rôle de l'entreprise dans la préservation et la

---

<sup>382</sup> Discussion « Wuala vs. Dropbox ». <http://forums.dropbox.com/topic.php?id=40403>

maintenance de celui-ci.

Premièrement, il est question de la nature des contenus stockés, et de la possibilité d'accès à ces contenus de la part d'entités tierces. Comme certains utilisateurs l'avaient déjà évoqué, l'équipe Wuala se concentre sur la nature éclatée et encryptée des contenus stockés dans le système, qui rend la question « légal/illégal » dépourvue de sens au niveau de la machine de l'utilisateur. Si un utilisateur active le troc de stockage et qu'un fragment de fichier illégal est stocké sur son ordinateur,

*Les fragments individuels sont juste des bits et des octets insignifiants tant que personne ne connaît la clé d'encryptage, la procédure de ré-assemblage, et les autres fragments. Les questions à propos des personnes qui ont stocké le fichier sur Wuala, avec qui le fichier a été partagé et qui ont accédé au fichier sont plus importantes que celles à propos des ordinateurs sur lesquels les fragments individuels d'un fichier illégal ont été stockés<sup>383</sup>.*

Dans la définition de la légalité de Wuala, soulignent donc ses développeurs, l'enjeu n'est pas tant la manière dont les ressources partagées par tel ou tel autre ordinateur sont exploitées, de façon automatique, dans les opérations de fragmentation et d'encryptage : les fragments ne sont pas des objets saisissables par le droit dans sa configuration actuelle. L'enjeu réside plutôt dans les opérations de réseautage et partage entre les utilisateurs qui, stockant pour la première fois des contenus dans le système et les échangeant au moyen des clés d'amitié, introduisent ou recomposent la totalité d'un fichier sur leurs machines.

*« (O)ur system is really not about anonymizing, our system is about building a legal platform, an online storage where everyone has his or her space to upload files. Just for the published files, it would be publicly known who has updated the file. So there might be some legal concerns coming up, for sure, but here we have the exact same problem as YouTube has. If there is an illegal file or a copyright-infringing file in the public area, then if someone reports it to us, we need to take it down. There are different methods for users to do so, they can report it directly via email, they can flag it as inappropriate from within the system so that they pop up much faster, and then we can also detect it manually, although it takes a lot of time, and then we also ask to take it down; but when it comes to the private files, everything that you see in my files, my friends and my groups, is not revealed to us because this is private and encrypted on their computers, we cannot see this data. Basically, we do not see the files, so there is nothing we can do about it<sup>384</sup>. »*

C'est donc à certains utilisateurs, et non à des entités de surveillance tierces, que la

---

<sup>383</sup> FAQ de Wuala, <http://www.wuala.com/fr/support/faq/c/16> (cette page n'est plus en ligne, vue pour la dernière fois en juillet 2011).

<sup>384</sup> Entretien avec K. Gueden, octobre 2009.

légalité des contenus circulants dans le système est confiée. Ce qui permet – reconnaît Kurt – beaucoup d’abus, comme dans tout système de partage de contenus ; mais cette « surveillance collaborative » entre utilisateurs<sup>385</sup>, plus facilement réalisable dans le cas des fichiers partagés en mode public, et entièrement dépendante de la volonté des utilisateurs dans le cas des fichiers privés, leur semble la seule solution techniquement possible au vu des fonctionnalités relatives à la sécurité et à la *privacy* du service.

---

<sup>385</sup> Une procédure de signalement de fichiers « suspects » est mise en place et établie dans les FAQ du site Wuala (<http://www.wuala.com/en/support/faq/c/20>): « Click on the file. Go to the action menu and click on ‘Flag as Inappropriate’. You may also send an email to [abuse@wuala.com](mailto:abuse@wuala.com). Please give us as many details as possible, e.g. link to the file, name of the file, owner of the file, etc. Thank you for reporting. »

#### SECTION 4. QUELLE DÉCENTRALISATION ? LE RÔLE CHANGEANT DES SERVEURS ET DES PAIRS

A plusieurs moments de mon récit, un enjeu s'est dessiné que j'ai choisi de laisser momentanément de côté et le moment est venu, dans cette section conclusive, de l'aborder. Il s'agit des modifications qui, au cours du temps, ont concerné la présence et le rôle dans le système Wuala de centres de serveurs appartenant à la firme ; à côté du réseau P2P d'utilisateurs, ces centres de serveurs ont été utilisés tour à tour comme sauvegarde de sécurité ou comme « nœuds stables » afin d'assurer la redondance pendant la phase d'amorçage (c'est-à-dire, jusqu'à ce que le nombre d'utilisateurs adoptant le système ne devienne suffisant pour assurer sa stabilité). Par ailleurs, entre 2007 et 2011, la firme Wuala réattribue progressivement une importance à la fois symbolique et très « matérielle » à ces centres de serveurs dans l'économie du dispositif – de façon particulièrement éclatante au moment où j'écris<sup>386</sup>.

Un des aspects les plus débattus au cours des premières années de vie du dispositif concerne la dépendance plus ou moins importante du système de stockage vis-à-vis du réseau P2P, et l'importance du recours additionnel à des serveurs. Ces aspects sont largement discutés non seulement au sein de l'équipe, mais avec les utilisateurs – contrairement à d'autres points, pour lesquels l'expression de divergences et leur discussion resteront essentiellement une « affaire interne » à l'équipe de développeurs. Notamment, la (dé)centralisation du dispositif, qui a pu apparaître plus ou moins prononcée aux différents stades de mon analyse, a fait l'objet de négociations complexes et de choix techniques et économiques vertement discutés au cours des quatre années de vie de Wuala, jusqu'à se profiler comme la caractéristique dont l'absence pourrait signifier le manque de « potentiel différenciant » de l'entreprise, et même mettre en péril sa survie.

L'architecture du service Wuala est pour la première fois expliquée de façon détaillée devant un public (même si très sélectionné comme l'est, certainement, l'audience de Mountain View) lors de la Google Tech Talk, au cours de laquelle Dietrich Poell souligne le rôle central du réseau P2P dans l'organisation technique et économique du dispositif :

---

<sup>386</sup> Le 28 septembre 2011, l'équipe Wuala a annoncé, de façon inattendue pour les utilisateurs (ainsi que pour moi qui suis de près le développement du service depuis deux ans) que la firme, désormais rachetée par le développeur français de matériaux de stockage StockFast, allait interrompre l'implémentation du mécanisme de troc dans l'application – décision qui réduit drastiquement, *de facto*, la composante collaborative et P2P du système, limitant désormais sa spécificité au système d'encryptage et décryptage en local. La controverse qui s'est ensuivie, illustrée par les discussions enflammées qui ont eu lieu depuis lors dans le forum Wuala (p. ex. <https://forum.wuala.com/viewtopic.php?f=39&t=2231&sid=98342d997d24c175479b6823fdbeb5c8>), ouvrent la voie à un nouveau terrain prometteur pour les questions abordées dans ce chapitre et dans la thèse, mais des contraintes de temps m'empêchent d'en poursuivre l'analyse.

“... (The) fragments are uploaded into the P2P network, and in addition to that, also the first end fragments are uploaded into our servers. For some time, this has been done to bootstrap the service, until there were enough users in the P2P network. At present, this wouldn't be necessary, users have become enough to ensure system sustainability, but in the future we might still keep it just to give a backup, in the very unlikely case that all the computers in the P2P network were to leave at the same time, you would still be able to retrieve the files from our servers.”

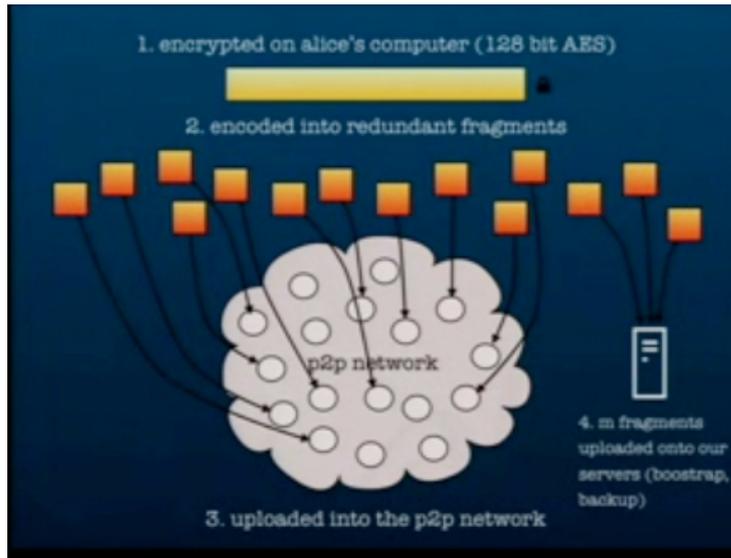


Tableau 8 et Image 8. Extrait de la conférence « Google Tech Talk » donnée au siège Google de Mountain View, Californie, par Dietrich Poell en octobre 2007.

Les paroles de Dietrich, et l'image qui les accompagne, affirment clairement la « philosophie » de décentralisation sous-tendant Wuala : le centre, le cœur du système est le réseau P2P. Les serveurs, propriétés de la firme, sont une bouée de sauvetage au cas où le système – encore instable puisque l'ensemble d'utilisateurs et de ressources est en train de prendre forme – serait abandonné par un nombre important d'utilisateurs à un même instant, au point de compromettre son fonctionnement. S'ils auront un rôle dans le futur, les serveurs resteront « peut-être » comme solution de sauvegarde. Lors de notre entretien au début de 2009, quand on parle du mécanisme d'encryptage et des codes d'effacement, Dietrich utilise un langage similaire :

*« In addition to that, the first 100 encrypted fragments are also stored on our servers, which is important at the moment to bootstrap the service. However, in the long run, it will be self-supporting, but we might still save the first 100 encrypted fragments just to ensure you the availability. It gets whatever it can from the P2P network, and the remaining fragments from our servers.*

*This means that we can save a lot of bandwidth over centralized services, and at the same time guarantee reliable storage<sup>387</sup>. »*

On voit comment, déjà au moment du dévoilement du *closed alpha*, le degré de décentralisation et de « P2P-isation » du dispositif s'annonce comme l'une des briques les plus importantes sur lesquelles va se fonder le rapport des usagers à Wuala le dispositif, et Wuala l'entreprise. Quel rôle pour le serveur ? S'agit-il d'une aide temporaire, en attendant que le système se dote d'assez de nœuds pour assurer sa propre durabilité ? Un « nœud du réseau parmi les autres », juste plus puissant ? Ou encore, une redondance supplémentaire ?

#### DES NŒUDS-USAGERS QUI DEVIENDRONT AUTO-SUFFISANTS...

Quelques temps après que Dietrich a incité l'audience de la Tech Talk à solliciter une invitation afin d'essayer le *closed alpha* de Wuala, quelques réactions et retours commencent à pointer sur les forums, portant sur la faisabilité et la facilité d'usage du « cloud de stockage en P2P ». Si la plupart d'entre elles ont pour objet le mécanisme de troc de l'espace disque pour l'espace en ligne, d'autres réactions concernent les aspects plus spécifiquement liés à l'architecture et à l'infrastructure décentralisée: en quoi le P2P peut-il améliorer la rapidité des téléchargements, augmenter le trafic, ou permettre de ne pas poser de limites à la taille du fichier qu'on veut télécharger :

« J'ai eu l'occasion d'avoir une vitesse d'upload de plus de 900 Ko par seconde et personnellement c'est le maximum que j'aie jamais vu sur un service de stockage de dossier en ligne<sup>388</sup>. »

« Ça a l'air plutôt sympa comme solution. Moi j'avais tester vite fait skydrive, mais j'avais était dégouté par le fait qu'on était limiter en donner par Upload [*sic*] (du genre limite max 80 Mo/Upload)<sup>389</sup> »

D'autres encore sont sceptiques quant à l'applicabilité du mécanisme des codes d'effacement au stockage en ligne :

« Le P2P marche généralement assez bien puisque chacun a une copie conforme du fichier (une fois fini de télécharger), ça réduit drastiquement les risques de pertes de données. Par contre tout découper et distribuer, pour du

---

<sup>387</sup> Entretien avec D. Poell, septembre 2009.

<sup>388</sup> « Dropbox + Mozy = Wuala », *HotlineMicro*, <http://www.hotlinemicro.com/windows-xp/dropbox-mozy-wuala/>

<sup>389</sup> Discussion de l'article « Wuala, le P2P pour votre stockage en ligne », *Le Journal du Geek*, <http://www.journaldugeek.com/2010/03/16/wuala-le-p2p-pour-votre-stockage-en-ligne/comment-page-2/#comments>

stockage, je suis pas convaincu... »<sup>390</sup>

Les débats se concentrent majoritairement, comme on peut voir, sur le réseau P2P d'utilisateurs que sous-tend Wuala. Peut-être à cause de la façon dont Dietrich a présenté le système lors de la Tech Talk (Encadré 3), les serveurs Wuala ne sont presque pas mentionnés par les utilisateurs, qui semblent avoir accepté et saisi leur statut de sauvegarde de sécurité et stabilité (temporaire ?), de « grand nœud » assurant la redondance pendant la phase d'amorçage, jusqu'à ce que le nombre d'utilisateurs adoptant le système ne devienne en lui-même suffisant pour assurer sa stabilité.

En 2010, quelque chose commence à changer dans le vocabulaire utilisé par l'entreprise lorsqu'elle fait référence au modèle technique sous-tendant le dispositif Wuala. L'équipe présente son travail lors de plusieurs manifestations, comme StartUp Camp Switzerland (janvier), Web Monday (mars) et la conférence LIFT 2010 en mai. C'est lors de ces événements, qui suivent de près la fusion de Caleido avec le constructeur français de périphériques de stockage StockFast « pour travailler sur l'avenir du *cloud storage*<sup>391</sup> », que le réseau d'utilisateurs/ressources sur lequel Wuala s'appuie commence à être qualifié de « P2P cloud », ou nuage P2P. Ce déplacement lexical amène une partie de la communauté à s'interroger sur ce que la fusion entraîne en termes d'éventuels changements dans la technologie, ce qui amène l'équipe Wuala à faire un communiqué à l'allure « rassurante » :

Avec StockFast nous pourrions pleinement réaliser la vision sur laquelle nous nous sommes lancés il y a plus de quatre ans, quand on faisait des recherches sur les fondements de Wuala à l'EPF de Zurich: la construction d'un nuage de stockage fiable et sécurisé, capable d'intégrer des millions d'appareils dans le monde entier. La fusion avec StockFast nous donne non seulement de nouvelles ressources opérationnelles, mais aussi une abondance de nouvelles idées et visions, qui nous n'aurions pas été en mesure de mettre en œuvre dans une société de logiciels pure. (...) Nous allons continuer à faire tous les efforts possibles pour vous écouter comme nous le faisons avant. Vous nous connaissez: nous ne changerons pas cette attitude<sup>392</sup>.

De façon intéressante, le communiqué se concentre aussi bien sur la valeur ajoutée du partenariat en termes commerciaux, que sur ses implications en termes d'infrastructure et de ressources – tout en réaffirmant le déplacement vers le concept de nuage, qui est en même temps en train d'acquiescer un véritable statut de « *buzzword* » au sein du secteur IT<sup>393</sup>, mais qui n'est pas la conception de service

---

<sup>390</sup> Discussion de l'article de Yannick Guerrini, « Wuala : le P2P comme solution de stockage », *Tom's Hardware*, <http://www.presence-pc.com/actualite/Wuala-stockage-cloud-P2P-39035/#xtor=RSS-11>

<sup>391</sup> Site web Wuala.

<sup>392</sup> Site web Wuala, FAQ, Section « Wuala and StockFast » (modifiée à l'heure actuelle, vue pour la dernière fois en l'état cité en novembre 2010).

<sup>393</sup> Bien que la définition même de *cloud* soit actuellement l'objet de vives controverses (une revue spécialisée a récemment réuni plus de vingt définitions différentes du concept), ce modèle indique généralement que le vendeur fournit l'infrastructure physique et le produit logiciel, abritant ainsi à la

internet préférée par certains représentants des communautés du logiciel libre et du peer-to-peer : ils l'ont déjà qualifiée sans détours de « pire que la stupidité<sup>394</sup> ».

...OU DES USAGERS-RESSOURCES, « OPTIMISATION » DES SERVEURS ?

En parallèle, l'importance et le rôle des serveurs d'entreprise dans le modèle semblent avoir été révisés. Une nouvelle version du site Wuala parue à la mi-2010, en faisant référence à la Google Tech Talk de 2007, signale :

*« One of the important differences to the somewhat outdated tech talk is that we now do a full backup of all files - in encrypted form - on our servers. We have built up a robust server infrastructure which guarantees that all files are always available<sup>395</sup> ».*

Aussi, quand la nouvelle version du site Wuala rend compte de la raison pour laquelle le dispositif est « fiable », le poids des centres de serveurs Wuala et du nuage/réseau de terminaux utilisateur se voit, une fois de plus, mis en question, dans les illustrations du site en tout premier lieu (Image 10). La fiabilité du système repose sur la « version complète » stockée sur les serveurs, tandis que le réseau P2P fait figure d'« optimisation ». De surcroît, parmi les clients/usagers de Wuala figurent pour la première fois des « organisations » plutôt que des individus :

*We make sure that your data is always safe. A complete backup of your encrypted files is always kept on our servers, so that it can always be accessed and restored. The Wuala cloud serves as an optimization that allows us to operate a large-scale infrastructure at low cost and to provide you with better performance. Also, it allows organizations to get access to reliable online storage by providing their own resources to the Wuala cloud<sup>396</sup>.*

---

fois les applications et les données dans un lieu inconnu de l'utilisateur (le fameux « nuage », *cloud* en anglais) et interagit avec ce dernier grâce à une interface client. On s'achemine dans ce cas vers un modèle de déportation sur des serveurs distants de traitements informatiques traditionnellement localisés sur le poste utilisateur. Voir aussi [Musiani & Schafer, 2011].

<sup>394</sup> <http://www.businessinsider.com/2008/9/gnu-founder-richard-stallman-cloud-computing-worse-than-stupidity->

<sup>395</sup> Blog de Wuala, page « Google Tech Talk », <http://wualablog.blogspot.com/2007/11/google-tech-talk.html>.

<sup>396</sup> Page « Sécurité » du site Wuala (version désormais hors ligne), <http://wuala.com/fr/learn/technology>.

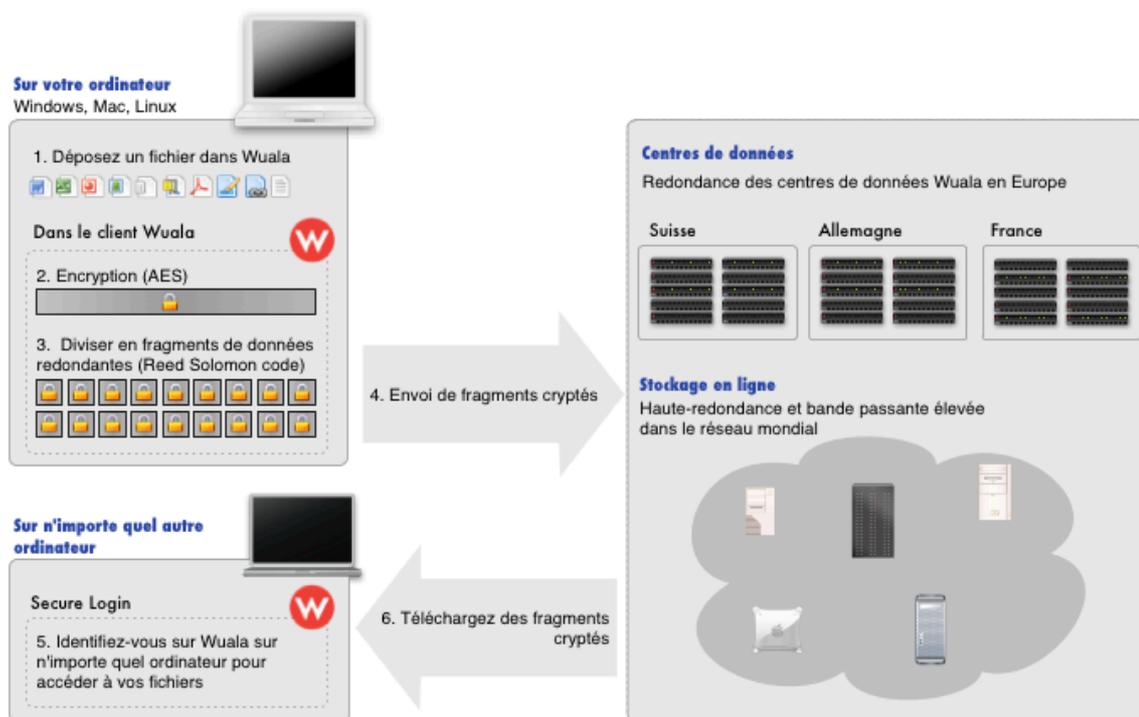


Image 10. Représentation graphique schématique du fonctionnement de Wuala, site Web de l'entreprise (<http://www.wuala.com/fr/learn/technology>), 2010.

Quand, à l'automne 2010, j'interroge Dietrich sur les implications techniques effectives de ce déplacement lexical, sa définition de l'architecture Wuala donne cependant beaucoup d'importance à la décentralisation, qui est encore bien présente en tant qu'aspect démarquant du dispositif « *in such a crowded market* » :

*« Wuala still profits significantly from the P2P grid storage network. We have decided to maintain the complete, redundant backup on our servers, but when it comes to downloading, the Wuala client is programmed to get as many fragments as possible from the grid. This is still the way in which we reduce our outgoing bandwidth and also electrical power significantly, which drives down our costs and is an integral part of how we seek to provide a better service<sup>397</sup>. »*

En même temps, Marc, le manager général et co-fondateur de StockFast – avec qui Caleido a récemment conclu une fusion définie par certains commentateurs comme « acquisition<sup>398</sup> » – exprime son intérêt pour Wuala de la façon suivante :

« On est en train de mettre en chantier un 'réseau Cloud' reposant sur la

<sup>397</sup> Entretien téléphonique avec D. Poell, septembre 2010.

<sup>398</sup> Mike Butcher (2009). « Wuala 'merges' with StockFast to take P2P cloud storage mainstream », *TechCrunch*.

technologie Wuala de Caleido. La technologie Wuala s'adapte parfaitement à notre désir de construire du stockage en cloud à la fois accessible et sûr par la mobilisation de ressources existantes, car elle fait fonctionner de concert du stockage centralisé et du stockage distribué. (...) StockFast apporte des nouveaux marchés pour cette technologie : des entreprises et des utilisateurs professionnels, qui pourront très bientôt tirer avantage des services Wuala à la fois au sein du cloud et au sein de leur datacenter. Le 'cloud storage' en pair-à-pair permet à la fois un échange de données faciles et une flexibilité complète des données<sup>399</sup>. »

Aux yeux de certains usagers, avec cette stratégie qui semble tenir à la fois du modèle d'affaires de l'entreprise, de ses efforts marketing vis-à-vis de nouveaux publics, et de l'organisation technique du dispositif, Wuala est en train de se déplacer dans une arène de compromis assez risquée. D'un côté, un modèle prévoyant la présence de serveurs destinés à un « *complete backup* » a un intérêt sûr pour les potentiels clients institutionnels ou organisationnels, auxquels les solutions « campus » ou « business » proposées dès l'automne 2010 sont destinées, et pour lesquels le P2P apparaît peut-être comme une technologie controversée, voir « incertaine ». D'un autre côté, au sein de la communauté d'utilisateurs pionniers qui s'étaient initialement intéressés au dispositif au vu de son positionnement dans le marché en tant qu' « *alternative* » P2P, des questions se posent. L'une d'entre elles, notamment, est beaucoup posée, soulevant toujours la même objection : comment les deux aspects de l'achat de stockage supplémentaire et, surtout, d'une présence des serveurs qui paraît toujours plus imposante, se conjuguent-ils avec la « philosophie de la décentralisation » qui est la pierre angulaire de Wuala ? L'équipe Wuala répond bientôt sous forme d'un ajout aux FAQs du site :

Wuala est et reste un système de stockage distribué, chaque utilisateur a la capacité de troquer du stockage et de gagner un stockage additionnel proportionnellement à sa contribution au système, sans avoir besoin de payer. Acheter un stockage additionnel est une option pour ces utilisateurs qui ne peuvent ou ne veulent pas troquer, ou qui veulent simplement plus que ce qu'ils peuvent troquer. Par ailleurs, tous les fichiers sont traités de la même manière. Il n'y a pas de différence entre les fichiers des utilisateurs qui utilisent le stockage gratuit, qui troquent du stockage ou ont acheté du stockage. S'il y a plus d'utilisateurs actifs sur Wuala et s'ils ont plus de stockage, cela est bénéfique à ceux qui troquent, parce que ces derniers peuvent gagner plus de stockage seulement quand les fichiers sont stockés<sup>400</sup>.

Cette réponse introduit en effet de nouveaux éléments dans la définition de décentralisation à l'œuvre dans le dispositif Wuala : celle-ci est pour l'utilisateur un choix et une option afin de profiter de la gratuité tout en conservant ou augmentant

---

<sup>399</sup> Entretien par e-mail avec Marc, octobre 2010.

<sup>400</sup> Site web Wuala, FAQ, Section « Online Storage » (modifiée à l'heure actuelle, vue pour la dernière fois en l'état cité en février 2011).

son espace de stockage en ligne ; la distribution signifie aussi la neutralité dans le traitement des fichiers, indépendamment de l'option de stockage choisie – et les clients qui choisissent l'approche distribuée peuvent tirer profit de l'ensemble de ces options, du fait que celles-ci se traduisent en une plus grande quantité de fichiers présents sur le réseau.

Les réactions à ce communiqué prolongent la controverse. Pour certains usagers, cela est une indication du fait que la partie de données stockées qui est désormais répartie sur le réseau P2P est en effet mineure, en proportion, de ce qui avait été envisagé dans la phase *closed alpha*. Impression renforcée par le fait que, comme on l'a vu, le dispositif est distribué, mais il est aussi un logiciel propriétaire – point très important pour la communauté du logiciel libre. L'idée que l'entreprise est maintenant réticente à s'investir pleinement dans la mobilisation des « marges du réseau » en tant que principe fondateur de son infrastructure technique persiste, et pour certains, stabiliser le dispositif d'une façon qui tiendrait à la fois de serveurs gérés de façon centralisée et de plateforme/nuage formée par les utilisateurs serait *de facto* une forme d'exploitation de ces derniers :

*« The source is closed, and so I don't know yet quite how tied the infrastructure is to Wuala the company (if Wuala disappeared tomorrow, would the network still exist?) ».*

*« I share some of your questions. From what I could tell, if wuala's servers go down, the whole system goes down. I believe they have most of the storage and they're a bit tentative about how much they rely on the edge. Hopefully their priority is making the filesystem rock solid. (...) Hopefully their priority after that is to make the system as independent of their servers as possible ».*

*« I like wuala and I want them to succeed. However, my current impression of their architecture is that it reminds me a lot of Mojo Nation<sup>401</sup> circa 2000. In particular the detail that they actually rely on centrally managed servers while trying to take advantage of users's machines<sup>402</sup>. »*

C'est ainsi que la plus récente<sup>403</sup> intervention de l'équipe Wuala (en la personne de Kurt Gueden) sur la question de l'architecture, vise à stabiliser le dispositif en tant qu'hybride entre approches centralisée et décentralisée, dont les frontières se confondent pour un meilleur rendu technique ; un compromis qui aurait été agencé dans le dispositif depuis son démarrage : « Le plus important est que Wuala a toujours eu une approche mixte. Notre but est de tirer le meilleur de la technologie peer-to-peer, qui fournit les meilleurs avantages aux utilisateurs, tout en se basant sur des serveurs centraux, qui sont plus efficaces<sup>404</sup> ». Marc contribue également à

---

<sup>401</sup> <http://www.oreillynet.com/pub/d/262> ; <http://www.moyak.com/papers/mojo-nation.html>

<sup>402</sup> Discussion de l'entrée de blog « Wuala » par Danny O'Brien, <http://www.oblomovka.com/wp/2008/08/15/wuala/>

<sup>403</sup> Avant l'annonce de septembre 2011.

<sup>404</sup> Blog Wuala, décembre 2010.

stabiliser cette conception de l'architecture et de ses possibilités techniques :

« Alors que d'autres opérateurs de nuages investissent en énormes centres de données, qui consomment mégawatts en électricité, pour alimenter leurs serveurs ainsi que les refroidir, Wuala construit son système de stockage en mobilisant des ressources situées aux 'bords' de l'Internet, et notre compétence permet justement de rendre un système de ce genre fiable et sûr, vraiment apte à intégrer des millions d'appareils dans le monde. Cela a de nombreux avantages pour nos utilisateurs, et le résultat est un produit qui réunit l'accès rapide propre du stockage local avec les commodités du stockage en ligne<sup>405</sup> ».

Cela sera plus tard renforcé par Igor, membre de l'équipe de support utilisateur de Wuala, comme le montre la réponse qu'il envoie à un membre de la communauté lors d'une demande de clarification sur l'architecture de Wuala :

*« The data is saved at both the users' network and Wuala servers. The p2p system does allow to download the files very fast. And the system is redundand (sic) and does not critically rely on one or the other, they mutually reinforce one another<sup>406</sup> ».*

Contrairement au cas de Faroo, examiné dans le chapitre précédent, le cas de Wuala nous raconte une histoire de négociations et de compromis entre une architecture centralisée, qui profiterait de la « simplicité » de la concentration et des points de passage obligés, et une architecture distribuée, reposant sur les ressources des usagers. Si l'information stockée sur un pair peut de toute façon être retrouvée ailleurs, sur une machine sous le contrôle direct de l'entreprise, « pour plus de sûreté », le pair est-il encore un égal parmi égaux, ou mieux, l'a-t-il jamais été ? Inversement, si un serveur cesse de se voir attribué un rôle central et incontournable dans le système pour devenir une aide à sa redondance et sa stabilité – aide dotée d'un poids plus ou moins important dans le réseau – est-il encore un serveur ? En cherchant et en négociant la réponse à ces questions, l'équipe Wuala et ses utilisateurs pionniers discutent du chemin vers une stabilisation du dispositif en tant que point d'équilibre entre deux mondes sociotechniques : hybride depuis le début, mélange de *cloud* et de P2P depuis le début, vertical et égalitaire depuis le début.

---

<sup>405</sup> Entretien par e-mail avec Marc, octobre 2010.

<sup>406</sup> Forum Wuala, Post « General Architecture Questions », <https://forum.wuala.com/viewtopic.php?f=34&t=1630>

## EN GUISE DE SYNTHÈSE

Wuala, une start-up suisse proposant un service de stockage en ligne communautaire, est créée en 2007, à un moment où l'industrie de stockage est foisonnante de propositions et de nouvelles applications. A l'instar de ce qu'on a pu observer dans le cas Faroo, un choix de fond sous-tend la démarche de développement du logiciel dès l'origine : travailler sur l'architecture technique du service en tant qu'aspect différenciant et original par rapport à ses concurrents, généralement basés sur un modèle client-serveur classique ou sur des plates-formes en « nuage ». L'idée centrale du dispositif est donc la suivante : les fichiers téléchargés par les utilisateurs dans le système seront stockés à la fois dans des centres de données classiques et sur un nuage de stockage composé d'une partie des disques durs de chaque utilisateur, reliés entre eux en architecture P2P.

Pourtant, la courte mais dense histoire du dispositif qu'on a exploré dans ce chapitre est celle d'une application dont les créateurs ont travaillé, selon les moments, avec l'hybridation entre différents modèles d'architecture technique, avec la représentation de ces architectures auprès de leur public de clients potentiels, et avec la volonté de ces clients/usagers de s'engager dans un « jeu » de partage de ressources aux nombreuses implications techniques, juridiques et sociales. En suivant, après Faroo, une deuxième histoire de ce que la décentralisation de l'architecture technique change aux relations entre développeurs, entre utilisateurs, entre les deux, ce chapitre a exploré une approche hybride du P2P. Il a suivi comment se construit le compromis entre une approche largement dépendante d'une plateforme de serveurs, et une approche d'architecture distribuée et décentralisée, où le fonctionnement du dispositif repose sur la mise à disposition de ressources informatiques de la part des usagers.

Le chapitre s'est tout d'abord concentré sur la manière dont l'architecture P2P est devenue l'aspect différenciant du dispositif, grâce à un travail sur les « tuyaux » qui reconfigure la sécurité du système, la privacy des données stockées, la vitesse des téléchargements, la rapidité des échanges effectués. En particulier, on a vu comment les « fragments encryptés » résultant du traitement des contenus stockés dans Wuala ont été saisis par les développeurs et les utilisateurs, tour à tour, en tant que symbole même de la solution de stockage permanent en P2P proposée par la firme, solution sur laquelle repose une bonne partie de la crédibilité et de la spécificité du produit ; en tant que mécanisme technique qui assure la sécurité des données par rapport non seulement aux utilisateurs non autorisés, mais à la firme elle-même ; ou encore, en tant qu'arène où se joue la controverse entre utilisateurs « méfiants » et « confiants » quant à la capacité du dispositif de maintenir sa promesse de sécurité et de confidentialité.

La première partie du chapitre a aussi abordé l'articulation des dimensions locale et globale qui constituent le rôle envisagé de l'utilisateur dans le dispositif Wuala : « la 'localité' et la 'centralité' s'échangent en de multiples endroits [...] parce que le réseau est agrégateur et qu'en même temps, il est construit sur un compromis » (Akrich, 1989 : 39). On a vu se construire un dispositif dans lequel, jusqu'à ce que

l'encryptage et la fragmentation des données ne soient complétés, le cœur du service est le terminal de l'utilisateur ; c'est sur ce terminal qu'ont lieu, grâce au client P2P, les opérations d'encryptage et de division en fragments. Une fois que les données sont sous forme de fragments encryptés, chaque fichier présent dans le système est stocké et partagé, « éparpillé » sur le réseau. C'est alors que s'effectue le passage de la dimension locale de la machine utilisateur à l'ensemble du réseau, et que les contenus stockés deviennent globaux – leur emplacement physique, ou plutôt leurs multiples emplacements, étant le réseau constitué par l'ensemble des ordinateurs. En découle une reformulation du lien entre l'internaute et le fournisseur de service : se dessine un modèle technique où l'encryptage en local d'abord, et la fragmentation après – avant toute opération de partage ou téléchargement qui implique une circulation des données dans le réseau – sont proposés par Wuala comme la preuve que la firme « n'a même pas les moyens techniques » de trahir la confiance des utilisateurs, car dans le pire des cas – celui où le réseau serait compromis – un éventuel attaquant en retirerait seulement des fragments non exécutable, et sans signification indépendante.

La section suivante du chapitre a suivi les « premiers pas » de Wuala dans un contexte et une communauté de recherche et de développement qui, depuis 2005, travaille à la mise au point de systèmes hybrides, au croisement entre réseaux P2P et réseaux sociaux. On a vu comment les développeurs de Wuala ont pris en compte la continuité entre la fonction de stockage et celle de partage, ainsi que la conformité de la technologie P2P à se mettre au service de cette continuité ; ils ont ainsi travaillé à une architecture pouvant mêler une structure distribuée à des dynamiques propres au réseautage social. Le déroulement de certaines opérations de sécurisation et conservation de données en local a donc été associé à une dynamique où des liens directs – directs dans les « couches inférieures » aussi bien qu'au niveau de l'interface – peuvent être établis entre les utilisateurs pour les opérations de partage. On retrouve dans cette démarche une conception de confidentialité qu'on a déjà appelé, pour le cas de Faroo, « privacy par la technique » : elle prend forme au moyen des contraintes et des opportunités qui dérivent du traitement et de l'emplacement des données, selon les moments et les opérations se déroulant au sein du système. On a enfin vu comment cette conception de privacy a tout particulièrement pris forme autour de deux « objets », le mot de passe et la clé d'amitié, ainsi que dans une attention particulière au rapport entre la confidentialité des données (personnelles, mais aussi les contenus stockés dans le cloud P2P) et le système d'allocation de ressources sur lequel se base le dispositif. En particulier, le mot de passe en vient à constituer une forme de dégagement du fournisseur de service par rapport aux problématiques de sécurité : un détail dont l'importance peut sembler modeste à première vue, mais qui finit par entraîner des changements dans les formes de solidarité technique (Dodier, 1995) qui s'instaurent entre les utilisateurs et fournisseur.

Le chapitre s'est ensuite concentré sur les manières dont le véritable cœur du modèle technique de distribution à l'œuvre dans Wuala est devenu aussi le cœur de son modèle socio-économique. Il s'agit là de l'idée qu'un utilisateur puisse échanger l'espace inutilisé du disque dur sur son propre ordinateur contre de l'espace de stockage sur le « nuage d'utilisateurs » de Wuala, pourvu qu'il reste en

ligne un certain temps par jour. On a vu comment le principe à la base de ce mécanisme de troc entre espace disque et espace en ligne n'a pas été de donner une plus grande capacité de stockage mais plutôt d'améliorer la qualité de l'espace de stockage dont l'utilisateur est normalement doté sur son disque dur, la possibilité de transformer son disque dur local en disque dur en ligne, potentiellement accessible en mobilité, et cela même si son propre ordinateur est éteint. La section a d'abord suivi le choix, de la part des développeurs, de prévoir une double « porte d'entrée » au système, mêlant client et démarrage Web, et comment ce choix a stabilisé le principe « faire le pont entre le bureau et le Web » comme pierre angulaire du modèle économique de Wuala. Ensuite, on a vu comment le type de connexion Internet de l'utilisateur s'est qualifié comme un enjeu de première importance pour que l'utilisateur puisse pleinement bénéficier du mécanisme de troc, jusqu'à constituer une véritable barrière à l'entrée du système, au cas où le client Wuala se heurte à un pare-feu refusant certains types de connexion. L'architecture technique fabrique, une fois de plus, des formes de représentation de la demande au sein du service.

L'enjeu principal pour le succès du mécanisme de troc a pourtant semblé être la manière dont les utilisateurs s'en saisissent et se l'approprient – leurs réactions au fait de stocker ses données dans un environnement qui est en même temps local et global, concentré et dispersé. On a vu comment les retours de la part des utilisateurs de la première heure se sont concentrés sur la manière dont la distribution du modèle au niveau technique a des effets sur le niveau d'implication, de participation et de « collaboration » des usagers au dispositif, et reconfigure ainsi leur statut de clients. Si certains ont apprécié la possibilité de troquer l'espace disque pour l'espace en ligne, et souscrit très volontiers à l'option qui constitue à leurs yeux la vraie nouveauté et l'avantage économique du dispositif, le caractère controversé des choix de distribution et de partage faits par Wuala a par ailleurs commencé à être discuté dans nombre d'autres interventions. On a donc vu comment les doutes sur la faisabilité, la durabilité et la « prise en main » du modèle Wuala se sont concentrées autour de deux grandes questions.

La première a concerné la stabilité socio-technique du dispositif : le fait qu'un nombre suffisant d'utilisateurs apporte suffisamment de ressources, de façon suffisamment constante, pour que les contenus stockés sur le système soient présents de façon durable et fiable. La deuxième a mis en discussion les frontières « nuageuses » de l'accès aux données personnelles ou stockées par d'autres utilisateurs, et les répercussions de la répartition des données et des ressources qui semble être inscrite dans l'outil sur la confidentialité des contenus et la sécurité des équipements. Cette question a également été à la base des débats sur la « légalité » du dispositif, et de la possible responsabilité de ses utilisateurs en cas de violations du droit – débat dont on a suivi les deux volets, un centré sur l'utilisateur et sa machine, l'autre concernant le réseau dans son ensemble. On a vu une définition de légalité du P2P prendre forme autour de la nature éclatée et encryptée des contenus stockés dans le système, qui rend la question « légal/illégal » dépourvue de sens au niveau de la machine de l'utilisateur, et de la « surveillance collaborative » associée aux opérations de réseautage et partage entre les utilisateurs.

La dernière section du chapitre a enfin analysé comment la (dé)centralisation du dispositif a fait l'objet de négociations complexes et de choix techniques et économiques vertement discutés au cours des quatre années de vie de Wuala, jusqu'à se profiler comme la caractéristique dont l'absence pourrait signifier le manque de « potentiel différenciant » de l'entreprise, et mettre en péril sa survie elle-même. Je me suis intéressée ici aux modifications qui, au cours du temps, ont concerné la présence et le rôle dans le système Wuala de centres de serveurs appartenant à la firme ; à côté du réseau P2P d'utilisateurs, ces centres de serveurs ont été utilisés tour à tour comme sauvegarde de sécurité ou comme « nœuds stables » afin d'assurer la redondance pendant la phase d'amorçage (c'est-à-dire, jusqu'à ce que le nombre d'utilisateurs adoptant le système ne devienne suffisant pour assurer sa stabilité), jusqu'à ce que la firme Wuala ne leur réattribue progressivement une importance à la fois symbolique et très matérielle dans l'économie du dispositif.

C'est là que l'on a vu les deux cas de Faroo et de Wuala se différencier le plus, dans la mesure où ce dernier nous raconte une histoire de négociations et de compromis entre une architecture centralisée, qui profiterait de la « simplicité » de la concentration et des points de passage obligés, et une architecture distribuée, reposant sur les ressources de chaque usager. Usager qui s'interroge, à son tour, sur son rôle au sein du dispositif distribué. Egal parmi des égaux, qui va être libéré des points de passage obligés dès que la stabilité du système le permettra ? Contributeur de ressources – « optimisateur », même – pour un système qui reste sous l'autorité d'un groupe de serveurs géré de façon centralisée ? En négociant la réponse à ces questions, l'équipe Wuala et ses utilisateurs pionniers discutent du chemin vers une stabilisation du dispositif en tant que point d'équilibre entre deux mondes sociotechniques : hybride depuis le début, mélange de *cloud* et de P2P depuis le début, vertical et égalitaire depuis le début.

Ce cas est aussi témoin d'un effet intéressant du croisement entre la décentralisation et une dynamique familière aux sociologues de l'innovation : un système décentralisé complique beaucoup la construction autour du service d'un ensemble d'indicateurs, destinés à permettre des ajustements progressifs entre l'offre et la demande. Ceux-ci ne peuvent qu'être, à leur tour, rassemblés et mesurés de manière décentralisée, transitoire, temporaire, partielle : ce qui peut aussi expliquer en partie la pénurie, dans mon analyse, d'indicateurs quantitatifs concernant les clients/usagers de Wuala et leur évolution au fil du temps. Ce sont les développeurs et entrepreneurs, en tout premier lieu, qui ont « des problèmes pour compter », même avant que le sociologue puisse éventuellement se réapproprier de leurs calculs.

En conclusion, ce chapitre a cherché à démontrer comment Wuala a constitué, au cours de ses quatre premières années d'existence, une arène d'expérimentation technique à la recherche d'un modèle socio-économique durable pour le stockage distribué. Cette expérimentation a eu lieu à la frontière de la dichotomie client-serveur/pairs, en redéfinissant en même temps la conception et la définition d'un système distribué et décentralisé. En travaillant au croisement de plusieurs tendances d'organisation et d'appropriation du dispositif – entre centralisation et décentralisation, entre réseaux P2P et réseaux sociaux, entre le local d'un terminal

utilisateur et le global du Web – on a observé comment les marges du réseau, où se déploient l'incertitude technique et l'opportunité socio-économique, ont été mobilisées pour leur richesse de ressources à la fois matérielles et « humaines ». Ressources qui, au delà de la curieuse appropriation de *buzzwords* telles que « cloud », ont été saisies par l'entreprise comme la clé pouvant l'aider à trouver sa niche et son succès, au sein d'un secteur à la concurrence farouche.

## CHAPITRE 6

### DE TRIBLER À P2P-NEXT, OU LA CONSTRUCTION DU PEER-TO-PEER COMME VALEUR POUR L'EUROPE

#### INTRODUCTION

Université technique de Delft, Pays-Bas, 29 mars 2010. Je suis en déplacement terrain auprès d'une équipe d'ingénieurs informatiques qui développent Tribler, un logiciel peer-to-peer pour le streaming de contenus audiovisuels. Le matin du deuxième jour, je suis reçue par Friedrich Bierbach, le directeur scientifique du projet, dans un très grand bureau rempli de matériel informatique et de papiers : le scénario auquel on pourrait s'attendre si on pensait par stéréotypes (ce que je suis décidée, par ailleurs, à ne pas faire). Friedrich m'explique au cours des trois quarts d'heure suivants comment une petite expérimentation avec les technologies de partage en P2P, née dans le milieu « *nerd* » d'une université technique, est devenue le référent scientifique d'un projet européen financé à hauteur de plusieurs millions d'euros et impliquant nombre de partenaires industriels des secteurs télécom et média. Bien qu'il soit clair que Friedrich ne s'est jamais trouvé dans une situation de ce genre – parler à quelqu'un qui s'est présenté comme sociologue et qui fait donc de la « science différente », ainsi me définit-il – il est tout de même habitué à parler à des journalistes, ce qui permet à l'entretien de s'écouler sans trop de silences, ni d'expressions perplexes. Cependant, je me rends compte avec quelque peu de frustration que mon entretien clé est en train de se résumer, ou quasiment, à quelques slogans de presse. C'est alors que je lis sur le tableau à côté de nous, ces mots gribouillés au feutre bleu :

« *Continue, until court determines Delft is working on forbidden innovations.* »

Je lui demande de m'en dire plus sur le sens de cette phrase. Elle a, sans doute, dû me paraître instinctivement prometteuse à ce moment : mais j'aurais difficilement pu prévoir que ces dix mots étaient ceux qui allaient m'ouvrir les portes de l'histoire de Tribler. En effet, la présence encombrante des tribunaux – si familière pour nombre d'autres applications P2P – ainsi que le mot « défendu » me semblaient *a priori* bien loin des priorités et de la démarche de ses développeurs : les contenus de leur site, que j'avais parcouru consciencieusement au moment de sélectionner Tribler comme

l'un de mes terrains, se focalisent sur les aspects sociaux et communautaires du logiciel, et le définissent de manière assez lapidaire comme « P2P légal ».

Pourtant – je le comprendrai par la suite, dans le reste d'un entretien bien plus fécond – ces dix mots ont rappelé à Friedrich nombre de décisions prises et négociées, d'opinions recueillies sur son travail, d'obstacles maîtrisés ou non. Négociations qui ont parsemé la démarche de développement d'un logiciel qui a, par moments, capitalisé sur son étiquette P2P, et a parfois dû s'en défendre : sous-tendant un modèle à la fois différenciant et risqué, dans un contexte où le principal « client » est l'Union européenne avec ses puissants moyens financiers et sa machine bureaucratique et juridique très articulée. En rendant compte de ces négociations, ce chapitre souhaite montrer comment un prototype de logiciel P2PTV est sorti de ses frontières académiques pour être saisi par des procédures bureaucratiques, et par des espaces de collaboration entre recherche publique et secteur privé. Ce faisant, il a contribué à la problématisation politique et technique du pair à pair en tant que modèle alternatif pour les services Internet, ainsi qu'à sa légitimation comme « valeur » dans le contexte européen. L'histoire de Tribler est celle de la recherche d'une réponse à une question centrale du P2P : comment le construire comme modèle alternatif crédible dans l'espace politique et, surtout, juridique ?

Tribler est défini par ses créateurs, aux premiers stades de son développement, comme une « *social community that facilitates content sharing through a P2P network*<sup>407</sup> », un logiciel de partage en P2P assez classique, permettant à ses usagers de trouver, de consommer et de partager des contenus numériques ; il s'inspire du protocole de communication BitTorrent<sup>408</sup> mais cherche à combiner son mécanisme sous-tendant avec les « aspects sociaux des communautés en ligne<sup>409</sup> ». Ces aspects sociaux sont déterminés par l'incorporation au logiciel d'un moteur de recommandation (« assez similaire à celui d'Amazon, » me dira l'un des développeurs<sup>410</sup>), qui propose des suggestions de contenus, et permet aux usagers d'obtenir des recommandations de la part des autres utilisateurs jugés fiables. Initialement, il s'agit d'un projet de recherche piloté conjointement par l'université technique de Delft et la Vrije Universiteit d'Amsterdam. Tribler a pu compter sur une équipe d'une trentaine de personnes, à géométrie variable (selon les départs et les

---

<sup>407</sup> Site web du projet Tribler, <http://www.tribler.org/trac/wiki/whatIsTribler> et Pouwelse et al., 2006.

<sup>408</sup> BitTorrent est un protocole de transfert de données en P2P, conçu en 2001 et lancé en 2002 par Bram Cohen avec la création de la société BitTorrent, Inc. Le protocole BitTorrent base son fonctionnement sur les principes suivants. Afin d'éviter la saturation qui se produit quand une information se trouve sur un serveur et *plus* elle est demandée, *moins* elle est accessible, BitTorrent tire avantage du principe P2P selon lequel chaque client qui a téléchargé l'information devient à son tour un serveur. Par ailleurs, BitTorrent facilite ultérieurement le commencement des téléchargements par rapport à d'autres réseaux P2P où un client doit posséder l'entièreté de la ressource afin de pouvoir la distribuer. BitTorrent découpe l'information à partager en segments, et les distribue à différents pairs afin qu'ils aient aussi rapidement, à leur tour, quelque chose à échanger.

<sup>409</sup> Site web de Tribler.

<sup>410</sup> Entretien du 30 mars 2010 avec Jingerzi, doctorant en informatique à l'université technique de Delft, travaillant dans l'équipe Tribler.

arrivées des étudiants de Master et de doctorat qui en ont fait partie), guidés par le professeur associé de TU Delft, Friedrich Bierbach. Le développement de Tribler a ses racines dans la convergence de plusieurs axes de recherche de longue date à TU Delft, portant sur les technologies de réseau P2P, de streaming vidéo, et d'interaction machine-usager.

Tribler est lancé en mai 2007, et pendant la première année où le logiciel client est disponible sur le Web, il est téléchargé plus de 130 000 fois selon les estimations de l'équipe<sup>411</sup>. À ses débuts, le développement de Tribler est financé en partie par le projet I-Share, composante du programme de recherche sur l'intelligence ambiante Freeband<sup>412</sup> – lui-même financé par le ministère de l'économie des Pays-Bas ; et en partie par la fondation hollandaise pour les nouvelles technologies STW<sup>413</sup>, jusqu'à l'été 2009. Dans ses premières phases, Tribler est donc un projet financé par de l'argent public, ce dont s'émerveillent plusieurs utilisateurs de la première heure au vu du caractère « *a priori* controversé » du P2P dans d'autres pays européens<sup>414</sup>. Cela semble pourtant beaucoup plus normal à l'un des développeurs de Tribler, Carlo, un Italien qui a désormais « passé suffisamment de temps aux Pays-Bas, » remarque-t-il ironiquement, pour savoir que « le gouvernement ici, en comparaison [avec d'autres pays européens] finance vraiment beaucoup de choses différentes. I-Share, le prédécesseur de Tribler, était littéralement créé et financé par un des ministères hollandais. Ce n'est pas ailleurs qu'on ferait ça<sup>415</sup>. »

En 2008, Tribler est lancé sur Tribbox – une plateforme *open source* développée par AHT, un fabricant hollandais de décodeurs pour télévision. À la fois pour l'équipe de recherche et pour l'entreprise, la rencontre entre le logiciel peer-to-peer et la télévision par Internet semble prometteuse. À l'AHT, on voit dans les fonctions « sociales » de sélection et manifestation des préférences, proposées par Tribler, une manière de faire face à la démultiplication de l'offre télévisée qui est en train de se mettre en place avec l'Internet, en permettant un meilleur repérage des programmes face à la croissance de l'offre, et en accompagnant la délinéarisation<sup>416</sup> des

---

<sup>411</sup> Entretien avec Friedrich Bierbach, professeur associé à l'université technique de Delft, responsable du projet Tribler.

<sup>412</sup> <http://www.freeband.nl/>

<sup>413</sup> <http://www.stw.nl/Default.htm>

<sup>414</sup> <http://elibtronic.ca/content/20101209/some-tools-overthrow-empire>

<sup>415</sup> Entretien du 30 mars 2010 avec Carlo, Business Contact/Valorization Officer pour Tribler depuis janvier 2010. L'entretien ayant été conduit en italien, notre langue natale à tous les deux, j'en traduis les extraits en français.

<sup>416</sup> On entend par délinéarisation une consommation libre des médias, dictée par la volonté du spectateur et non plus du média source. Ceci passe par un certain nombre de pratiques que l'utilisateur se crée ou qui lui sont permises par les fournisseurs de contenus. Par exemple, consulter directement une partie d'un programme ; choisir où, quand et sur quel support (internet, mobile, tablette, TV connectée) regarder le contenu ; en parler avec d'autres spectateurs via les réseaux sociaux par exemple. Voir aussi <http://www.hexaglobe.com/en/news/44-comment-faire-de-la-delinearisation-des-contenus-une-veritable-opportunite-pour-les-medias.html> et [Missika, 2006].

programmes et donc les transformations de la consommation :

*« television is being opened up all over the world via the Internet, the supply will increase enormously. This means that in future it will be hard to find the programmes you find interesting in the midst of all the content. What interests us about Tribler is its characteristic to combine social aspects with peer-to-peer so people can help each other find the best programmes. The viewer may no longer be schedule-dependent, but find all programmes via a network of friends<sup>417</sup>. »*

De leur côté, Friedrich Bierbach et son équipe concentrent leur intérêt sur les aspects techniques du modèle de distribution. L'organisation du modèle de distribution classique des contenus télévisés, transmis d'un point central à un nombre important de spectateurs, ne se prête pas, à leur avis, à la transmission via l'Internet – soit parce que les données vidéo sont lourdes et « goinfres » en bande passante, soit parce que un point central devrait transmettre ces données de manière répétée à beaucoup de terminaux<sup>418</sup> :

*« a two-party structure isn't optimal. Television is generally broadcast from a central point to a large number of viewers. On the Internet, this means that the television server has to set up image transmission separately to each user. Television pictures comprise so much data that only the very largest server systems can broadcast internet television to a large number of viewers. With peer-to-peer technology, however, this limitation does not apply. Each viewer in a peer-to-peer television system takes over part of the server's task and in turn serves another viewer<sup>419</sup>. »*

Par ailleurs, lors du lancement de la première expérience Tribbox, Friedrich montre qu'il n'est pas insensible aux aspects du projet qu'il n'hésite pas lui-même, d'ailleurs, à définir comme « idéologiques<sup>420</sup> ». En 2007 et 2008, au cours de plusieurs interventions orales et écrites au sujet de Tribler et de Tribbox<sup>421</sup>, il relie explicitement le projet Tribler à des questions d'efficacité de distribution de matériaux culturels, de réorganisation des programmes, de neutralité du réseau, d'abattement des barrières à l'entrée de nouveaux acteurs dans le domaine de l'industrie numérique, et même de l'« effet démocratisant » produit par ces dynamiques. Le protocole BitTorrent devient, suivant ses mots, le symbole de la « capacité suffisante » de chaque spectateur de télévision potentiel à profiter de son

---

<sup>417</sup> TU Delft, « Dossier Tribler », <http://tudelft.nl/en/current/dossiers/archive/tribler/>

<sup>418</sup> La diffusion hertzienne présente aussi, par ailleurs, des problème d'occupation de la bande passante (les ondes en l'occurrence), et aussi celui des répéteurs (mais sans doute dans une moins bien grande proportion).

<sup>419</sup> Entretien avec F. Bierbach.

<sup>420</sup> Id.

<sup>421</sup> « Dossier Tribler », cité.

expérience de télévision en ligne. Il conclut même à une occasion: « *It will then become very simple to begin your own TV station. Who will need YouTube?*<sup>422</sup> »

Pendant ses deux premières années de vie, Tribler se positionne donc comme dispositif au potentiel novateur dans plusieurs arènes économiques, politiques et techniques. Un projet guidé par l'intérêt du public, mais dans lequel des acteurs du secteur privé voient une opportunité de marché. Un réseau socio-technique pouvant faire face, à moyen et long terme, aux modifications de l'offre et de la demande dans un secteur qui est en train de changer profondément. Un système qui, construit sur une technologie aux usages politiquement controversés, en propose une relecture aux connotations, fortes et également « politiques », d'ouverture, de transparence et d'interopérabilité. Ce faisant, Friedrich établit des partenariats à la fois dans le public et dans le privé, ouvre des portes (et en referme d'autres), entretient l'intérêt du public hollandais, tandis que son équipe continue à se concentrer sur le développement de la partie technique du logiciel. Selon Friedrich, c'est le fait de s'être activé sur cette pluralité de fronts, et peut-être d'avoir su, par moments, en brouiller les frontières, qui lui vaudra l'attribution d'un financement de quatorze millions d'euros dans le cadre du septième programme de l'Union européenne. À partir de 2009, l'équipe Tribler sera chef de projet scientifique et l'un des principaux partenaires d'une complexe coalition, désormais appelée P2P-Next Consortium. L'objectif de P2P-Next est tout sauf modeste : construire le réseau européen de télévision par Internet du futur, basé sur la technologie de streaming vidéo en *peer-to-peer* de Tribler. Le chapitre qui suit est l'histoire de la poursuite de cet objectif.

2005 : Début de la collaboration entre l'université technique de Delft et la Vrije Universiteit d'Amsterdam (Pays-Bas), autour d'un intérêt commun pour le développement d'un « BitTorrent social ».

2005: Premier *fork* du client BitTorrent ABC, qui donne naissance à Tribler (version 3.x), projet de recherche piloté conjointement par TU Delft et la Vrije Universiteit.

[2006-2009: Tribler est financé par le projet I-Share].

2006 : Publication du papier « Social-based paradigm in P2P » (Pouwelse et al., 2006).

2007 (février): Tribler 3.3.4 (première version avec l'algorithme Buddycast).

2007 (mai): Lancement du *public beta* de Tribler.

2007 (fin): Début de la série 4.x de Tribler. Collaboration avec l'université de Harvard pour la fonctionnalité de streaming video.

---

<sup>422</sup> « Dossier Tribler », cité.

2007-2008: 130 000 téléchargements du client Tribler.

2008: Tribler est publié sur Tribbox, une plateforme *open source* développée par le fabricant hollandais de décodeurs AHT.

2009 (janvier): Début de la série 5.x de Tribler. Les fonctionnalités sociales ne sont plus visibles par l'utilisateur.

2009 : Naissance du consortium P2P-Next et début de son financement par l'Union européenne. L'équipe Tribler est chef de projet scientifique.

2009 : Première participation de P2P-Next à l'International Broadcasting Convention.

2010 : Deuxième participation de P2P-Next à l'International Broadcasting Convention.

[2011, septembre : mon enquête de terrain prend fin]

2011 : Revue de mi-projet pour P2P-Next.

2012 (juin) : Début de la série 6.x de Tribler.

*Tableau 9. Chronologie du développement de Tribler et de P2P-Next, 2005-2012.*

## SECTION 1. UN PROJET UNIVERSITAIRE POUR UN PARADIGME DE P2P « SOCIAL-BASED »

En 2005, plusieurs chercheurs en sciences informatiques de l'université technique de Delft et de l'université Vrije à Amsterdam se réunissent autour d'un intérêt commun : le développement de « BitTorrent *with a social twist*<sup>423</sup> ». Pendant les années qui ont suivi le lancement de BitTorrent, les projets de modification et d'amélioration du protocole originaire se sont multipliés au sein de la communauté *open source*, au point que ABC, l'un d'entre eux, est labellisé de manière élogieuse « *Yet Another BitTorrent Client*<sup>424</sup> ».

### DE ABC À TRIBLER : LE « SOCIAL IMMÉDIAT »

C'est donc « encore un autre client BitTorrent » que l'équipe mixte des deux universités hollandaises s'attache à modifier. A la fin de 2005, un premier *fork*<sup>425</sup> de ABC voit le jour : au mécanisme de distribution propre de BitTorrent, les développeurs ont cherché à ajouter des caractéristiques « sociales », telles que les notions d'amitié, de « *taste buddies* » (personnes qui partagent les mêmes préférences), de recommandations faites par des usagers à d'autres usagers, et d'amélioration de la rapidité des téléchargements basée sur la réputation des usagers dans le système.

D'un point de vue visuel, la première version de Tribler – comme est appelé le logiciel nouveau-né, en voulant rappeler la dimension communautaire des interactions au sein des « tribus » – ne se différencie pas beaucoup de son « père », ABC<sup>426</sup> : la dimension sociale qui doit donner sa spécificité au logiciel n'est presque pas visible pour l'utilisateur, si ce n'est dans l'intitulé du logiciel lui-même (Image 10).

---

<sup>423</sup> Entretien avec F. Bierbach.

<sup>424</sup> ABC (Yet Another BitTorrent Client), <http://pingpong-abc.sourceforge.net/>

<sup>425</sup> Un *fork*, ou embranchement (dérivation, bifurcation), est un nouveau logiciel créé à partir du code source d'un logiciel existant.

<sup>426</sup> Blog «Serializable Thoughts» (26 août 2011). «History of Tribler» (<http://blog.shinnoir.nl/2011/08/history-of-tribler.html>), entrée de blog proposée par un ancien développeur à l'occasion du lancement imminent de la version 5.4 du logiciel.

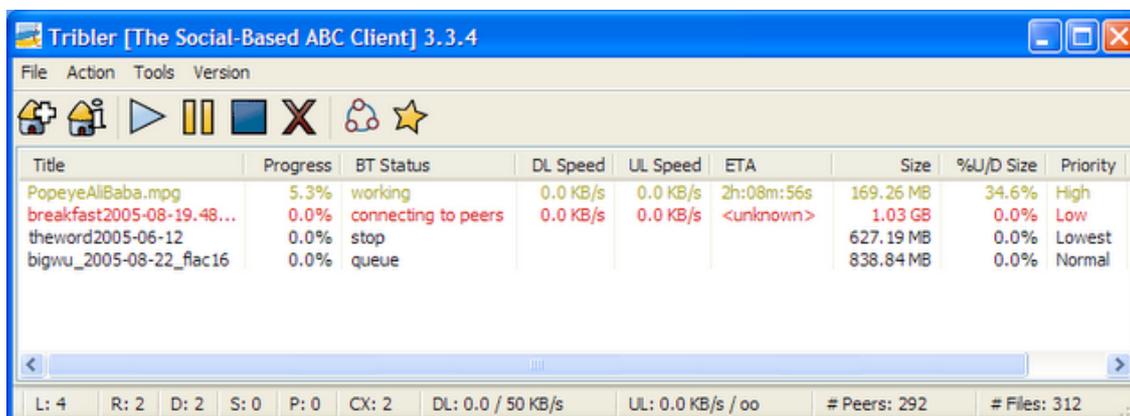
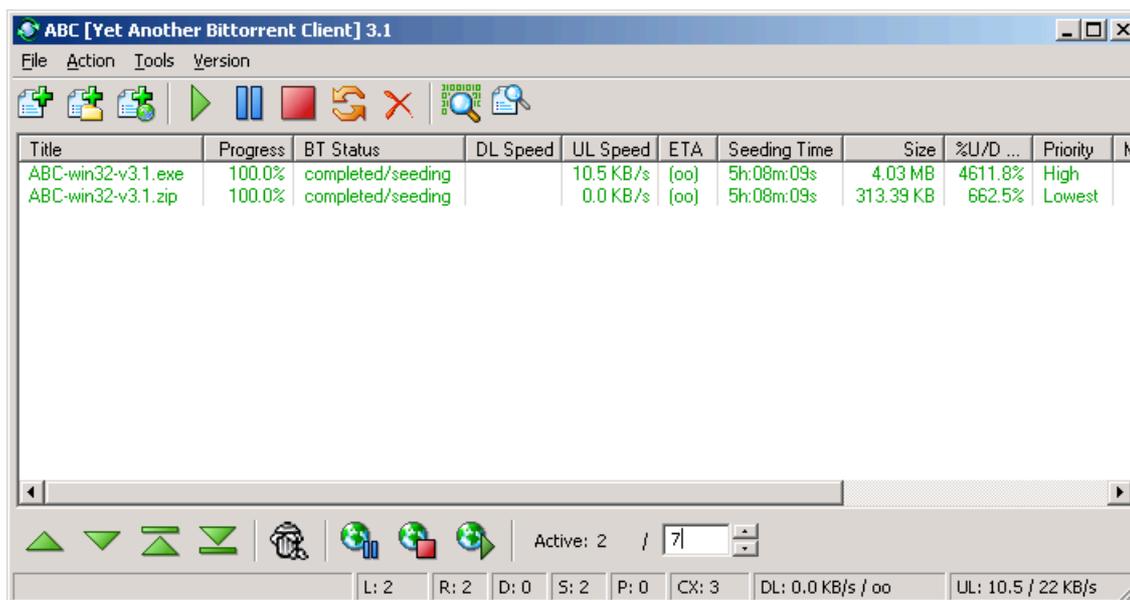


Image 10. De ABC à Tribler 3.x (première version).

Pourtant, des embryons de dynamiques « sociales » telles que la recommandation et l'identification de préférences partagées sont déjà présents. En effet, l'utilisateur peut suivre en direct non seulement les pairs auxquels il est en train de se connecter, mais aussi les fichiers que soit lui-même, soit d'autres pairs possèdent simultanément : ces préférences communes seront la base de recommandations ultérieures, qui se baseront sur l'ensemble de contenus possédés par l'un ou l'autre pair. La discussion s'engage dans le groupe de développeurs sur les implémentations techniques qui pourront mieux servir les trois buts déclarés de Tribler : trouver, profiter et partager des contenus.

Certaines caractéristiques doivent – selon Hjalmar, chef programmeur de Tribler – faciliter la recherche de contenus appropriés, à la fois dans Tribler et sur d'autres portails comme YouTube ou Liveleak, surtout en phase d'amorçage du dispositif, quand les contenus matériellement stockés dans l'ensemble des logiciels clients Tribler sont encore trop peu nombreux pour rendre des résultats de recherche satisfaisants. Une recherche par mots clés, plus classique, est donc intégrée à des

fonctionnalités, comme la navigation par catégories, qui vise à améliorer la capacité de l'utilisateur de trouver des contenus qui correspondent à ses préférences<sup>427</sup>.

Un dispositif rudimentaire de réseautage social est également prévu : les liens d'amitié, l'entrée en contact avec des usagers partageant les mêmes goûts et la possibilité de la part de l'utilisateur de recommander de manière proactive, à son tour, des contenus, deviennent des instruments de recherche et de découverte de contenus potentiellement intéressants. Une convergence entre types de services – partage, recherche et réseautage – se profile dans les négociations des choix techniques des développeurs, sous la bannière du « *social-based P2P* ».

Enfin, dans le choix des composantes de la deuxième version de Tribler, se profile quelque chose qui, même si proposé de manière presque accessoire, aura par la suite une grande importance pour le succès du dispositif. Une note de Friedrich Bierbach à son équipe, début 2006, souligne :

*« There should be some kind of integrated video and audio player that makes it possible to almost immediately start watching or listening to stuff. BitTorrent allows us to do that. We should exploit it. You don't get any more social than that<sup>428</sup>. »*

On voit ici la première trace de ce qui deviendra, par la suite, la caractéristique principale de Tribler – la mise en place d'un flux de données en streaming associé au compte de chaque utilisateur, qui donne la possibilité à celui-ci de commencer à regarder ou écouter en direct, dès le début du téléchargement, les contenus partagés. Si cela permet le visionnage en streaming dès le programme choisi, Friedrich voit aussi bientôt, dans cette immédiateté, l'une des fonctionnalités sociales du dispositif. En effet, il s'agit aussi, selon lui, d'une « mise en valeur » des goûts de chaque utilisateur vis-à-vis des autres – puisque les préférences de chacun s'explicitent depuis le début de l'action de partage<sup>429</sup>.

C'est l'établissement d'une affinité immédiate, à la fois entre usagers et entre l'utilisateur et les contenus, qui est donc renforcé par le son et par l'image et se reflète au niveau technique dans la capacité de BitTorrent à découper et distribuer rapidement des segments de données aux pairs. La dimension collective des téléchargements permise par BitTorrent d'un côté, et d'un autre côté, l'infrastructure de distribution P2P, construite par les ressources fournies par les usagers, sont dans la vision de Friedrich et de son équipe les deux composantes d'une couche inférieure qui, même invisible, est déjà « sociale<sup>430</sup> ».

En même temps, plusieurs commentateurs techniques et usagers pionniers voient dans

---

<sup>427</sup> Entretien du 31 mars 2010 avec Hjalmar, chef programmeur de Tribler.

<sup>428</sup> Notes de travail de Tribler, 2006, que Friedrich Bierbach me montre lors de mon séjour à Delft.

<sup>429</sup> Entretien avec F. Bierbach.

<sup>430</sup> Id.

la décentralisation du streaming la caractéristique qui pourrait constituer le véritable potentiel novateur de Tribler par rapport à d'autres systèmes de diffusion et échange basés sur BitTorrent. Les inconvénients qui pourraient résulter d'un manque de « masse critique » d'utilisateurs – peu de torrents, téléchargements plus lents – sont par ailleurs soulignés :

*« Tribler as a BitTorrent client isn't that flash. It doesn't have anywhere near the number of features of µTorrent or Vuze [...] What Tribler does is provide an excellent demonstration of the future of BitTorrent, where torrents are truly decentralised and don't rely on web sites to post files or trackers to manage downloads. While your favourite BitTorrent sites could be closed down, filtered or go offline, Tribler will always continue to work and we expect that soon µTorrent and Vuze will follow it down that path<sup>431</sup>. »*

*« At times it happens that while downloading a torrent using BitTorrent client, the main servers suddenly stop responding causing user to wait until the servers get back online. Since downloading of torrents are taking place through centralized servers, it is always on the cards that servers hiccups would move the torrent data you're sharing with other. [It is t]o address this [that] Tribler has been created [and its] vision of capitalizing on only P2P sharing is great<sup>432</sup>. »*

#### PAIRS OU « BUDDIES » : DES SOLUTIONS SOCIALES POUR DES DÉFIS TECHNIQUES

*« That was in the very beginning. We're kind of lazy in updating the website, »* rit Friedrich en indiquant la page du site Tribler qui présente la philosophie qui sous-tend le dispositif<sup>433</sup>. En effet, le début du texte que j'ai à ce moment sous les yeux et que j'avais déjà pu consulter avant ma visite à Delft me fait penser, non sans un sourire, à la Déclaration d'Indépendance du Cyberspace (Barlow, 1996) pour son style où le technique se mêle au politique, avec la proclamation d'un Internet « égalitaire » aux accents vaguement naïfs :

*...any person should be able to use a content distribution system to spread his or her content directly and efficiently to an audience of millions all over the world, and that P2P technology can be used to realize this vision.*

*The main idea of P2P technology, at the core of Tribler, is that users pool resources to create a service. Because users donate resources, the content distribution infrastructure is very cost-effective. P2P is currently used mainly for the illegal spreading of copyrighted files, but we believe that P2P is the future of content distribution and will mature from the current 'wild west' into a respectable business*

<sup>431</sup> <http://www.pcuser.com.au/downloads-apps/internet/2011/aug/find-files-with-tribler>

<sup>432</sup> <http://www.addictivetips.com/windows-tips/tribler-bittorrent-client-brings-decentralized-p2p-content-sharing/>

<sup>433</sup> Entretien avec F. Bierbach et page « Vision » du site Tribler, <http://tribler.org/trac/wiki/AboutTribler>

*solution. At least at the present stage, we focus on the technical problems of P2P and consider more traditional aspects of business models such as payments and advertisements outside of our scope. Rather, inspired by authors arguing that P2P will create new business opportunities, (we) aim at spreading specialized content (such as video recordings of local concerts, soccer matches, and city council meetings) to more or less small audiences, which will suddenly become feasible due to the low cost of using P2P over the Internet.*

*In Tribler's vision, software for video file sharing should have a basic understanding of, and be enhanced by the real-life functioning of, human friendships, user tastes in content, and Internet connectivity between users.*

Tableau 10. Extrait de la page du site Tribler expliquant la philosophie qui sous-tend le dispositif (<http://tribler.org/trac/wiki/AboutTribler>, page en l'état vue pour la dernière fois en mars 2011).

Si ce programme à l'allure de « Charte des Droits de l'internaute » a profondément marqué les débuts du dispositif, Friedrich reconnaît qu'il s'agissait en grande partie d'un positionnement stratégique par rapport aux financeurs potentiels du projet. Ceux-ci devaient à la fois pouvoir se démarquer de la vision dominante du P2P dans les arènes politiques, et en proposer une nouvelle, dans laquelle la technologie serait au service d'une sorte de « démocratisation » de la distribution de contenus, avec des modèles d'affaires « respectables ». L'équipe s'est pourtant concentrée, lors des premières phases du projet (les premières idées remontent au début de 2004<sup>434</sup>), sur les aspects de développement technique du dispositif – ces aspects qui sont mentionnés de manière presque secondaire dans la déclaration.

C'est peut-être la raison de l'amusement de l'équipe lorsque, à la fin de mon séjour à Delft et après m'être entretenue avec une bonne partie de ses membres, on discute de combien leur démarche centrée sur la technique a été « politique ». Parfois, cela a été en accord avec leur philosophie de départ, et parfois, leur démarche a fini par suivre d'autres parcours. Notamment, l'équipe a identifié nombre de défis de recherche et d'ingénierie pour le partage de contenus au moyen du P2P, arrivant à la conclusion qu'un recentrement sur les usagers et leurs interactions pouvait aider à résoudre ces défis.

Un trait marquant de cette expérience pour l'équipe Tribler a consisté à développer leur dispositif, comme on l'a dit, sur un système déjà existant, ABC. Hjalmar explique que cela a permis, pour certains aspects, de procéder plus rapidement – mais a ralenti la démarche pour d'autres, car les développeurs ont dû rechercher la meilleure manière de retravailler des couches de l'application – notamment les façons d'incorporer la partie « sociale » caractéristique de Tribler à ce qui existait déjà :

*« One of the problems was that we took an already-existing client (BitTorrent) and started modifying that. It helped speed up things, but it also delayed them,*

---

<sup>434</sup> Selon Hjalmar, un des « vétérans » du projet qui a été en charge de la coordination des programmeurs depuis lors.

*as we had to do maintenance and basically reshaped the entire thing to separate the P2P engine part from the GUI (Graphic User Interface). A challenge was also to find the optimal way to build things on top of what was already existing. Currently the Tribler client is a core with three or four applications running on it. There is the search feature, a stand-alone video player, a version of it that actually runs as plug-in in the browser<sup>435</sup>. »*

Le défi de recherche le plus important pour l'équipe s'avère bientôt être la décentralisation de la fonctionnalité du système P2P entre les différents pairs. De façon similaire à ce qu'on a pu voir pour les cas analysés dans les chapitres précédents, la recherche d'un compromis se met en place. Une décentralisation complète pourrait éliminer le besoin de points de passage obligés dans le réseau, qui pourraient éventuellement constituer des points de défaillance ou des menaces à la sécurité. Par ailleurs, il semble bientôt clair pour les programmeurs de l'équipe que les choix de programmation se concentrant sur la stabilité et la fiabilité des connexions au réseau, ainsi que sur la reconnaissance et la validation de l'identité des usagers (par conséquent, de leurs activités dans le système), nécessitent la mise en place de quelques éléments centralisés. La structure de BitTorrent elle-même, ainsi que, de manière plus générale, l'histoire des dispositifs hybrides P2P les plus répandus et efficaces, sont convoquées pour soutenir cette idée :

*« We ended up thinking, to date, no P2P file-sharing system exists which fully decentralizes all functionality efficiently and without a high risk of a loss of integrity. BitTorrent, the very protocol Tribler stems from, is not fully decentralized as it depends on web sites and trackers for finding content<sup>436</sup>. »*

En revanche, si les éléments « centralisants » du client BitTorrent sont préservés au bénéfice de l'efficacité du système, l'introduction de dynamiques d'interaction et de regroupement social est identifiée comme une manière de décentraliser l'application de manière « naturelle », en accord avec la manière même dont les personnes – les pairs – échangent de l'information dans des contextes où ils se connaissent :

*« Social groups became the guiding principle of our programming as a sort of natural method so as to efficiently decentralize P2P systems, because it reflects the way people, peers, who know each other tend to exchange information<sup>437</sup>. »*

Le deuxième problème rencontré par l'équipe est d'assurer la disponibilité du système P2P dans son ensemble. Les opérations du système ne devraient pas, selon Hjalmar, dépendre trop largement de la disponibilité d'aucun pair en particulier, ou d'aucune composante centrale, car la défaillance éventuelle de cette composante peut s'avérer perturbatrice pour le bon fonctionnement du service au point de le compromettre. Les périodes, pour la plupart très courtes, pendant lesquelles les pairs

---

<sup>435</sup> Entretien avec Hjalmar.

<sup>436</sup> Id.

<sup>437</sup> Id.

sont disponibles, mettent la question de la disponibilité (*availability*) au centre de la démarche de développement. Encore une fois, c'est une solution « par le social » qui est envisagée, en rendant visibles au niveau de l'interaction entre usagers des éléments, comme les « récompenses » pour les pairs vertueux, qui sont de toute manière déjà inscrits dans l'architecture technique pour des raisons d'optimisation de performance :

*« Proven social incentives such as rewards and social recognition could stimulate users to leave their P2P software running for longer periods, thus improving the overall availability of the network<sup>438</sup>. »*

En relation étroite avec la question de la disponibilité, les développeurs de Tribler abordent également la question de la confiance entre pairs, une dynamique qui est fondamentale pour l'intégrité du système. Le principe selon lequel, par définition, un système P2P utilise des ressources mises à disposition par ses usagers, ne devrait pas amener à négliger le fait que ce statut de « donneurs » ne les rend pas pour autant automatiquement dignes de la confiance des autres utilisateurs. Les développeurs Tribler s'appuient, encore une fois, sur l'histoire de dispositifs semblables pour parvenir à la conclusion que la préservation de l'intégrité du système a toujours été l'un des principaux problèmes des systèmes P2P opérationnels, étant donné que les attaques peuvent se situer à différents niveaux allant des marqueurs de contenus (*tags*) jusqu'aux métadonnées et aux contenus eux-mêmes. Introduire des dynamiques sociales dans le dispositif permet à cet égard de rendre la confiance un instrument de préservation de son intégrité, où les utilisateurs seraient les premiers à avoir un intérêt dans le comportement vertueux de leurs pairs, contrairement à un scénario où ils tireraient avantage d'un éventuel statut de « *free-rider* » :

*« This significant problem, often ignored by P2P system designers, can be solved with a social-based network, in which users actively help to clean or remove polluted data and users can select trustworthy representatives<sup>439</sup>. »*

La reconnaissance sociale est également identifiée comme une dynamique pouvant fournir les bases d'un mécanisme d'attribution de récompenses et de crédits, mécanisme fondamental pour favoriser la coopération et obtenir une bonne performance. Si celle-ci dépend de la capacité et de la volonté des pairs à fournir des ressources, les pairs peuvent, par ailleurs, décider de manière autonome combien de ressources ils vont contribuer au système. Rendre le mécanisme de motivation plus explicite aux yeux des usagers qui, au moyen de leurs ressources informatiques, participent au fonctionnement du système en tant que nœuds/pairs, peut être une façon efficace de réconcilier ces deux tendances.

Enfin, il s'agit pour l'équipe Tribler de s'attaquer à un autre défi commun à beaucoup de systèmes P2P – la gestion des éléments qui contribuent à sa volatilité d'un point de

---

<sup>438</sup> Entretien du 30 mars 2010 avec Andreas, programmeur dans l'équipe Tribler.

<sup>439</sup> Id.

vue technique : les adresses IP dynamiques, la *Network Address Translation*, et la présence des pare-feu dans les ordinateurs de plusieurs pairs. *De facto*, en arrivent à conclure Hjalmar et son équipe, les changements fondamentaux de l'Internet entraînés par l'usage répandu de ces trois dispositifs techniques<sup>440</sup> – et souvent permis par défaut par les systèmes d'exploitation installés sur les machines des utilisateurs – amènent les pairs à ne plus avoir la liberté d'envoyer des données aux autres pairs sans que un ou plusieurs autres pairs aient un rôle d'intermédiaire. Les développeurs Tribler en sont amenés à prendre le parti d'accepter la présence dans le système de ces intermédiaires, comme résultat de choix extérieurs concernant l'infrastructure de l'Internet dans son ensemble, et se concentrent plutôt sur la manière dont les choix de programmation qu'ils opèrent permettront à un usager de choisir des intermédiaires « de confiance » :

*« Social networks enable communicating peers to automatically select trusted mediators from the members of their social proximity that are online, thus eliminating the need for fixed mediators<sup>441</sup>. »*

Par ces déplacements, comme on a vu, l'équipe Tribler cherche à répondre par la mise-en-dispositif de dynamiques « sociales » afin de s'attaquer à certains des problèmes et fragilités classiques du P2P. Cet ensemble de choix socio-techniques est présent et prend forme dans les « couches inférieures » de l'algorithme qui va constituer la base du logiciel Tribler depuis le lancement de la première version publique, Tribler 3.3.4, au début de 2007 : Buddycast.

#### *BUDDYCAST*, OU COMMENT LE « SOCIAL » SE TRADUIT EN ALGORITHME

L'introduction de l'algorithme Buddycast, dont la caractéristique principale et distinctive se trouve dans le mécanisme de recommandation décentralisée<sup>442</sup>, marque définitivement leur prise de distance avec BitTorrent et, plus particulièrement, le client ABC qui a servi de base au développement de Tribler : c'est le dispositif qui qualifie les pairs composant le système de « buddies », compagnons et amis, stabilisant l'ajout de la dimension « sociale » aux nœuds de réseau P2P que sont les clients Tribler.

Avec le façonnage et la mise en place de Buddycast, s'ouvre pour l'équipe Tribler une phase de réajustements progressifs du dispositif autour de différentes dynamiques d'inspiration « sociale » : plusieurs mécanismes sont testés et certains écartés, avec pour finalité ultime de renforcer un système possédant encore une allure de P2P « classique<sup>443</sup> ». Les développeurs mettent en place une stratégie à plusieurs volets :

---

<sup>440</sup> Voir Chapitre 3.

<sup>441</sup> Entretien, Andreas.

<sup>442</sup> Entretien, Hjalmar.

<sup>443</sup> D'après le site Tribler (<http://tribler.org/trac/wiki/AboutTribler>): "When the Tribler application program is started, it automatically starts searching other users that have Tribler running on their

regroupements d'utilisateurs sur la base de leurs préférences, réseautage social, stockage temporaire en local d'informations destinées à identifier les échangeurs et faciliter les échanges, attention à l'interface usager (*Graphic User Interface*, GUI), recommandation et téléchargement collaboratif.

La première dynamique que Hjalmar et son équipe s'attachent à intégrer dans Buddycast est le mécanisme selon lequel « l'affinité encourage la coopération », c'est à dire, que des goûts similaires et partagés par les utilisateurs pour certains types de contenus peuvent constituer les fondations d'une communauté globalement dotée d'un comportement altruiste – et donc, plus durable à moyen et le long terme, un des principaux problèmes du P2P étant justement l'accumulation de comportements égoïstes dus à un manque de motivation à se comporter de manière plus « vertueuse ». Afin d'implémenter des regroupements sociaux efficaces dans Tribler, les développeurs choisissent d'éliminer l'anonymisation des pairs – ce qui n'implique pas pour autant que l'utilisateur doit figurer au sein du réseau avec son véritable nom, mais que ce nom sera choisi une fois pour toute et demeurera stable au fil du temps pour le système et les autres utilisateurs, en tant que base sur laquelle construire la réputation, la fiabilité et les préférences :

*“(We needed to be able to) distinguish friend, foe – like polluters – and newcomers. For this, we wanted peers to be de-anonymized and social-group formation facilitated. De-anonymization is achieved by having every user choose a nickname; Tribler was designed to transfer user nicknames between users automatically and we added afterwards to Buddycast a social networking module, responsible for storing and providing information regarding social groups: the group members, their recently used IP numbers, and so on<sup>444</sup>.”*

En effet, comme on a pu voir, une des principales limites des systèmes de partage de fichiers auxquels l'équipe de développeurs Tribler doit faire face est la durée limitée, et non définie *a priori*, des sessions pendant lesquelles l'utilisateur est actif sur le réseau. Les informations « de contexte » ou de présence, c'est-à-dire les données en train d'être partagées par un nœud à un moment donné<sup>445</sup>, se perdent quand un utilisateur se déconnecte du réseau : ce qui ouvre la voie à un manque de stabilité du système. Cette déstabilisation est ultérieurement accentuée par les adresses IP dynamiques, qui rendent difficile de stocker les informations liées au contexte, pour chaque pair et au fil de différentes sessions de connexion.

---

*computer. When a connection is established it starts exchanging information. (...) The information about the discovered files and persons is available in the Tribler program. By browsing through the files and persons each user can find their preferred files and users. (...) When a user presses the download button, that user's computer makes an inventory of which computers actually have this file (or a part of it) and then downloads the parts from the different computers.”*

<sup>444</sup> Entretien, F. Bierbach.

<sup>445</sup> Voir aussi la deuxième section du chapitre 2 (*Echange et utilisation partagée d'information*).

L'une des possibilités qui s'ouvre aux développeurs est de stocker les informations de contexte dans des bases de données<sup>446</sup> (qui commencent bientôt à être appelées, au sein de l'équipe, les « *Megacaches*<sup>447</sup> »), ce qui faciliterait l'existence de groupes sociaux basés sur des mécanismes de confiance et de réputation. Cependant, pour que la mise en place de ces bases de données soit possible, une condition doit être remplie : il faut que les identités des utilisateurs au sein du réseau soient stables, même si leur présence ne l'est pas. Pour essayer de remédier à ce problème, l'équipe Tribler introduit dans l'algorithme Buddycast les « PermIDs » : il s'agit de codes qui doivent servir d'identifiants permanents et uniques sur le réseau, à échanger par e-mail entre utilisateurs. Mais l'augmentation du risque de phénomènes de *spoofing* (la falsification de l'identité appartenant à un PermID) qui s'ensuit complique cette démarche. Finalement, Hjalmar et ses collaborateurs aboutissent à une solution qui rappellera au lecteur le mécanisme d'encryptage et d'échange de clés déjà analysé dans le chapitre précédent pour le cas de Wuala :

*« We end up implementing this challenge-response mechanism for validating PermIDs. In this mechanism, a random number selected by the challenger is encrypted by the challengee with its private key and subsequently decrypted by the challenger with the corresponding public key—authentication succeeds when the result is identical to the original number<sup>448</sup>. »*

Enrichis de ce système de validation et de sécurisation, les PermIDs permettent de remédier au manque de « mémoire » persistante sur les activités précédentes d'un usager dans le réseau, limitation partagée par la quasi-totalité des systèmes de partage de fichiers en P2P. Les pairs échangent généralement des interrogations concernant les fichiers et les métadonnées concernant celles-ci, mais ignorent les autres types d'information ; par ailleurs, les informations de contexte qui doivent être sauvegardées afin d'améliorer la performance du réseau concernent les fréquences de contact entre les nœuds, les niveaux d' « altruisme » et de présence sur le réseau, et les similarités dans les goûts. Une fois stabilisée l'identité d'un pair au sein du réseau, il devient possible pour l'équipe Tribler de mettre à profit les différents types

---

<sup>446</sup> Id. Hjalmar précise qu'il y a quatre bases de données, servant des objectifs différents, et trois d'entre elles restent d'une dimension relativement réduite : *“The four Megacaches in Tribler are the Friends List with information on social networks, the Peer Cache with information on superpeers and peers in general, the Metadata Cache with file metadata, and the Preference Cache with the preference lists of other peers. The sizes of the Friends List, the (Super) Peer Cache, and the Preference Cache are below ten megabytes at any time, which for the Preference Cache is already enough to store the preference lists of thousands of taste buddies.”*

<sup>447</sup> Friedrich Bierbach fait remonter l'origine de cette appellation à l'arrivée dans le groupe, en septembre 2006, de Jingerzi, un doctorant et développeur qui deviendra par la suite le responsable principal des statistiques sur le comportement des usagers dans Tribler. Une mémoire cache (ou antémémoire) est, une mémoire qui enregistre temporairement des copies de données provenant d'une autre source de donnée, afin de diminuer le temps d'accès (en lecture ou en écriture) d'un matériel informatique (en général, un processeur) à ces données. La mémoire cache est plus rapide et plus proche du matériel informatique qui demande la donnée, mais plus petite que la mémoire pour laquelle elle sert d'intermédiaire.

<sup>448</sup> Entretien, Hjalmar.

d'informations de contexte reçues par chaque pair, et les stocker en local dans les *Megacaches*. Ces informations seront ensuite échangées à l'intérieur d'un groupe, et éventuellement entre groupes, à l'aide d'un protocole épidémique<sup>449</sup> intégré dans Buddycast.

Reste cependant un problème lors de l'utilisation des *Megacaches*, qui est difficile pour l'équipe : ces mémoires doivent être constamment mises à jour, ce qui nécessite de la mise à disposition de ressources de mémoire et de réseau importantes (« *The main problem concerning the Megacaches is the overhead traffic required to keep them up-to-date*<sup>450</sup>. ») Cependant, le choix d'allouer une grande quantité de ressources à ces opérations n'est à partir de ce moment-là plus mis en question : la stabilisation des identités des nœuds-usagers, comme pré-condition de l'enrichissement des informations nécessaires à la durabilité du système, est aux yeux de l'équipe un but suffisamment important pour que cette forme de gestion du réseau soit justifiée.

L'interface usager est également vue comme un élément important du dispositif, afin de rendre « utilisable » le réseau *social-based*. Une première version de l'interface comporte cinq composantes : *Map*, *Downloads*, *My Friends*, *Files I Like*, et *Taste Buddies*, qui se réfèrent respectivement à la localisation d'un usager, à ses téléchargements en cours, à ses amis, aux contenus qu'il préfère, et aux autres usagers qui partagent les mêmes préférences. Les développeurs discutent par ailleurs, pendant longtemps, de l'aspect que l'interface devra prendre afin de faciliter la formation des « groupes » qui sont à la base de la valeur ajoutée de la couche sociale de Tribler. Hjalmar propose que le module « My Friends » montre plus clairement non seulement les amis, mais aussi le réseau d'amis-des-amis, ainsi que les *taste buddies*, avec l'idée que cette proximité visuelle permet à l'utilisateur d'avoir un contact plus personnel avec ses pairs, et peut aider à diminuer les comportements antisociaux (et non durables). Friedrich est de son côté particulièrement concerné par la manière dont le design de l'interface peut rendre plus facile pour l'utilisateur l'identification visuelle de collaborateurs potentiels. Il propose donc d'introduire un mécanisme selon lequel quand un usager est en train de télécharger un fichier, les adresses IP des pairs à partir desquels le téléchargement a lieu soient géo-localisés et rendus visibles à l'utilisateur au moyen du module *Map* ; un module de géo-localisation est donc ajouté au dispositif, construit sur la base d'un moteur *open source*<sup>451</sup>. Enfin, Andreas propose que l'interface usager comporte des fonctionnalités destinées à faciliter l'usage du

---

<sup>449</sup> Un protocole épidémique (ou gossip-based) est caractérisé par des échanges de messages périodiques entre les nœuds du système. Il est fondé sur la connaissance locale, et partielle, du réseau pour chacune des entités le composant. Un mécanisme d'échange périodique de ces visions locales conduit à la création de systèmes exhibant des propriétés globales auto-organisantes (routage, maintenance, surveillance, communication complexe, partage de ressources). Ce type de protocole s'adapte dès lors très bien aux besoins des systèmes repartis à large échelle ; il est utilisé, par exemple, dans Gossple, le moteur de recherche P2P collaboratif dont je parle dans le quatrième chapitre. Voir [Rivière, 2008] et [Gramoli, Le Merrer & Kermarrec, n.d.], [Bertier, Busnel & Kermarrec, 2010].

<sup>450</sup> Entretien, Hjalmar.

<sup>451</sup> Il s'agit de <http://www.hostip.info/>

module pour le téléchargement collaboratif : l'interface devra permettre à l'utilisateur de suivre quels amis l'ont déjà aidé dans le passé, à quels amis il a déjà permis d'utiliser part de sa bande passante, et quels amis actuellement en ligne peuvent l'aider à accélérer ses nouveaux téléchargements.

Buddycast doit également comporter des éléments propices à l'amélioration de l'efficacité du dispositif dans la phase d'amorçage<sup>452</sup>. Hjalmar estime en effet que cette phase n'est pas gérée de manière optimale dans BitTorrent :

*“In BitTorrent, peers have to repeatedly connect to a tracker in order to discover other peers; and more, the original BitTorrent protocol restricts communication to within the swarms of peers that download the same files, making the bootstrapping process unnecessarily repetitive.”*

Pour résoudre le problème de l'amorçage dans Tribler, Jingerzi, un doctorant qui a récemment rejoint l'équipe, propose que soit mis en place un mécanisme permettant à un pair de contacter automatiquement, au moyen d'un message Buddycast, un « super-pair » parmi un ensemble déjà connu, une première fois et immédiatement après l'installation, afin d'obtenir une liste initiale d'autres pairs dans le système et pouvoir tout de suite commencer à participer à la diffusion épidémique de l'information dans le réseau Tribler.

Localiser et rendre disponibles des contenus est un des points critiques pour tout système P2P. L'équipe identifie bientôt le cœur de l'algorithme Buddycast dans le mécanisme de découverte et de recommandation des contenus, qui doit refléter l'approche et la spécificité de Tribler : celle-ci consiste à connecter tout d'abord des usagers et leurs préférences similaires, plutôt que donner la priorité aux fichiers et aux contenus en soi<sup>453</sup>. Un protocole épidémique est donc développé afin d'échanger les listes de préférences des usagers et découvrir, de manière efficace, ses *taste buddies*. Les listes de préférence, afin de pouvoir amorcer le mécanisme, se remplissent au début, par défaut, avec les premiers téléchargements faits par les usagers, mais elles s'affinent au fil du temps, au fur et à mesure que le dispositif est utilisé et que les contenus « préférés » par certains groupes d'utilisateurs commencent à émerger. Buddycast prend bientôt la forme d'un algorithme collaboratif, qui filtre les résultats moins pertinents en comparant périodiquement les caractéristiques de téléchargement des différents usagers :

*“Tribler's architecture has a module we call Peer Similarity Evaluator, that is destined to compare preference lists and determine the similarity in taste of two peers. The Recommendation Engine module, that is indeed the heart of the system, is able to compile a list of files a user most likely wants: recommendations are made using a collaborative filtering algorithm. This*

---

<sup>452</sup> Ou *bootstrapping*: la recherche des autres pairs au sein du système P2P après que l'utilisateur a installé le client pour la première fois et doit donc se “former un réseau”: cfr. les deux chapitres précédents.

<sup>453</sup> Entretien, Hjalmar.

*algorithm will compare users' download history to that of other Tribler users they are connected to. If a user has content in his download history that another connected user has not downloaded, it will recommend it to the first user. The recommendation value assigned to a content item depends on how similar the first user's download history is to the second. So the higher the value, the more the second user is predicted to like it. In a nutshell, that's how it works: constant, implicit content comparisons<sup>454</sup>.*”

## SOCIAL ET FAIR-PLAY : LA BANDE PASSANTE COMME DEVISE

Le mécanisme de collaboration à l'œuvre dans Buddycast s'étend dès lors à la fonctionnalité de téléchargement collaboratif, 2Fast : l'idée sous-tendant ce dispositif est la possibilité pour un utilisateur d'obtenir une augmentation sensible de la rapidité de ses téléchargements, en exploitant la capacité de téléchargement non utilisée par ses amis qui sont en ligne à un moment donné. Les ensembles d'utilisateurs qui se sont formés par regroupement autour de préférences communes constituent l'arène où les membres du système trouvent d'autres membres dans lesquels ils ont confiance, et qui acceptent de collaborer à l'amélioration des performances de téléchargement. Hjalmar et son équipe développent tout d'abord 2Fast de manière à ce qu'il soit possible pour des pairs « altruistes » de joindre un groupe et de contribuer à ses activités par le « prêt » de leur bande passante inutilisée, même s'ils ne sont pas intéressés par les contenus distribués dans ce groupe à ce moment-là. Mais les premiers tests ne donnent pas des résultats satisfaisants : le fait d'avoir, en phase de développement, assumé un niveau d'altruisme suffisant pour le bon fonctionnement du réseau ne trouve pas de contrepartie dans l'environnement « réel » des pairs Tribler. Hjalmar décide donc d'introduire un système de motivations/avantages (*incentives*) à l'intérieur des groupes :

*« Peers from a social group that decide to participate in a collaborative download take one of two roles: they are either collectors or helpers. A collector is the peer that is interested in obtaining a complete copy of a particular file, and a helper is a peer that is recruited by a collector to assist in downloading that file. Both the collector and the helpers start downloading the file using the classical BitTorrent tit-for-tat and the collaborative download extensions. As in BitTorrent, a helper selects a chunk of the file for downloading based on the rarest-first policy among the chunks in possession of its bartering partners. However, before actually downloading this chunk, it asks the collector for approval, which will only be granted when the chunk is unique, that is, when no other helper already downloads the same chunk. After downloading a file chunk, the helper sends the chunk to the collector without requesting anything in return. In addition to receiving file chunks from its helpers, the collector also optimizes its download performance by dynamically selecting the best available data source from the set of helpers and other peers in the BitTorrent network using the default mechanisms of*

---

<sup>454</sup> Entretien, Hjalmar.

*BitTorrent, which prefer peers that upload at higher rates. As helpers give priority to collector requests, they are preferred as data sources<sup>455</sup>. »*

Avec 2Fast, l'équipe Tribler met en place ce qui peut être défini comme un système de « bande passante comme devise » (*bandwidth as currency*), afin d'encourager et faire respecter un code de bonnes pratiques comme pré-condition à la stabilité du système. Cela semble réussir : les partageurs altruistes y gagnent une meilleure vitesse de téléchargement, soit vers l'amont soit vers l'aval, tandis que les *leechers* (sangsues), ceux qui cherchent à profiter du système plus qu'ils n'y contribuent, sont pénalisés.

Autour de ce code de bonnes pratiques, centré sur la bande passante comme monnaie d'échange, prend forme la collaboration de l'équipe de Delft avec l'école d'ingénierie de l'université de Harvard : c'est à ce moment que les chemins de Tribler et de la vidéo *via* l'Internet se croisent pour la première fois<sup>456</sup>. David Parkes, professeur associé à Harvard, est en effet très intéressé par le développement d'un système qui calcule de manière fiable la bande passante nécessaire à l'ensemble des téléchargements passant par un dispositif P2P. Ayant l'idée que cela pourrait devenir par la suite un « *model for safe and legal electronic commerce that uses Internet bandwidth as a global currency*<sup>457</sup> », Parkes suggère de travailler en premier lieu sur le partage et le téléchargement de vidéos :

*« Using bandwidth as a currency can remove some of the problems seen in file-sharing systems such as BitTorrent. (...) In peer-to-peer, I can build up credit by offering upload capacity and then use the credit for download in the future. There is still a balance, but the balance is on the order of days rather than seconds and this time-shifting can be welfare enhancing<sup>458</sup>. »*

Le prototype du système d'allocation de bande passante qui en résulte, utilisant les « minutes TV » comme monnaie d'échange, marque le premier déplacement de Tribler vers une spécialisation dans le partage et le streaming vidéo. L'introduction de la série des versions 4.x, à la fin de 2007, reflète ce changement<sup>459</sup>. À côté d'améliorations de l'interface usager et d'une articulation plus complexe des fonctionnalités « sociales » que je viens d'analyser, les utilisateurs peuvent désormais commencer à lire en streaming les fichiers vidéo quelques secondes après que le partage commence, au lieu de télécharger les fichiers dans leur totalité (Image 11).

---

<sup>455</sup> Entretien, Hjalmar.

<sup>456</sup> <http://tv.seas.harvard.edu/>

<sup>457</sup> <http://tv.seas.harvard.edu/press.php>

<sup>458</sup> Entretien de Colin Barras avec D. Parkes pour BBC, cité en "File sharers forced to play fair", <http://news.bbc.co.uk/2/hi/6971904.stm>

<sup>459</sup> Blog "Serializable Thoughts" (26 août 2011). "History of Tribler" (<http://blog.shinnoir.nl/2011/08/history-of-tribler.html>).

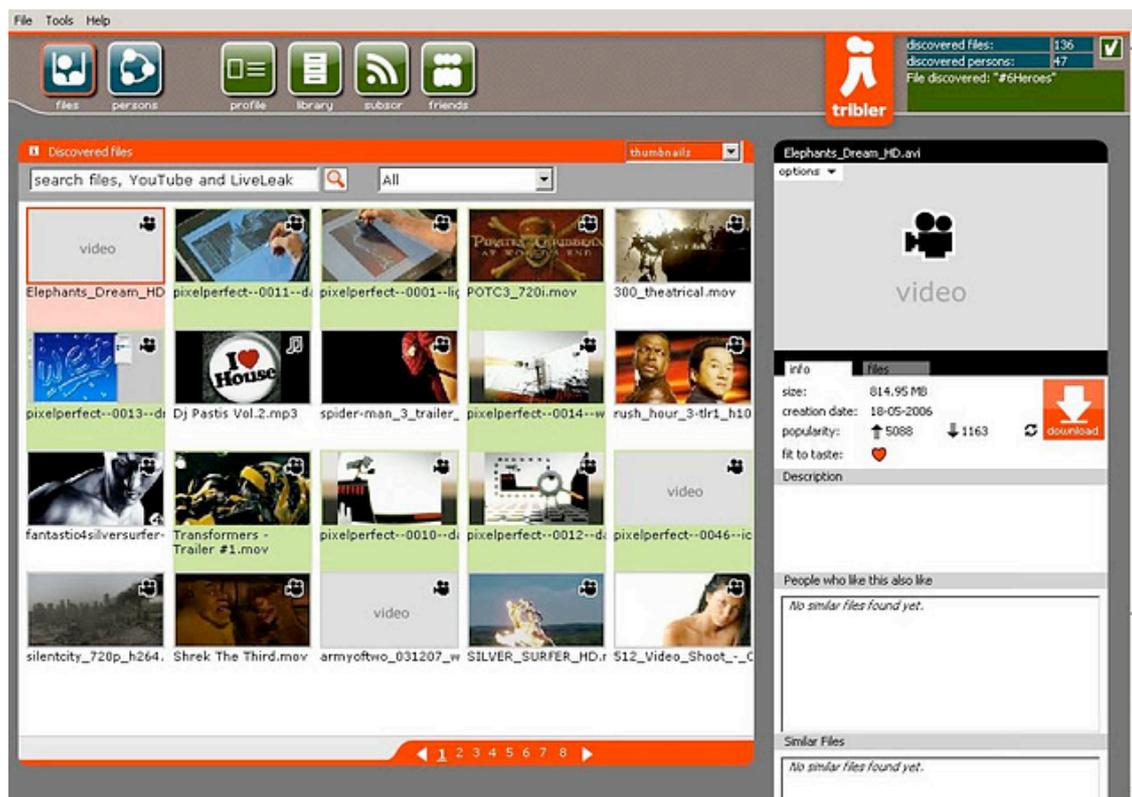


Image 11. Interface usager de Tribler, série 4.x.

La série 4 fera aussi le premier lien, bien que de façon très temporaire, avec les grands acteurs du streaming, en permettant aux usagers de visionner des contenus provenant d'autres sources, comme YouTube<sup>460</sup>. Tribler est publié sur Tribbox – une plateforme *open source* développée par AHT, un fabricant hollandais de décodeurs pour télévision<sup>461</sup>. L'Union européenne de radio-télévision manifeste, elle aussi, de l'intérêt pour l'architecture Tribler comme base possible pour la standardisation d'un système de diffusion européen pour la télévision via Internet<sup>462</sup>.

Tous ces facteurs font que la deuxième grande phase de développement de Tribler – pendant laquelle le produit commence à être reconnu et repris par un marché, bien que demeurant à plus d'un titre un travail de recherche universitaire – est marquée par l'élimination, au moins au niveau de l'interface, d'une grande partie de ces dynamiques « sociales » qui avaient marqué les premiers pas du logiciel, et par une concentration des efforts de développement et de « marketing » sur ce qui semble susciter le plus d'intérêt des partenaires commerciaux et de recherche possibles. Le suivi et l'analyse de cette phase font l'objet de la deuxième section.

<sup>460</sup> Cette fonctionnalité sera de courte durée et sera déjà supprimée avec l'introduction de la série 5.x, les développeurs étant surtout intéressés à la rencontre entre le streaming et le P2P.

<sup>461</sup> Voir l'introduction de ce chapitre.

<sup>462</sup> <http://www.digitaltrends.com/computing/tribler-brings-fairness-to-p2p/>

## SECTION 2. « OUR PRODUCT OUT THERE » : DU SOCIAL À LA VIDÉO À LA DEMANDE

La série 5.x de Tribler, dont la première version est lancée au début de 2009, est donc marquée par d'importants changements dans le rapport entre l'interface et l'architecture du dispositif. Au cours de 2008, l'interface de la série 4.x a commencé à présenter une série de problèmes aux yeux des développeurs, liés à sa « lourdeur » :

*« For Tribler 5.0, a minimalistic look was chosen because the old GUI of the 4.x series was deemed to be clunky. Just an example, in 4.x, when you searched for files, the results were presented in a grid with thumbnails. This approach, however, had problems like most results not having a thumbnail at all, so it was dropped in 5.x and instead the search results were displayed in a list. A number of these things gave the programme a clunky look and we could no longer afford that. After the Harvard collaboration, we now had our product out there<sup>463</sup>. »*

C'est donc à la première « économisation » du logiciel<sup>464</sup>, au fait que les développeurs le considèrent désormais un « produit », que semble être dû en tout premier lieu cet important choix de développement et de rapport à l'utilisateur : tout en restant *open source*, et en gardant son identité de projet universitaire (ce qui, comme on le verra, aura aussi un poids dans la décision de changement), Tribler est désormais à plus d'un titre un « *product out there* », un produit, saisi par le marché, intéressant un public spécifique dont les souhaits semblent être une plus grande attention à la fonctionnalité de streaming et à la simplicité.

Ainsi, si nombre de fonctionnalités telles que, par exemple, le téléchargement collaboratif, continuent de faire partie du système comme mécanisme techniquement efficace pour assurer la stabilité du système P2P, elles ne sont désormais plus destinées à être visibles par l'utilisateur, à moins que celui-ci ne le décide spécifiquement. Le paradigme « *social-based* » n'est pas abandonné, mais n'est plus présenté à l'utilisateur comme la caractéristique qui différencie Tribler, dans un panorama très chargé, des autres systèmes de partage de fichiers : les développeurs se proposent désormais de le reléguer aux seules « couches inférieures » du dispositif.

En parallèle, les développeurs choisissent de concentrer l'attention de l'usager sur la fonctionnalité qui a semblé attirer le plus d'attention de la part de potentiels financeurs et collaborateurs du projet : le service de vidéo à la demande et de streaming vidéo. Les usagers qui démarrent Tribler 5.0 sont accueillis par une interface très simplifiée, voire épurée, ne montrant qu'un espace de recherche, la mise à jour de la « réputation de partageur » de l'usager, une fenêtre sur la droite destinée à la lecture en streaming des fichiers vidéo échangés (en quasi-direct lors du partage) et

---

<sup>463</sup> Entretien, F. Bierbach.

<sup>464</sup> Qui se rapproche de la description faite par Alexandre Mallard de la commercialisation comme troisième étape clé de l'entrée en usage d'un dispositif, avec l'innovation et l'appropriation [Mallard, 2011].

deux liens pour changer ses paramètres et visionner la liste de ses téléchargements (Image 12).

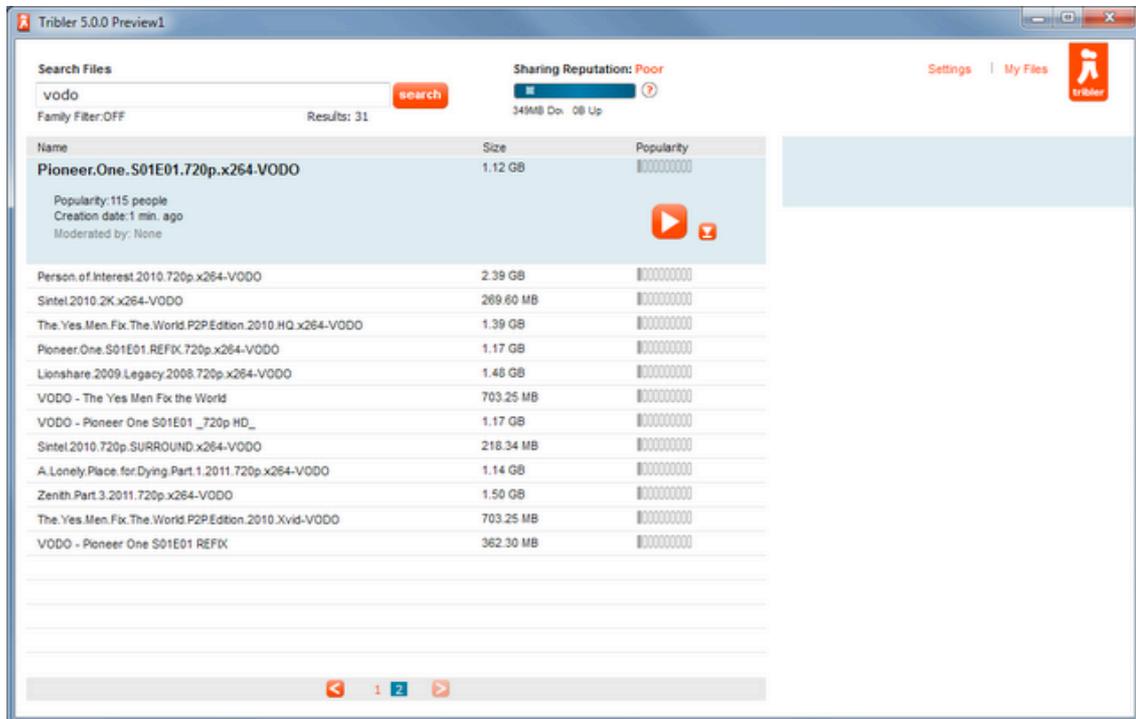


Image 12. Interface de Tribler 5.0 du point de vue de l'utilisateur (2009), montrant les résultats d'une recherche.

#### UNE INTERFACE ÉPURÉE, UNE ARCHITECTURE QUI RESTE « SOCIALE »

Quand, lors de mon séjour à Delft en mars 2010, je m'apprête à discuter avec l'un des programmeurs senior de l'équipe Tribler, Andreas, je n'ai que très peu d'éléments en ma possession pour essayer de reconstruire les changements intervenus dans Tribler au tournant de 2009 : tout ce que je sais, c'est que, partant d'un papier universitaire cité près de trois cents fois (Pouwelse et al., 2006) et proposant nombre de fonctionnalités « social-based », la démarche de développement s'est tournée vers le logiciel on-ne-peut-plus-sobre que j'ai récemment téléchargé. C'est donc avec une certaine curiosité d'utilisateur, avant que de sociologue, que je m'enquiers des causes qui ont donné lieu au choix d'abandonner un outil auparavant explicitement centré sur des dynamiques « sociales » pour revenir à cette interface élémentaire de la série 5. Cette démarche est-elle allée jusqu'à l'« épuration » de l'architecture, et si tel n'est pas le cas, quelle est maintenant la relation entre interface et architecture?

L'abandon des fonctionnalités sociales au niveau de l'interface est, selon Andreas – qui se présente tout de suite comme l'un des principaux responsables de ce choix – d'abord liée à la question de la composition « à géométrie variable » de l'équipe

Tribler, qui est et est restée depuis ses débuts un projet universitaire. Les changements dans les priorités de développement sont donc dus à la durée limitée des financements et des contrats attribués aux étudiants de Master et de doctorat qui travaillent au projet, ainsi qu'à leurs intérêts et priorités de recherche. Lors des phases initiales du projet, ces intérêts étaient majoritairement concentrés sur les aspects sociaux – sur la manière dont ceux-ci pouvaient être à la fois implémentés au niveau de l'architecture et rendus visibles à l'utilisateur au niveau de l'interface :

*« At that moment, there were a lot of people in the team that were very interested in the socially-enhanced part of the work, like epidemic protocols, algorithms based on affinities, etc. So they were aiming, not only within the client but on the interface as well, to represent peers as actual persons, they put images and names, and sent those things around the network<sup>465</sup>. »*

Lors du remodelage de l'interface à l'occasion de la sortie de la série 5.x, la plupart des personnes intéressées par les aspects sociaux a quitté l'équipe, et un principe guide le nouveau design : la volonté de rendre les choses plus simples pour l'utilisateur. En effet, les retours des premiers usagers indiquent qu'un nombre relativement faible d'entre eux se déclarent intéressés par l'utilisation de Tribler pour des raisons spécifiquement liées aux fonctionnalités sociales. Certains des développeurs de Tribler, y compris Andreas, font dériver de ces retours un choix qui restera controversé au sein de l'équipe : rendre Tribler plus facilement accessible et utilisable pour les utilisateurs moins techniquement compétents, et cela au détriment d'un outil plus articulé et plus complexe dans lequel l'apport à la fois du P2P et des dynamiques sociales sont explicitement visibles pour l'utilisateur.

Nombre de changements dans l'interface répondent donc à cette volonté, comme par exemple le fait de donner une certaine prééminence à la partie de l'écran destinée à la recherche de fichiers, explicitement inspirée à un « *very popular search engine*<sup>466</sup> ». Mais surtout, le but principal de la démarche est de placer la fonctionnalité de streaming vidéo au centre du dispositif, et de rendre plus simple son utilisation – ce qui se traduit *de facto* en une dépersonnalisation, en induisant une série de déplacements de vocabulaire aussi bien que visuels. Les pairs sont représentés par des chiffres plutôt que par des personnes ; les contenus sont représentés par une liste plutôt que par des icônes.

*« At the core was the idea of privileging the video on demand aspect and making clear that the video was there, ready to be played, and you could just click on it and here it is, ready to go, without needing first to find people, or make friends, so all the social concepts, we felt that they reached too far for an average user. We didn't do any research on this, maybe we should have, but we felt, just the most simple design was what most people actually wanted. That's what happens now with the current version of Tribler: people go, they tape, they click, they get a list back, they select from what's most popular, and the way they figure out what is most popular is because they see a number, not*

---

<sup>465</sup> Entretien, Andreas.

<sup>466</sup> Une claire référence à Google. Entretien, F. Bierbach.

*a list of persons, they join swarms and download it and that's it. All the social aspects can be very powerful, but at some point we felt it was too much at the moment. »*

La notion de réseau pair-à-pair elle-même passe au deuxième plan par rapport à celle de lecteur vidéo. Ce dernier aspect a, bien sûr, des implications en termes de différenciation du « produit » Tribler par rapport à ses concurrents. Si pendant les premières deux années d'existence de Tribler ses développeurs avaient attribué une grande importance à la spécificité qui dérivait de l'union entre le P2P et des fonctionnalités de visionnage vidéo en streaming, le message que la nouvelle interface veut communiquer à l'utilisateur est tout simplement que s'il souhaite voir une vidéo, il a en Tribler un outil simple et rapide pour atteindre son but.

*“(... Also,) we wanted to see what it would take to target those who have no idea what P2P was and who didn't care, we didn't reference P2P at all, stopped doing that, and we just 'sold' Tribler as a video player that delivered content over the Internet. What they want is to see a movie. Then, knowing that they are seeing it over the Internet, well that is detailed enough. If they have to install a programme, that's more than enough, difficult enough for most people. They don't need to understand what peer to peer is<sup>467</sup>.”*

Le paradigme « social » qui a informé le développement de Tribler ne disparaît pas pour autant. Caché aux yeux des utilisateurs, Buddycast continue à constituer l'ossature du système et à imprimer le « P2P social » aux interactions qui y ont lieu. Au fil des différentes versions dans la série 5.x, la visibilité des fonctionnalités sociales devient *opt-in* : le but principal du nettoyage de l'interface étant la simplicité, Hjalmar décide que pour les utilisateurs qui verraient dans les aspects sociaux une valeur ajoutée, ceux-ci peuvent être réintroduits sur une base volontaire au niveau de l'interface.

*“The social-based part never actually disappeared. Personally, I was among the 'socially-interested' Andreas talks about, I was from the very beginning, and I thought it was unfortunate that the interface part was disappearing, especially as beneath the new interface, the code is still there, we are just not displaying users and images anymore. So, now there is always the possibility to put the features back in place. I mean, even now, if you edit the profile, you could still upload the picture and so on. It's still more or less there, just less frequently used<sup>468</sup>.”*

Les utilisateurs qui discutent de l'apparition et de la disparition du « social » au fil des différentes versions de Tribler semblent, de leur côté, y relier des questions de *privacy* et les placer au centre de leurs préoccupations. Si Tribler est avant tout une variation de BitTorrent, le mécanisme de recommandation et de regroupement affinitaire n'a pas de place dans le logiciel, soutiennent certains d'entre eux. En quoi l'expérience sociale peut-elle améliorer ce qui demeure la tâche principale de ces

---

<sup>467</sup> Entretien, Andreas.

<sup>468</sup> Entretien, Hjalmar.

types d'outils : repérer et coordonner le téléchargement des *torrents* ? Si l'amélioration de l'efficacité technique du système grâce au social ne leur paraît pas évidente, certains utilisateurs sont d'autant plus perplexes qu'ils perçoivent les aspects de communication et de « sociabilité » comme un facteur d'exposition et de révélation de soi, dont ils ne veulent pas :

*« integrating 'human friendship' into a torrent software? not for me, at least not in a torrent software. there are instant messengers, forum, etc. for that. Who knows what I might be exposed to ? »*

*« Sorry, but the whole 'social sharing' thing is too much. As someone who has had trouble with ISPs getting all up in my business about file sharing, I prefer as much anonymity as possible in a torrent client. I don't need to add friends or even see any information about who I'm sharing with.<sup>469</sup> »*

Un autre utilisateur fait encore plus clairement le lien entre le mécanisme de recommandation basé sur les préférences en commun, et une possible détection des contenus stockés et échangés de la part de tiers. C'est l'occasion pour les développeurs de clarifier la politique de confidentialité du logiciel – quelles informations personnelles concernant l'utilisateur sont partagées et dans quel but. La logique à la base du système veut que l'histoire des téléchargements faits par un usager soit partagée avec les autres utilisateurs afin de « nourrir » l'algorithme de recommandation ; une quantité suffisante et fiable d'information relative aux préférences est une pré-condition au bon fonctionnement de l'algorithme, comme souligne un autre utilisateur. Dans la série 4, l'utilisateur pouvait choisir de ne pas utiliser cette option ; dans la série 5, il peut au contraire choisir de l'activer.

*« Another feature that is bugging me is that this software 'recommends torrent' to me. Is it monitoring on me? It just doesn't feel right somehow. »*

*« From our license: [Tribler] will by default exchange your download history with others. This feature can be disabled by disabling the recommender in the Preference menu. So all it does is send your history to others, we don't keep it on a server (we are a decentralized P2P app), and you can disable it. Anyway these are just features you only use if you want. »*

*« The recommendations don't work yet for me, I think I need more history and it needs more users contributing<sup>470</sup>. »*

La priorité des développeurs reste par ailleurs, à ce stade, l'amélioration de la qualité technique du système<sup>471</sup>. Face à des retours ambigus des utilisateurs – deux groupes, ceux qui s'expriment positivement au sujet de l'interface « élémentaire » et ceux qui y voient une perte de valeur pour le dispositif, coexistent<sup>472</sup> – Andreas décide de ne

---

<sup>469</sup> <http://ubuntuforums.org/archive/index.php/t-234700.html>

<sup>470</sup> Id.

<sup>471</sup> Entretien, Andreas.

<sup>472</sup> Entretien, Hjalmar: "We did get a few emails saying that they liked the new clean interface and

pas s'engager dans une véritable recherche de consensus auprès des usagers, pour en revenir à la première « raison d'être » du logiciel : celle de champ d'expérimentation des technologies P2P dans un contexte de projet universitaire et d'une équipe de développeurs en évolution constante.

*« We were (still) mostly interested in the technical aspects. Had we had at the time students who, like a few years ago, were interested in the more social aspects of it, it would have been nice to know more about it, we would have had more ground to base our decision on. But as it is, it was just an educated guess, with the word 'educated' loosely used, more like getting the flavour to identify the direction the project should take in order to get more users, which of course is very important. But what remained more important was that all the tests we made gave us more credibility while we were writing the papers we then presented to the academic community<sup>473</sup>. »*

Et Hjalmar renchérit :

*« All the protocol-level things are still there. Indeed, all the stuff is still there, you still exchange nicknames, it still works according to the preferences/affinities based protocol. The Buddycast protocol is still there, being improved upon, and works according to the principle of finding people you have the most similar tastes with<sup>474</sup>. »*

Le paradigme « social-based » des origines – la démarche de développement consistant à unir les deux approches de pair à pair et de réseautage social comme une façon d'améliorer et d'enrichir l'expérience utilisateur – finit donc par se stabiliser comme un des principes fondateurs du logiciel qui, tout en restant important, se fait discret. L'architecture reste « sociale » derrière une interface épurée, et l'utilisateur peut choisir de la rendre tangible et visible, s'il le souhaite. Les développeurs doivent par ailleurs faire face, au delà de leurs propres priorités de recherche et d'amélioration de l'outil technique, aux objections et aux perplexités de ces utilisateurs qui voient dans la mise à disposition d'informations « personnelles » (profils, préférences, interactions) un risque pour la préservation de leur vie privée plutôt qu'une possibilité d'enrichissement de leur expérience de partage.

## LA VIDÉO À LA DEMANDE DÉCENTRALISÉE ET L'OUVERTURE DE PARTENARIATS R&D

En même temps que le « social » dans l'interaction architecture-interface du dispositif

---

*were happy with that, but we had just about the same amount of emails complaining that they didn't like it".*

<sup>473</sup> Entretien, Andreas.

<sup>474</sup> Entretien, Hjalmar.

se voit reconfiguré, la série 5.x marque aussi un pas ultérieur vers le développement d'un autre choix important pour le futur du système : le fait de centrer les efforts de développement, ainsi que l'attention de l'utilisateur, sur la fonctionnalité de streaming vidéo et de *video-on-demand*. La collaboration avec l'équipe d'économistes et informaticiens de Harvard avait fait en soi que l'équipe Tribler commence à se concentrer tout particulièrement, parmi les différentes fonctions du dispositif, au streaming vidéo en *peer-to-peer*. Par ailleurs, ce sont d'un côté la participation de Tribler à deux projets de recherche centrés sur la diffusion de contenus vidéos, et d'autre côté la démarche de l'Union européenne de radio-télévision (UER), qui souhaite sélectionner un partenaire technique pour le développement d'une télévision européenne via Internet<sup>475</sup>, qui stabilisent définitivement le rôle central de cette fonctionnalité pour le futur du dispositif.

Comme on a pu le voir, et comme cela a déjà été le cas pour les fonctionnalités « sociales », la première insertion du streaming vidéo dans le système n'était pas prévue *a priori* et ne devait pas servir à différencier le dispositif de ses concurrents. Elle découle, en fait, de l'intérêt de recherche principal d'un doctorant :

*« Actually, initially, Tribler did not have this VoD thing, it was added a few years ago by a PhD student at the time, he also added some live streaming features to Tribler, and I took that and worked on that to make it fit the application. But before that, Tribler was just that, a BitTorrent client and the interest was to add some social features to it<sup>476</sup>. »*

Par la suite, l'obtention de financements publics hollandais dans le cadre de deux projets, I-Share et P2P-Future<sup>477</sup>, fait faire un pas supplémentaire dans la direction du streaming. Ces deux projets, dont l'université technique de Delft est l'un des partenaires techniques, se fondent effectivement sur une même idée de base. D'un côté la majorité des contenus multimédia circulant sur Internet sont des vidéos « net-goinfres<sup>478</sup> », et de l'autre, une architecture centralisée s'adapte moins bien au passage à l'échelle résultant d'une augmentation importante des usagers au sein du système : le résultat est que plus il y a d'usagers, moins il y a de ressources à la

---

<sup>475</sup> [http://www.ebu.ch/en/union/news/2006/tcm\\_6-46781.php?display=EN#0](http://www.ebu.ch/en/union/news/2006/tcm_6-46781.php?display=EN#0)

<sup>476</sup> Entretien, Andreas.

<sup>477</sup> <http://www.cs.vu.nl/ishare/>: "The I-Share Project, funded by the Dutch national Freeband Programme (itself funded by the Ministry of Economic Affairs of The Netherlands), has investigated the sharing of resources in virtual communities consisting of heterogeneous and transient nodes and networks: (virtually bounded) groups of nodes willing to share resources and to help each other for the benefit of individual and system performance. As a particular case, the project has considered the sharing of resources for the processing – such as compression – of multimedia streams, as this faces the sharing nodes with significant challenges in terms of volume of data, processing requirements, and security. With Philips and the Universities of Eindhoven and Twente, the Tribler team has been one of the primary partners of the project."

<sup>478</sup> Etiquette récemment créée par le directeur général de la Fédération française des télécoms, Yves le Mouël, pour se référer à ces contenus numériques dont la transmission et diffusion nécessite une grande quantité de bande passante. Pour ses usages récents dans les débats français sur la neutralité du Net, voir [Schafer, Le Crosnier & Musiani, 2011].

disposition de chacun d'entre eux. Comme un très grand pourcentage de trafic sur l'Internet est du P2P, les deux projets font donc l'hypothèse que l'on peut essayer de mettre à profit ce trafic pour faciliter la tâche des serveurs qui distribuent des contenus multimédia, en exploitant le principe du pair à pair suivant lequel chaque usager supplémentaire qui intervient dans le réseau aide à la distribution du contenu. Tribler, qui réunissait déjà dans ses premières versions des éléments de pair à pair et de diffusion de contenus, devient donc un banc d'épreuve pour la transmission des contenus vidéo en particulier, dans le cadre d'I-Share et de P2P-Future<sup>479</sup>.

En parallèle, la nécessité de se différencier d'Abacast<sup>480</sup> et de Rawflow<sup>481</sup>, les deux autres technologies de diffusion de contenus vidéo via Internet prises en compte par l'UER dans le processus de sélection, amène les développeurs de Tribler à redéfinir la spécificité du logiciel en s'appuyant sur son architecture décentralisée ; ce qui les conduit à définir de manière plus précise les avantages et les difficultés rencontrés pour implémenter une fonction de streaming vidéo sur du pair à pair, qu'on y introduise du « social » ou pas<sup>482</sup>.

Tout d'abord, il est nécessaire de modifier le modèle de téléchargement des fragments proposé par BitTorrent dans sa version de base, car celle-ci prévoit la possibilité de télécharger les fragments dans un ordre aléatoire si cela optimise la performance technique. Avec la fonctionnalité de streaming au premier rang dans l'application, il devient très important de développer un système qui télécharge d'abord les premiers fragments – premiers d'un point de vue chronologique dans l'ensemble du fichier vidéo – afin que l'utilisateur puisse commencer à le visionner aussi tôt que possible au cours du téléchargement. Plus généralement,

*« Data structures, some of which were very slow, needed to be optimized for some specific uses, to fit the first-first model rather than the rarest-first model<sup>483</sup>. »*

Le deuxième défi de développement consiste à gérer, avec un modèle décentralisé, la durée *a priori* non déterminée des flux de streaming en direct ; c'est-à-dire que si la version classique de BitTorrent connaît depuis le début la durée et le poids des fichiers vidéo avec lesquels elle doit travailler, une séance de streaming en direct peut, théoriquement, se prolonger indéfiniment. Il faut donc assurer une certaine redondance du modèle décentralisé pour que celui-ci ne soit jamais en manque de ressources :

---

<sup>479</sup> D'après Carlo, le membre de l'équipe qui, tout en ayant des tâches d'ingénieur et de programmeur, est depuis son arrivée en 2010 le contact pour les entreprises et le responsable de la « valorisation » de Tribler dans le consortium P2P-Next (voir Sections 3 et 4).

<sup>480</sup> <http://www.abacast.com/abacast-technologies>

<sup>481</sup> <http://www.rawflow.com/>

<sup>482</sup> Je dois l'explication de la stratégie de développement Tribler dans cette phase à Hjalmar, Carlo et Jingerzi.

<sup>483</sup> Entretien, Hjalmar.

*« For live streaming it was more difficult to get infinite duration, in the sense that with BitTorrent, the client works with files of pre-determined length, you know right away how big or long a file will be. For live streaming, you don't know how long a session will be, may be infinite, you have to make the tool so you can overcome the limited duration feature, and that was tricky to do<sup>484</sup>. »*

L'attention aux défis techniques à l'œuvre dans le streaming décentralisé – et, en parallèle, plusieurs déclarations qui rendent manifeste l'intérêt de Tribler pour ce champ de recherche et d'application<sup>485</sup> – amènent à Tribler, au cours de 2009 et pendant que la série 5.x est implémentée, plusieurs propositions de partenariats R&D centrés sur le développement de la télévision via Internet. Il s'agit d'une phase d'intenses « relations publiques » pour Friedrich Bierbach<sup>486</sup>. Le coordonnateur scientifique de Tribler veut capitaliser rapidement sur la popularité (qui pourrait bien être de courte durée dans le volatile secteur IT) de Tribler auprès d'autres universités, entreprises et communautés *open source* pour compléter la redéfinition et reconfiguration de Tribler et passer de client P2P pour partage de fichiers à une véritable plateforme d'échange et diffusion en direct de contenus multimédias. Son but principal est de continuer à assurer la survie du projet : les financements qui dérivent d'I-Share et de son projet jumeau sont destinés à s'achever à l'automne 2009.

À la mi-2009, l'équipe Tribler est impliquée dans plusieurs partenariats R&D, au sein desquels son rôle est soit la coordination scientifique et technique, soit le développement de code. La néerlandaise Publieke Omroep (association de diffuseurs de télévision publique hollandais) planifie avec Tribler un essai à large échelle d'une infrastructure de streaming. AHT International, un fabricant hollandais de décodeurs pour télévision, a publié comme on l'a vu une des versions 5 de Tribler sur son dernier décodeur, appelé Tribbox, et collabore plus largement avec l'équipe pour la programmation d'une plateforme ouverte de télévision P2P. C'est à ce moment que Friedrich reçoit une nouvelle déterminante pour le développement de Tribler : dans le cadre du septième programme-cadre de l'Union Européenne<sup>487</sup>, un financement de quatorze millions d'euros est attribué à l'équipe hollandaise et à sa complexe coalition de partenaires, laboratoires scientifiques et entreprises, désormais appelée le P2P-Next Consortium.

---

<sup>484</sup> Id.

<sup>485</sup> Voir par exemple les déclarations de Friedrich Bierbach à <http://www.tgdaily.com/software-features/52967-bittorrent-gets-decentralized>.

<sup>486</sup> Entretien, Carlo.

<sup>487</sup> [http://cordis.europa.eu/fp7/home\\_en.html](http://cordis.europa.eu/fp7/home_en.html)

### SECTION 3. P2P-NEXT : VERS UN RÉSEAU EUROPÉEN DE TÉLÉVISION DISTRIBUÉE

Au cours de 2008, le croisement de plusieurs facteurs – l’expérience de développement de Tribler, avec ses défis techniques et plusieurs années de travail consacrés à les résoudre ; la mise en place de canaux de communication privilégiés avec les institutions publiques hollandaises ; les réseaux tissés avec le monde des moyennes et grandes entreprises de la télévision et de la distribution de contenus – a en effet convaincu Friedrich Bierbach que le moment se prêtait bien à essayer d’obtenir un important financement public au niveau international. Il s’agit d’exploiter nombre de changements en cours dans la manière dont les utilisateurs des médias s’approprient des contenus, les font circuler, en découvrent de nouveaux :

*“These were the first landmarks for the P2P-Next project. (...) we built on our Tribler experience of several years to figure out the contribution that could be made to what is often called the ‘digital convergence’. It looked to us like this was the deal: we were moving from classical, linear TV to a non-linear model, that would privilege things like on-demand, anytime-anywhere, cross-device and all of this was going to change the roles of publishers and consumers. Plus, a model of media consumption that could also be called linear, I suppose, was giving room to personal, active behavior at home, and also in mobility situations. An offer of content that became popular by default as it was being offered in a monopoly situation was leading to selective disclosure of content... and the appropriation of this content by communities with varying and diverse interests. In parallel with all this, the publicity models are also changing, for sure, and we could leverage that aspect with our commercial partners. Forms of promotion are becoming more targeted to specific audiences, non-disruptive, and oriented to specific channels. All of this called for TV on P2P. For sure, we were not going to be the only ones interested<sup>488</sup>.”*

On a vu comment, pendant ses premières années de vie, Tribler s’est positionné comme dispositif au potentiel novateur dans plusieurs arènes économiques, politiques et techniques, grâce à la double dynamique du « P2P social » et de la « télévision P2P ». Au fil de ces années, Tribler a porté plusieurs casquettes : celle de projet principalement financé par de l’argent public, mais auquel des partenaires privés prêtent intérêt comme clé d’entrée dans un marché jusque là relativement inexploré. Celle de réseau, à la fois de personnes et de machines, pouvant supporter à moyen et long terme les modifications de l’offre et de la demande dans un secteur profondément bouleversé par les nouvelles pratiques de consommation et de distribution. Celle de système qui, construit sur une technologie aux usages fortement controversés au même moment dans les arènes politiques<sup>489</sup>, en propose une relecture aux connotations, fortes et également « politiques », d’ouverture, de transparence et

---

<sup>488</sup> Entretien, F. Bierbach.

<sup>489</sup> Il s’agit du moment déjà décrit dans la section 2 du chapitre 5.

d'interopérabilité.

Ces différentes facettes ont été présentes pendant la mise en place de partenariats à la fois dans le public et dans le privé, pendant que l'équipe continuait à se concentrer sur le développement de la partie technique du logiciel, en ouvrant des portes et en fermant d'autres. Cette pluralité de fronts – et une certaine habileté à brouiller les frontières aux moments opportuns – est d'après Friedrich ce qui lui vaudra l'attribution de ce financement de quatorze millions d'euros dans le cadre du septième programme-cadre de l'Union européenne :

*« I spent Tribler's first two years planting a lot of seeds. It was certainly really, really good news that it was the biggest one to pay off<sup>490</sup>. »*

Le P2P-Next Consortium se dote d'un objectif très ambitieux : construire NextShare, un réseau européen de télévision via l'Internet, basé sur une technologie de streaming vidéo en pair à pair, aux connotations sociales et qui « tient en considération le cadre législatif européen existant<sup>491</sup> ».

Deux autres aspects sont censés guider la démarche du consortium. En premier lieu, pour faciliter l'identification et l'élaboration de nouveaux modèles d'affaires applicables éventuellement à d'autres contextes, le code et les résultats scientifiques issus par le projet seront rendus disponibles en clair sur le site de P2P-Next<sup>492</sup>. Deuxièmement, la formulation du projet révèle la volonté de se servir des quatre premières années de P2P-Next en tant qu'arène d'expérimentation pratique pour résoudre nombre de questions ouvertes concernant la légalité et le potentiel commercial de la technologie pair à pair, qui est, pour le moment, assimilée à des pratiques douteuses :

*« P2P-Next is also meant to address a number of outstanding challenges related to content delivery over the internet, including technical, legal, regulatory, security, business and commercial issues<sup>493</sup>. »*

*« P2P still has a somewhat dubious reputation as an illegal file sharing mechanism akin to Napster, Kazaa, Glockster, etc. Nevertheless, (...) Broadcasters and content providers consider P2P as a future-proof, universal, and ubiquitous two-way (interactive) distribution mechanism. Initially, P2P will complement the existing distribution mechanisms such as satellite, cable and terrestrial networks, but ultimately it may supersede them<sup>494</sup>. »*

---

<sup>490</sup> Entretien avec F. Bierbach.

<sup>491</sup> <http://www.tribler.org/trac/wiki/P2P-Next/19Million-for-P2P>

<sup>492</sup> CORDIS (2008). “European research project to shape next-generation Internet TV”, [http://cordis.europa.eu/search/index.cfm?fuseaction=news.document&N\\_RCN=29153](http://cordis.europa.eu/search/index.cfm?fuseaction=news.document&N_RCN=29153) : “All core software technology will be available as open source, enabling new business models”.

<sup>493</sup> Id.

<sup>494</sup> P2P-Next, “Objectives”, <http://www.p2p-next.org/?page=content&id=73F87D854E37E0F75C68C69304535873&mid=EE056556C09ED2492B>

La notion de « mise à l'épreuve » d'un cadre technique et législatif émergeant par l'intermédiaire des activités de P2P-Next est également présente dans les discours des autorités européennes responsables de l'attribution du financement :

*« This ambitious project is investigating new ways to deliver video using the P2P paradigm (...) It will be a live trial of the future media Internet we are envisioning. (...) the project we are launching today will help European stakeholders to be at the frontline in pioneering the looming media revolution enabled by the Internet<sup>495</sup>. »*

À ses débuts, le projet P2P-Next est donc défini, à la fois par ses promoteurs et ses financeurs, comme un laboratoire où le public et le privé, liés en partenariat, souhaitent se démarquer des politiques et des modèles d'affaires dominants au niveau européen par rapport au P2P. Il s'agit là du premier pas vers le dépassement d'une conception du pair à pair comme une technologie dont il faut, d'un point de vue politique et légal, se « défendre », se protéger, et vers sa construction comme « valeur » pour l'Europe.

Cette démarche est confrontée cependant à nombre de défis inédits. D'un côté, les développements techniques ultérieurs de Tribler dans le cadre de P2P-Next doivent tenir compte des réactions du public européen à un projet dans lequel, comme plusieurs journaux et périodiques se plaisent à le rappeler, les organismes publics de l'audiovisuel sont en train de faire des expérimentations sur les technologies pair à pair tant aimées des pirates<sup>496</sup> grâce à des financements très largement publics. Si l'opportunité politique de construire le peer-to-peer comme « valeur » finit par influencer la démarche des développeurs, l'organisation du travail entre les partenaires doit à son tour tenir compte à la fois de la meilleure manière d'avancer dans le travail technique, et des nécessités de positionnement dans le marché des partenaires privés du consortium. Sont en jeu des questions telles que la transition entre le travail expérimental sur le peer-to-peer, technologie souvent considérée comme incompatible avec l'économie marchande, et les modèles d'affaires existants ; ou encore l'opérationnalisation de dispositifs destinés à assurer la « légalité » du futur système et sa compatibilité avec le cadre juridique européen.

---

[0FF130393D054B](#)

<sup>495</sup> Déclaration de Luis Rodríguez-Roselló, responsable de l'Unité *Networked Media Systems* ([http://cordis.europa.eu/fp7/ict/netmedia/home\\_en.html](http://cordis.europa.eu/fp7/ict/netmedia/home_en.html)) de l'Union Européenne, à l'occasion du lancement de P2P-Next (<http://www.tribler.org/trac/wiki/P2P-Next/19Million-for-P2P>).

<sup>496</sup> Staff Member (2010, 4 mars). "The Bigger Picture. Video on the Internet: Why are public broadcasters experimenting with the 'peer-to-peer' technology beloved of online pirates?" *The Economist Technology Quarterly*, <http://www.economist.com/node/15582215>

## UNE COALITION COMPLEXE POUR UN OBJET CONTROVERSÉ

À ses débuts, le projet P2P-Next réunit vingt-et-un partenaires repartis dans douze pays européens : des acteurs de la distribution de contenus comme la BBC jusqu'à des fabricants de *hardware* comme Pioneer, en passant par des laboratoires de recherche en informatique tels que les équipes de Delft et d'Amsterdam, impliquées dans Tribler, il s'agit d'un ensemble hétérogène et complexe à gérer. Le premier défi de taille pour l'équipe Tribler, en tant que coordonnateur scientifique du projet, se révèle dès le début d'assurer une certaine stabilité et une certaine efficacité aux activités du consortium, regroupé autour de cet objet à la fois complexe et controversé, le P2P<sup>497</sup>.

Du point de vue de l'équipe Tribler, mettre en place un projet de cette ampleur, avec des partenaires nombreux et variés, était la seule manière possible d'obtenir un consensus avec les pouvoirs politiques européens autour du peer-to-peer en tant qu'objet de recherche, mais aussi en tant qu'opportunité économique<sup>498</sup>. Il fallait donc que les possibilités de création de nouveaux modèles d'affaires, et le potentiel de changement et d'innovation pour les entreprises concernées, soient clairement inscrits dans la demande de financement. En même temps, il s'agissait de s'assurer la collaboration d'acteurs du secteur privé de premier plan, se déclarant suffisamment intéressés par l'objet de recherche pour y investir de leur côté des subsides qui s'ajoutent au financement européen :

« Dans ce modèle des financements européens, si les universités et les acteurs du monde de la recherche pure sont complètement couverts par le financement – s'il y a, mettons, cinq personnes à plein temps sur le projet, elles sont payées complètement avec des fonds européens – les entreprises et les acteurs du monde de l'industrie ne sont payés qu'à moitié et doivent subvenir à leurs besoins pour l'autre moitié. En fin de compte, cela donne au secteur privé une

---

<sup>497</sup> La liste complète des partenaires de P2P-Next est la suivante (source: site web du projet, <http://www.p2p-next.org/>, "Consortium"): AG Projects - Haarlem, Netherlands (<http://www.ag-projects.com/index.php>); British Broadcasting Corporation (BBC) - London, United Kingdom (<http://www.bbc.co.uk/>); DACC Systems - Taby, Sweden (<http://www.dacc.se/index.html>); Technische Universiteit Delft - TU Delft - Delft, Netherlands (<http://www.tudelft.nl/>); FABChannel - Amsterdam, Netherlands (<http://www.fabchannel.com/>); Institut für Rundfunktechnik - IRT - Munich, Germany (<http://www.irt.de/>); Josef Stefan Institute - Ljubljana, Slovenia (<http://www.ijs.si/ijsw/JSI>); Kendra Foundation - London, United Kingdom (<http://www.kendra.org.uk/>); Kungliga Tekniska Högskolan - KTH - Stockholm, Sweden (<http://www.kth.se/?l=en>); Markenfilm - Wedel, Germany (<http://www.markenfilm.de/>); Norut- Tromsø, Norway (<http://en.itek.norut.no/>); First Oversi - Petach Tikva, Israel (<http://www.oversi.com/>); Pioneer Digital Design Centre Limited - London, United Kingdom (<http://www.pioneer.co.uk/uk/body.html>); RTV Slovenia - RTVSLO - Ljubljana, Slovenia (<http://www.rtv slo.si/>); STMicroelectronics - Milan, Italy ([www.st.com](http://www.st.com)); The European Broadcasting Union (EBU) - Geneva, Switzerland (<http://www.ebu.ch/>); University of Klagenfurt - Klagenfurt, Austria (<http://www.uni-klu.ac.at/>); University of Lancaster - Lancaster, United Kingdom (<http://www.lancs.ac.uk/>); University of Rome - Rome, Italy (<http://www.uniroma1.it/>); University Politehnica of Bucharest - Bucharest, Romania (<http://www.pub.ro/English/eng.htm>); VTT – Technical Research Centre of Finland, Tampere, Finland (<http://www.vtt.fi/index.jsp?lang=en>).

<sup>498</sup> Entretien, Carlo.

occasion d'investir dans quelque chose à l'avant-garde, de très expérimental, dans lequel ils n'auraient normalement pas investi... et de pouvoir disposer de beaucoup plus que ce qu'ils ont investi. Dans ce cas, le peer-to-peer<sup>499</sup>. »

Friedrich et son équipe se trouveront par ailleurs confrontés, à plusieurs reprises, à une vague de critiques émanant du public et des arènes politiques nationales et internationales – qui, en dépit de plusieurs marques d'intérêt de l'opinion publique hollandaise, estiment que l'argent des contribuables ne doit en aucun cas être utilisé pour soutenir des efforts relatifs au P2P<sup>500</sup> ; ces critiques vont dans un cas jusqu'à étiqueter le travail de l'équipe Tribler comme « innovations interdites », ce qui finira par avoir un impact performatif sur le projet<sup>501</sup>. Friedrich finira par éclaircir la position de Tribler, et plus généralement de P2P-Next, sur les pages du populaire site TorrentFreak :

*“Tax payer money is going into Internet research, which happens to use a very powerful technology called BitTorrent. That's different. On a wider scale a few hundred million euros of research money is being spent on making computer networks more robust and improving video streaming. I think that is money well spent<sup>502</sup>.”*

Du côté des entreprises, la présence d'universités et de centres de recherche académiques dans le consortium légitime, grâce à un certain nombre de publications et de livrables, leur intérêt commercial pour l'objet peer-to-peer : cela donne l'occasion au secteur privé d'utiliser des résultats de recherche qui sont généralement, selon Carlo, d'une meilleure qualité que ceux auxquels l'industrie est habituée<sup>503</sup>. En même temps, les entreprises sont souvent conduites, dans les premières phases du projet, à justifier de leur engagement dans le développement du P2P : le fait que ces acteurs du secteur privé décident d'investir et de s'investir dans des activités de recherche fondamentale, extérieures à l'entreprise, et sur un objet controversé, est lu de plusieurs côtés comme une « action de marketing<sup>504</sup> » en soi, impliquant une prise de position politique.

Si en règle générale, le rapprochement de plusieurs entreprises réputées et de leurs partenaires académiques semble obtenir des retours favorables dans le monde de la

---

<sup>499</sup> Id. Comme cela a déjà été le cas, l'entretien ayant été conduit en italien, notre langue natale à tous les deux, j'en traduis les extraits en français.

<sup>500</sup> Trent Nouveau (2010, 4 décembre). “BitTorrent Gets Decentralized.” *TGDaily*, <http://www.tgdaily.com/software-features/52967-bittorrent-gets-decentralized>

<sup>501</sup> Entretien, F. Bierbach. Je ne suis pas arrivée à connaître le contexte précis de cet échange, qui a pourtant été très important pour le cadrage des activités de Tribler par la suite (voir l'introduction à ce chapitre).

<sup>502</sup> <http://torrentfreak.com/tribler-bittorrent-client-adds-magic-search-110829/>

<sup>503</sup> Entretien, Carlo.

<sup>504</sup> Voir les articles de *TG Daily* et de l'*Economist Technology Quarterly*, ci-dessus.

recherche<sup>505</sup>, quelques discussions et incidents centrés sur l' « engagement P2P » de certaines entreprises marquent les débuts du consortium. Le chef du département R&D de la BBC fait part à l'équipe Tribler du scepticisme du conseil d'administration sur l'investissement réalisé par la société publique de diffusion britannique<sup>506</sup>. Bien qu'en exprimant peu de doutes quant à l'efficacité technique du modèle de distribution décentralisée, la société publique de diffusion norvégienne NRK, partenaire de P2P-Next, exprime quant à elle ses craintes de souffrir d'une publicité négative en cas d'utilisation d'un système basé sur BitTorrent, donc « associé au piratage<sup>507</sup> », pour rendre le documentaire-marathon *Bergensbanen*<sup>508</sup> accessible à un plus large public après que sa première diffusion télévisée a rencontré un succès inattendu. Les caractéristiques techniques du fichier vidéo (sept heures de film en haute définition, pour un volume de 250 gigaoctets) risquent pourtant d'encombrer les serveurs de l'entreprise au point d'en compromettre le bon fonctionnement : la distribution du documentaire servira finalement comme un des premiers test à large échelle du système NextShare, mais pas avant que NRK n'ait précisé à son audience, par un communiqué de presse, la neutralité (*value-neutral*<sup>509</sup>) du mécanisme de BitTorrent.

D'un autre côté, certains partenaires semblent réussir à intégrer le P2P dans leurs stratégies marketing, ainsi que dans leurs activités de recherche et développement, de manière moins problématique. En 2009, Pioneer annonce le lancement, dans le cadre de P2P-Next, d'un prototype de décodeur (ou *set-top box*) compatible avec la prise en charge du streaming *via* BitTorrent<sup>510</sup>. IT Microelectronics, une firme de Milan partenaire du projet, concentre ses efforts sur les algorithmes de codage nécessaires au développement d'un produit très similaire<sup>511</sup>. Markenfilm, partenaire allemand du projet, explore des modèles d'affaires possibles autour d'un système de publicité ciblée et d'un système de *pay-per-content*<sup>512</sup>. Le texte du projet lui-même, d'ailleurs, prévoit un équilibre entre la recherche pure et le développement de produits, avec un plafond maximal pour le premier type d'activité qui ne peut pas dépasser 80%. En général, souligne Carlo,

« tout cela donne des produits qui, en l'état, ne finiront pas sur le marché,

---

<sup>505</sup> Par exemple au NEM (Networked & Electronic Media Technology Platform) Summit de Saint-Malo et à IBC (Amsterdam), où le consortium présente les premiers résultats du projet en 2009.

<sup>506</sup> Entretien, Carlo.

<sup>507</sup> *Economist*, cit.

<sup>508</sup> « La ligne Bergen », un documentaire de sept heures et demi produit par NRK, retraçant le parcours du train historique qui, de Bergen à Oslo, couvre la quasi-totalité de la côte Ouest norvégienne : <http://www.newsinenglish.no/2009/11/30/marathon-documentary-marks-bergen-oslo-lines-100th-year/>

<sup>509</sup> *Economist*, cit.

<sup>510</sup> <http://torrentfreak.com/pioneers-live-bittorrent-streaming-device-080911/>

<sup>511</sup> Entretien, Carlo.

<sup>512</sup> Id.

parce qu'il y a trop de recherche à la base pour qu'on puisse investir à large échelle sur un produit ; mais ils ouvrent la voie à des générations futures de produits. Et une fois que (P2P-Next) est terminé, si les entreprises veulent continuer à investir dans la même direction, elles peuvent le faire avec moins de problèmes<sup>513</sup>. »

Ces jeux de négociation pendant la première année d'existence du consortium contribuent à stabiliser la structure générale de P2P-Next en cinq « couches<sup>514</sup> ». La première de ces couches concerne la relation entre les différentes parties prenantes du consortium et les exigences et spécificités techniques du système. Guidé par le Northern Research Institute (Norut), qualifié comme directeur stratégique du projet, cette couche naît de la nécessité d'analyser et comprendre finement l'environnement dans lequel P2P-Next souhaite s'inscrire, et de le traduire en contraintes et possibilités techniques. Les moyens à la disposition de directeur stratégique sont ainsi identifiés :

*“identifying stakeholder requirements, exploring future business models, continuous technology watching, and understanding end-user requirements using scenarios based on reality and centred on the user<sup>515</sup>.”*

La deuxième couche, sous la responsabilité du directeur de création (la BBC), se concentre sur la gestion des contenus spécifiquement créés pour la plateforme de distribution NextShare, à la fois par les professionnels impliqués dans le projet, et par les usagers. L'équipe de Delft, coordinateur technique et scientifique de P2P-Next, assume la tâche de traduire les résultats des deux premières couches, exigences techniques de système et contenus, en un *enabling framework*, un cadre qui facilite leur conversion en applications concrètes. Ce sont donc, en résumé, les fondations architecturales du projet qui sont en jeu avec le travail de Friedrich et son équipe. Pioneer, en tant que directeur de l'innovation, a la charge de la démonstration et de la mise en œuvre des résultats intermédiaires : cette quatrième couche comporte la mise en place d'un environnement d'expérimentation des capacités du système à large échelle. En constituant l'arène de mise à l'épreuve du caractère intuitif des usages, de la fiabilité technique et de l'interopérabilité des premiers prototypes du système NextShare, les *Living Labs* sont censés être le

*« reality check on our vision, platform, technology, scenarios, and possibly the business models we have in mind, (with the idea of) targeting, eventually, the living room's TV screen<sup>516</sup>. »*

---

<sup>513</sup> Id.

<sup>514</sup> *Streams*, peut-être mieux rendus en français par “couches”. Entretien, F. Bierbach.

<sup>515</sup> Entretien, F. Bierbach, reformulation d'après le premier livrable Tribler. Le paragraphe suivant se base aussi sur cet entretien.

<sup>516</sup> Entretien, F. Bierbach.

Enfin, la couche centrée sur la gestion et l'impact, gérée par le directeur de projet (*Technical Research Centre of Finland*), est destinée à assurer le bon déroulement du projet et ses activités de diffusion des résultats, avec la création possible d'une fédération P2P-Next réunissant, au-delà du cadre chronologique du projet, les principaux acteurs de la diffusion de contenus sur un modèle distribué.

On voit comment, en explicitant les liens entre les objectifs généraux du projet et ses domaines d'opération pratique, les cinq couches problématisent aussi, de manières différentes, le P2P : il apparaît comme un ensemble de contraintes techniques, mais aussi comme une architecture pouvant faciliter, à plus d'un titre, l'efficacité du modèle de diffusion ; comme une démarche de création de contenus spécifiquement adaptés aux architectures distribuées ; comme une ingénierie de l'environnement dans laquelle les pairs deviennent à la fois décodeurs et écrans, pour former, du point de vue de l'utilisateur, partie intégrante de l'expérience de télévision ; et, finalement, comme un modèle de gouvernance du secteur de la diffusion distribuée, dont l'efficacité et le bon fonctionnement sont strictement liés aux apports de chacune des parties prenantes. La clé du projet européen de télévision par Internet sera l'enchevêtrement de ces différents niveaux de distribution, entre lesquels le P2P se déploie à la fois comme objet technique complexe, système d'interaction, modèle de gestion et dynamique de production.

#### RESTER TRIBLER DANS P2P-NEXT : GÉRER LA « MISE-EN-GLOBALITÉ » D'UN PROJET LOCAL

Avec l'importance croissante, stratégique et chronophage, du rôle de coordination et gestion scientifique assumé par l'équipe Tribler, celle-ci en vient à négocier une fois de plus son statut de projet universitaire, fondé sur l'*open source*, et conduit par une équipe aux dimensions relativement réduites. Deux stratégies coexistent désormais au sein de l'équipe, répondant à deux « identités » différentes<sup>517</sup>. L'une est guidée par les exigences de recherche à un niveau local : l'équipe Tribler est encore essentiellement composée d'étudiants de Master et de doctorat, qui choisissent la « brique » sur laquelle travailler de façon relativement indépendante et selon leurs intérêts propres. L'autre, beaucoup plus reliée aux nécessités de développement de prototypes au sein du consortium, est largement dépendante des négociations que l'on vient de décrire : une partie essentielle de cette démarche est l'ajustement progressif de l'algorithme Buddycast à l'architecture globale de P2P-Next. Hjalmar tient cependant à ce que la démarche du groupe de Delft maintienne une certaine originalité et indépendance, étroitement liée à l'identité de recherche, tout en étant conscient que les négociations conduites dans le projet européen sont à la fois nécessaires et opportunes :

*« We try to keep on doing our own thing and try to make it fit on what P2P-Next wants as much as possible. So, we are basically research-driven and we try to put what comes out of it into P2P Next. Tribler keeps on coming first (...) as a university, our main objective keeps on being research, in the end*

---

<sup>517</sup> Entretien, Hjalmar.

*despite the European project we are judged by our research outputs. Primarily, I am busy with making Tribler fit into P2P-Next as a programmer, while others are more specifically focused on Tribler's core features<sup>518</sup>. »*

La valeur ajoutée dérivée des conversations et activités conjointes avec les partenaires consiste par ailleurs dans la possibilité de sortir de la « pure recherche » et de soumettre à des épreuves de réalité un dispositif qui, à la différence de plusieurs autres projets universitaires de recherche et expérimentation P2P, est déjà un service saisi par un public d'utilisateurs. L'idée que l'avantage comparatif de Tribler est d'être un « *product out there* » (v. Section 2) est à nouveau présente dans les paroles d'Hjalmar. À différence de plusieurs projets P2P universitaires qui restent retranchés dans une pure logique de laboratoire, y compris dans la phase de test, Tribler est appropriable et saisissable – en effet, déjà approprié et saisi – par d'autres acteurs, sans que le chercheur ou l'entité de recherche qui l'a créé soit la seule autorité en matière et son *gatekeeper* exclusif. Les concepteurs de Tribler se confrontent de fait avec des usagers et des usages :

*« What's interesting about Tribler is that our product is out there and has to work, so we are not at the stage in which only that particular researcher can make it work and nobody else can because they don't know all the tweaks; our output is already good or at least proven quality, so it makes it easier for the partners to appropriate and use it. In my past years as a 'pure researcher' at a couple of universities, it didn't work that way, everybody was coming with his own prototype but they were all less consolidated things<sup>519</sup>. »*

De son côté, Friedrich Bierbach – qui s'occupe de plus en plus des aspects organisationnels du consortium, mais ne cesse pas pour autant de s'intéresser au développement technique de Tribler – voit dans le projet européen une occasion de remettre sur la table, d'une manière différente et à plus grande échelle, le paradigme « social » dont il avait été le premier inspirateur et promoteur aux débuts de Tribler :

*« One of the underlying principles of NextShare still remains the participation of several peers in a common activity, which in this case, for example, would be sharing a video clip. P2P-Next is a new chance to facilitate the introduction of several social features in content distribution, this time to support concrete, actual user and business communities. With P2P-Next, we have the chance to organize an incremental approach, by early and short-cyclic releases involving actual user communities, making use of the Living Labs<sup>520</sup>. »*

Les financements dérivés du projet vont être finalement l'occasion de fonder les choix de développement sur des essais et des expérimentations concrètes, conduites

---

<sup>518</sup> Entretien, Hjalmar.

<sup>519</sup> Id.

<sup>520</sup> Entretien, F. Bierbach.

au moyen des *Living Labs*, ces collectifs de volontaires recrutés sur internet et opérant par son intermédiaire : ceci avait manqué, faute de temps et de financements suffisants, lors de la première phase, conduisant à l'abandon des fonctionnalités sociales au niveau de l'interface<sup>521</sup>. Friedrich espère ainsi réduire, par la prise en compte des retours des utilisateurs pionniers des Labs, l'écart souvent présent entre la vision des développeurs et la perception des usagers.

## CONSTRUIRE UNE « LÉGALITÉ » DU P2P ENTRE LICENCES LIBRES ET DROIT EUROPÉEN

Les premières expériences avec le consortium dans un cadre de projet européen amènent l'équipe Tribler, à la fois de façon indépendante et en relation avec les partenaires, à reconfigurer et définir avec plus de précision leur positionnement déclaré de P2P « légal ». J'utilise les guillemets pour le mot « légal » pour souligner que l'équipe Tribler (et le consortium P2P-Next) ne cherchent pas à s'insérer pleinement (et à s' « effacer ») dans la définition du terme mobilisée par les ayants droits de l'industrie des contenus numériques, renvoyant à des contenus dont la diffusion est autorisée par les auteurs, mais construisent des définitions de légalité qui mettent l'accent sur ses aspects locaux, contestés, recherchés, ambigus – pour suggérer des pistes de régulation techno-juridique alternative pour les réseaux P2P et la variété d'objets qui y circulent. À partir d'ici, le mot « légal » ne sera plus entre guillemets, mais sera utilisé dans cette acception (sauf précision du contraire).

Né comme logiciel totalement inséré dans le domaine du libre et de l'*open source*, Tribler est initialement publié sous une licence GNU Lesser General Public License (LGPL)<sup>522</sup>: construit à partir des clients ABC et BitTorrent, le code est en partie soumis aux licences d'origine. La philosophie du libre semble aussi imprégner la philosophie initiale sous-tendant le projet, qui, comme on a pu voir dans la première section de ce chapitre, répond à la question de la responsabilité juridique par l'argument du « pur instrument » : les efforts de l'équipe sont concentrés sur la mise en place d'un outil de distribution efficace et rapide, et la responsabilité liée à d'éventuels usages illicites de cet outil retombe entièrement sur ceux qui les utilise. Cet argument, comme on l'a déjà vu à plusieurs reprises dans cette thèse, ne s'est pas révélé suffisant pour contourner la question, qui revient donc à la surface sous d'autres formes.

Par la suite, au cours de 2008, l'insertion de Tribler dans le domaine du P2P légal semble chercher d'autres formes de validation aux yeux du public. Elle en trouve dans les collaborations et partenariats établis avec des acteurs du marché et des politiques publiques ; du fait de leur statut « institutionnel » ou de collaboration avec des institutions, ces acteurs légitiment d'une certaine façon l'ambition de Tribler de se qualifier comme P2P légal. Ils contribuent aussi à souligner qu'une variété de contenus, légaux comme illégaux, peuvent être distribués sur les réseaux P2P : par

---

<sup>521</sup> Qui a fait l'objet de la deuxième section de ce chapitre.

<sup>522</sup> Dans la version 2.1 (<http://www.gnu.org/licenses/lgpl-2.1.html>).

exemple, d'autres logiciels libres ou des contenus audiovisuels qui sont déjà dans le domaine public, et donc libres de droits d'auteur (Tableau 11).

**Legal P2P, does that exist ?**

Yes, and it is rapidly picking up speed. Tribler will be cooperating with the Dutch public broadcasters to make content available via peer-to-peer, in the same line of sites like [www.vuze.com](http://www.vuze.com), that distributes movie trailers in HD via BitTorrent, [www.publicdomaintorrents.com](http://www.publicdomaintorrents.com) and [www.legaltorrents.com](http://www.legaltorrents.com), and the distribution of free software carried out via BitTorrent (for example, the distribution of the OpenOffice office suite). The Internet Archive currently seeded by BitTorrent includes a number of classic films now in the public domain and modern experimentations with open source video tools, such as Elephants Dream (2006), the world's first animation movie created entirely with open-source video tools, and Sita Sings the Blues (2009).

*Tableau 11. Extrait de la foire aux questions (FAQ) du site Tribler (<http://tribler.org/trac/wiki/faq#LegalP2Pdoesthatexist>), consultée en décembre 2009.*

Si ces arguments peuvent suffire à Tribler pendant sa « première vie » de projet universitaire et *open source*, l'élaboration d'une définition de légalité du P2P plus complexe et articulée devient bientôt nécessaire lorsque le portail web de P2P-Next est ouvert au public, en 2009. Le travail sur la présentation du projet dans le site cherche donc à expliciter l'opposition entre le discours public prévalant sur le P2P comme technologie illégale, « innovation interdite », et les caractéristiques techniques qui font du P2P une technologie particulièrement adaptée à l'échange de lourds contenus audiovisuels. La question des types de contenus échangés par ses moyens sera ensuite dans les mains de partenaires fiables et de premier plan dans le secteur : la réputation « douteuse » du P2P peut donc passer en deuxième plan face au fait que

*« today, (P2P) is considered by many as an efficient, reliable, and low cost mechanism for distributing any media file or live stream<sup>523</sup>. »*

En parallèle, des négociations sont engagées avec les financeurs du projet – c'est-à-dire l'Union européenne – pour « vendre » l'idée de distribution libre et efficace de vidéos d'une manière compatible avec le cadre juridique existant en matière de droit d'auteur<sup>524</sup>. Encore une fois, l'accent est mis sur le fait que les principaux « distributeurs » du système NextShare seront des partenaires reconnus et légitimes ; du côté des usagers, on discute par ailleurs de la possibilité d'introduire une fonctionnalité de localisation, non pas personnelle mais géographique, qui doit faciliter l'application à NextShare des différentes politiques nationales en termes des

<sup>523</sup> <http://www.p2p-next.org/?page=content&id=73F87D854E37E0F75C68C69304535873&mid=EE056556C09ED2492B0FF130393D054B>

<sup>524</sup> Entretien, Carlo.

modalités de distribution des flux en streaming<sup>525</sup>.

P2P-Next en arrive donc à définir une partie de ses objectifs vis-à-vis de l'Europe en termes juridiques – d'un côté pour stabiliser son propre statut de P2P « légitime » et de l'autre, pour apporter une contribution à l'affinement d'un cadre juridique pour le P2P qui, tout en les reliant, ne confonde pas la technologie avec ses usages<sup>526</sup>. P2P-Next estime pouvoir s'attaquer à trois types de problème d'ordre à la fois technique et juridique :

*« So far, the three major problems that have prevented a fuller development of authorised P2P are the lack of efficient technical solutions for rights management, the lack of business models suited to the content industry and a stable legal framework that might take into account differences between countries, but also be applicable across them. For the first aspect, the relation to digital rights management... ok, let's say concisely that any video content can be protected on a P2P system provided that the end points are black boxes. But once decrypted, an anonymous P2P system can also spread this content fast. This sounds like a chicken-and-egg problem, and I think it can only be solved by applying strict secure computing, and secure transmission techniques. So far, these techniques are far from successful, mainly because they hamper the users freedom to copy content for proprietary use, and on the other hand they can easily be prevented<sup>527</sup>. »*

Les premières phases de P2P-Next cherchent donc à jeter les bases d'un modèle économique pour la distribution vidéo sur Internet qui puisse dès le départ intégrer au système des éléments à la fois juridiques et techniques. L'idée qui prend forme de manière relativement explicite au sein du consortium est que cette démarche peut dépasser le cadre du projet, pour contribuer à la reformulation de la définition de P2P elle-même, en opposition à des tentatives de régulation *ex post* et surplombants qui ont caractérisé la réponse juridique au P2P jusque là.

Si le texte du projet P2P-Next, tel qu'il avait été soumis à l'Union européenne, centrait son attention sur la possibilité d'inclure dans le modèle des mécanismes d'authentification, de *pay-per-use*, de souscription – embryons techniques d'un modèle économique pour le streaming en P2P – ce n'est qu'avec le premier livrable que la clé du succès du projet est identifiée autour de deux aspects qui engagent une vision et problématisation politique du P2P à un niveau bien plus profond. D'un côté, l'acceptation progressive du P2P en tant que technologie « légitime » de la part des industries des contenus, dont les partenaires de P2P-Next ne constituent qu'une avant-garde avec un penchant pour l'expérimentation ; de l'autre, la nécessité d'un

---

<sup>525</sup> C'est par exemple le cas d'une limitation imposée par BBC, qui pose comme condition que les vidéos visionnés sur son site ne soient pas accessibles en dehors du territoire national anglais.

<sup>526</sup> Voir (Elkin-Koren, 2006) pour un traitement clair et efficace de ce sujet, à partir d'une perspective juridique, en rapport au P2P et à ses "vertus".

<sup>527</sup> Entretien, F. Bierbach.

cadre juridique stable, inspiré par l'Europe, pour encadrer cette légitimation.

La quatrième et dernière section de ce chapitre s'attachera à montrer comment, au fil du déroulement du projet, ces deux aspects se sont articulés dans la démarche du consortium et de l'équipe Tribler en particulier ; ce faisant, on espère donner quelques pistes d'analyse de la problématisation politique du *peer-to-peer* en tant que modèle alternatif pour les services Internet, et finalement, de sa légitimation comme « valeur » dans le contexte européen.

#### SECTION 4. DE PROBLÈME À VALEUR : LE PEER-TO-PEER DANS L'EUROPE

*This looks great. Now I suppose the things to ask are:*

- a) how long before it is available to everyone as a real alternative to the p2p technology being used atm [at the moment]?*
- b) how long before the entertainment industries stop it from developing further?*
- c) how long before all those involved in this project are sued for 'facilitating copyright infringement'?*
- d) how long before the EU is sued for funding the project ?*
- e) how long before the EU is sued for 'facilitating copyright infringement'?*
- f) how long before the entertainment industries want this technology for themselves only, so they can charge people of the Earth for using it?*
- g) how long before the EU does the right thing and tells the entertainment industries to 'f\*\*k off'!!*

Une année après le début du projet P2P-Next, un usager poste en ligne le commentaire ci-dessus, à la suite d'un article qui annonce le lancement d'une nouvelle version de Tribler.

*« There they were, all the questions we didn't want to hear – and knew we were going to hear again and again as the project went on. So we just had to hold on to them<sup>528</sup>. »*

C'est en effet le résumé des conditions de succès pour Tribler et pour P2P-Next que vient de faire cet utilisateur, de manière pittoresque mais pertinente et complète, lors de l'entrée du projet dans sa phase centrale. L'innovation technique reste, certes, une question très importante pour l'originalité et la spécificité de leur démarche : il faut que le projet soit une véritable « alternative » aux systèmes existants, soit aux autres systèmes P2P de distribution de contenus, soit aux systèmes de streaming vidéo concurrents.

Cependant, il y a autre chose qui se dessine dans la réponse des usagers à Tribler, comme l'illustre bien l'extrait ci-dessus, qui trouve un écho dans les préoccupations de ses développeurs. La première année de vie du consortium l'a de plus en plus démontré : l'argument du « pur instrument » risque de refermer *a priori* les portes du succès pour Tribler. Les connotations politiques fortes dont le P2P est déjà doté auprès des institutions et du public européen sont impossibles à ignorer par les développeurs, à moins de se retrancher à nouveau dans la réalité universitaire « pure » des débuts. C'est une sorte d'agenda politique pour l'Europe et le P2P qui prend forme, faite d'équilibres entre les acteurs du marché, de législation en matière de droit d'auteur, de la responsabilité politique européenne d'assurer la concurrence. Autour de cet agenda, le public, les autorités nationales et internationales, les entreprises

---

<sup>528</sup> Entretien, F. Bierbach.

concernées ou potentiellement concernées, doivent pouvoir s'engager, se confronter, construire leurs propres définitions du P2P.

L'engagement politique et juridique avec son objet devient donc pour l'équipe Tribler une nécessité autant qu'une opportunité, et les premiers résultats officiels du projet le reconnaissent. Avec P2P-Next, ils sont désormais impliqués dans une double démarche : d'un côté, travailler à l'acceptation progressive du P2P en tant que technologie « légitime » de la part de l'industrie des contenus numériques, au delà des expérimentateurs qui font partie du consortium ; de l'autre, contribuer techniquement à l'innovation du cadre juridique européen relatif au P2P. Le déroulement de la suite du projet est étroitement lié à la validation du P2P non seulement comme modèle technique alternatif légitime, mais aussi en tant que « valeur » politique et économique pour l'Europe du prochain futur.

#### LE P2P, UNE ARCHITECTURE POLITIQUE ET ÉCONOMIQUE POUR LA TÉLÉVISION EUROPEENNE

La présentation de NextShareTV, la plateforme de diffusion de contenus en streaming vidéo décentralisé développée par P2P-Next, a lieu pour la première fois à l'*International Broadcasting Convention* (IBC), en 2009. À cette occasion, le prototype du décodeur qui constituera l'unité de connexion et accès au réseau européen de streaming vidéo en P2P est présenté ; il deviendra bientôt, dans les mots des différents partenaires de P2P-Next, l'objet-frontière (Star et Griesemer, 1989 ; Trompette et Vinck, 2011) autour duquel s'articulent les formulations et reformulations des objectifs des partenaires au sein du consortium – et, enfin, les définitions de système P2P qu'ils mobilisent dans leur démarche et la « valeur » de l'approche distribué et décentralisé dans leur vision.

Pour Lieven Vermaele, le directeur technique de l'UER (qui, on se rappelle, avait sollicité Tribler comme candidat possible pour la construction d'un réseau TV via Internet européen avant l'obtention du financement P2P-Next), la *personnalisation* du système que le décodeur représente est au centre de l'intérêt de la société de diffusion européenne. Le vocabulaire « social » des communautés et des interactions, caractéristique du premier Tribler, revient dans ses propos :

*“[The set-top box] create[s] a viable broadcast platform that would enable large audiences to stream and interact with live and on-demand (VoD) content [...] In addition, it is our intention to allow audiences to build communities around their favourite content via a fully personalized system”<sup>529</sup>.*

Mark Stuart de Pioneer Digital Design utilise un langage plus explicitement technique, qui reflète la position de directeur technique occupée par son entreprise

---

<sup>529</sup> Cet extrait et les trois suivants sont tirés des déclarations prononcées par différents partenaires de P2P-Next à IBC 2009.

dans le consortium. La *set-top box* occupe selon ses propres mots un rôle central pour l'*interopérabilité* du système ; celle-ci dépasse par ailleurs sa signification strictement technique pour qu'y soient reliées la facilité d'appropriation de la part de l'utilisateur, et l'efficacité d'un système qui revendique une approche P2P intégrale ; une priorité qui est partagée par George Wright de la BBC :

*“Pioneer recognises that consumers are demanding greater choice and quality content on demand, together with integrated devices that are easy to use. The next step is to agree Open Standards for interoperability between CE peer devices across Europe.”*

*“We [at BBC] are pleased to be able to inform industry of our results at IBC and help inform work towards creating an interoperable standard for Open Internet TV devices of the future.”*

C'est à Jari Ahola, *technology manager* au VTT *Technical Research Centre of Finland*, qu'il revient d'établir explicitement le lien entre deux « opportunités européennes », le financement de la Commission et un scénario européen très vivant en termes de recherche et développement dans le domaine de la distribution des contenus numériques. L'accent est mis dans ce cas sur NextShare et l'approche P2P à la télévision comme réponse à la convergence numérique et aux changements dans les modèles d'affaires du secteur médias :

*“[The NextShareTV prototype] combines European wide technology potential with significant European Commission funding and provides the means to create Internet based future television for global adoption. As the European television is going digital by 2012 we need to be prepared for the technical, legislative and regulatory challenges of digital convergence in time to lead the imminent change of the media distribution value chain.”*

De son côté, Friedrich Bierbach introduit dans le débat la question de la durabilité de l'infrastructure de l'Internet dans son état actuel, et de sa pertinence pour un scénario de diffusion pervasive de contenus numériques audio et vidéo. NextShare est dans ce cas la brique sur laquelle se fonde le système P2P, la porte d'entrée dans un réseau décentralisé qui promet de pouvoir mieux gérer les usages « non-linéaires » dérivant du déplacement « de l'approche collective et passive à l'approche personnalisée et active » à l'expérience de télévision<sup>530</sup>. C'est une discussion enflammée sur les choix d'architecture et sur l'importance d'un modèle Internet durable qui en découle, quand Friedrich suggère qu'une évolution fondamentale a déjà eu lieu dans les pratiques et l'appropriation de la variété de dispositifs qui permettent l'accès à la distribution de contenus audiovisuels, et qu'il s'agit pour l'industrie des contenus de le reconnaître et de construire des modèles alternatifs au lieu d'essayer de contrer ces tendances émergentes :

*“Already as we speak, TV set no longer has the monopoly of delivery of audiovisual content, the PCs, mobile phones, and potentially other new devices are all becoming increasingly important. So far this has been dealt by*

---

<sup>530</sup> Intervention de F. Bierbach à IBC 2009.

*the industries in ways that are, according to us, inadequate. There has been a proliferation of terminals and networks. This is not sustainable in the long run. We should focus on using a medium as the Internet that can be seen as unifying*<sup>531</sup>.”

Pourtant, selon Friedrich, ce ne sera pas l'Internet dans sa forme actuelle qui pourra atteindre ce but. L'infrastructure du « réseaux des réseaux » en l'état n'est pas adaptée à la transmission simultanée d'événements en direct à des millions de personnes, car un flux de données dédié doit être envoyé à chaque utilisateur. Avec des millions d'utilisateurs potentiels, ces flux de données transmis simultanément représentent un risque très élevé de congestion.

Une autre manière de contrer ce risque, qui s'adapterait mieux à la topologie de l'Internet dans sa forme actuelle, comme un participant à la conférence le note, est le modèle du *multicasting* : un système hybride où les flux de données sont distribués à une couche intermédiaire de serveurs locaux, qui redistribuent ensuite les contenus aux usagers. Pourtant, Friedrich ne voit pas dans ce modèle une alternative viable, et la raison réside encore une fois dans l'organisation actuelle de l'infrastructure de l'Internet : la plupart des routeurs IP de l'Internet ne prennent pas en charge le multicasting et à l'heure actuelle, il n'y a pas de motivations économiques pour que les FAI introduisent cette possibilité. La réponse « durable » à l'hétérogénéité de l'environnement Internet, destiné à devenir toujours plus articulé, serait donc un système qui permette d'ajouter des pièces à la mosaïque de manière relativement facile et ouverte. Au niveau technique, cela se traduit dans

*“a platform approach (that) allows modular development and modular applications, enables knowledge sharing and facilitates technology integration, code- and skill re-use*<sup>532</sup>”,

Ceci rappelle l'interopérabilité et l'intégration préconisées par le représentant de Pioneer.

Le projet de P2P-Next se révèle par ailleurs aussi profondément politique, au delà des déclarations délibérément telles qui semblent être le moment moins convaincant de l'intervention de Friedrich et qui reprennent à leur compte le langage assez artificiel des communiqués de presse institutionnels européens<sup>533</sup>. C'est en premier lieu le choix d'une architecture de réseau en pair-à-pair, représentée par le décodeur

---

<sup>531</sup> Id.

<sup>532</sup> Id.

<sup>533</sup> Par exemple lors d'un passage sur l' "effet démocratisant" de l'Internet: "To an increasing extent, the Internet is a part of the media supply used by a mass public to learn about the world around them. At the same time, people also increasingly use multimedia on the Internet as their means of cultural orientation. Both of these dynamics contribute to the 'democratising effect' of the Internet. The fact that the Internet is fulfilling this role is a positive development because in principal it is accessible worldwide and can also help in viewing issues from opposing standpoints. The Internet is also attractive to news distributors. Anyone can express his or her opinion and make it available cheaply to anyone else who is interested."

NextShare, qui est politique dans la démarche de P2P-Next : le consortium propose en effet que le déplacement d'un modèle linéaire de distribution à un modèle de plateforme centrée sur l'utilisateur ne se limite plus à la « couche supérieure » des interactions et des pratiques, mais devienne un principe véritablement intégré à la production de valeur effectuée au sein du réseau, à l'organisation de ses ressources matérielles, à l'allocation de ses coûts de distribution et de maintenance<sup>534</sup>.

#### VERS UNE CONTRIBUTION À L'ENCADREMENT JURIDIQUE DE LA TÉLÉVISION VIA P2P

En parallèle à sa démarche « orientée produit » centrée sur NextShare<sup>535</sup>, le consortium P2P-Next s'engage de plus en plus délibérément dans la recherche d'une approche qui, en plus du travail purement juridique et de gouvernance mené dans les arènes appropriées, puisse proposer et stabiliser une contribution à un cadre juridique pour le P2P – capable de dépasser les tentatives de régulation surplombants et assez maladroits qui l'ont, jusque à ce moment, caractérisé. Le contexte spécifique de l'application de cette technologie au streaming vidéo qui, au contraire, bénéficie d'une régulation plus précise et moins agressive au niveau européen, devient une arène d'expérimentation pour la régulation du P2P en tant que « plomberie » sous-tendant un système non-linéaire de distribution de contenus.

À partir de la deuxième année de vie du consortium, s'engage une discussion centrée sur l'intégration au système NextShare de la directive Télévision sans frontières, principal instrument de politique audiovisuelle de l'Union européenne, initialement votée en 1989 et ayant fait l'objet d'importantes mises à jour en 2007<sup>536</sup> :

*“The most significant changes in the new Directive, from our point of view, are the introduction of the concept of audiovisual services, the division between linear and non-linear audiovisual services, and the concept of a minimal regulation applying to both linear and non-linear audiovisual services. It also encourages co- and self-regulatory regimes in the fields that it covers<sup>537</sup>.”*

---

<sup>534</sup> “As broadband Internet becomes ubiquitous, all content distribution services will be combined and conveyed to the general public via a common “pipeline”, the Internet. (...)Also, P2P-Next is an overlay to the existing infrastructure as it is an application-layer media delivery system, which can in principle be overlaid to any communication system. Media delivery does not in fact, necessarily require a dedicated network. So, P2P Next does require much lower infrastructure investments, management costs, and maintenance costs compared to dedicated distribution networks, what they call ‘streaming farms’, such as the YouTube ones”.

<sup>535</sup> Voir le chapitre 2, section 1: il s'agit de mobilisations hybrides d'innovateurs, consommateurs, utilisateurs et entrepreneurs autour d'objets techniques et scientifiques complexes, pouvant avoir un rôle déterminant dans la construction et la diffusion de formes alternatives de culture matérielle (Hess, 2005).

<sup>536</sup> [http://europa.eu/legislation\\_summaries/audiovisual\\_and\\_media/l24101\\_en.htm](http://europa.eu/legislation_summaries/audiovisual_and_media/l24101_en.htm)

<sup>537</sup> Entretien, F. Bierbach.

Si la mobilisation du concept de « service audiovisuel » et la création de la catégorie « non-linéaire<sup>538</sup> » constituent un point de départ prometteur pour P2P-Next, plusieurs points précis de la directive sont susceptibles d'influencer la démarche du consortium. En premier lieu, la directive préconise que les systèmes de régulation pour les services de diffusion « linéaires » et « non-linéaires » aient un niveau de contrainte différent, plus léger pour les seconds, qui comprennent la vidéo à la demande. Deuxièmement, il est question d'un assouplissement des règles dans le domaine de la publicité télévisée, applicable au premier groupe de services mais pas au deuxième. Troisièmement, la directive spécifie pour le deuxième groupe le principe « du pays d'origine », qui établit qu'un canal de communication peut être régulé dans le pays à partir duquel il opère. Pour les fournisseurs de services de médias audiovisuels non-linéaires, comme certains partenaires de P2P-Next, cela pourrait être bénéfique : ce principe leur permettrait d'être soumis uniquement à la régulation du pays européen dans lequel ils résident, tout en explorant, avec un degré d'autonomie qui n'était pas possible auparavant, des régimes de co-régulation et auto-régulation pour les aspects collaboratifs du système.

L'enjeu central de la directive, du point de vue de P2P-Next, est qu'on y introduit des principes fondateurs de régulation de tous les services de médias audiovisuels, mais que l'on distingue le niveau de contrainte préconisé par la régulation en accord avec le degré de contrôle que l'utilisateur exerce sur le dispositif. L'Europe serait, avec la nouvelle directive, en voie d'admettre que les outils permettant un plus grand contrôle de la part de l'utilisateur, comme des nouveaux modèles de distribution personnalisée, pourraient bénéficier d'une régulation moins stricte. Pour Friedrich et son équipe, il s'agit là d'un principe très intéressant – et pas seulement pour le domaine spécifique du streaming : en tant que système de distribution et échange « centré sur l'utilisateur », le peer-to-peer dans ses multiples applications pourrait en tirer des leçons, et son encadrement juridique pourrait en être précisé. P2P-Next s'engage donc à proposer, au fil du développement de NextShare, une définition précise des composantes qui rendent le système « non-linéaire » et à expliquer comment le développement de service en jeu dans le projet peut entrer dans le cadre de la directive.

Pendant que le consortium cherche à obtenir des indications valables sur l'adaptation du cadre juridique du streaming au P2P, il négocie des questions bien plus « classiques » concernant ce dernier : notamment, le droit d'auteur et l'anonymat, qui inquiètent tant certains utilisateurs par rapport à la survie même du projet<sup>539</sup>. Pour Carlo, qui assure avec Friedrich la grande majorité des relations extérieures de l'équipe Tribler, le premier enjeu nécessite encore de beaucoup s'investir dans la recherche de solutions durables, ce que ni la pluralité de régimes de droit d'auteur existants, potentiellement contradictoires, ni les mesures techniques restrictives

---

<sup>538</sup> Définie dans la directive comme “any audiovisual media service where the user decides upon the moment in time when a specific program is transmitted on the basis of a choice of content selected by the media service provider”.

<sup>539</sup> Voir l'intervention citée au début de cette section, et les débats (p. ex.) à <http://torrentfreak.com/truly-decentralized-bittorrent-downloading-has-finally-arrived-101208/>

destinées à empêcher le trafic de contenus spécifiques n'ont su faire :

« En ce moment, on est dans une situation où ils se sont rendus compte que le copyright n'a plus beaucoup de sens en l'état. C'est ainsi que pour restreindre certains comportements, ils ont commencé à proposer d'autres voies, par exemple, le repérage des adresses IP des téléchargeurs pour la pédopornographie. Pour sûr, personne n'est d'accord avec la pédopornographie, et le législateur obtient automatiquement le consensus [pour mettre en œuvre la mesure restrictive]. Mais ce n'est pas ce que je définirais comme une manière très subtile de trouver une solution au problème<sup>540</sup>. »

Selon Carlo, le problème central est que ces mêmes acteurs qui devraient s'attaquer avec des alternatives plus structurelles et structurantes à la question – les distributeurs et les producteurs de contenus numériques – sont aussi ceux qui, à court terme et au vu du cadre de régulation actuel, risquent le plus à les expérimenter et les adopter. Dans ce contexte, le consortium P2P-Next a donc l'opportunité de se positionner comme une arène de « bricolage avec des solutions techniques et juridiques, qui commencent vraiment par les couches inférieures<sup>541</sup> » : par l'architecture du système, en expérimentant avec des dispositifs, des modèles d'affaires et des partenariats de recherche et développement inédits.

Enfin, l'équipe met en place une stratégie complémentaire à l'argument du « pur instrument » prédominant pendant les premières phases du projet ; en effet, celui-ci se révèle de plus en plus insuffisant à garantir la légitimité du système préconisé par le projet, face aux législations nationales les plus contraignantes. Si l'équipe Tribler est basée aux Pays-Bas, pays doté d'un cadre juridique notoirement libéral en la matière, il est parfois nécessaire pour d'autres partenaires de revoir leur rôle au sein du consortium, s'ils sont basés en Italie ou au Royaume-Uni<sup>542</sup>. L'Europe elle-même suggère parfois un changement de route sur des points particulièrement controversés :

« L'attention à l'anonymat, par exemple, n'est pas exactement parmi les aspects préférés de la Commission européenne... mais il s'agit là d'une question de priorités de recherche. En ayant le P2P comme objet de recherche, il est inévitable de procéder en parallèle sur plusieurs fronts, de l'anonymat à l'optimisation dans les échanges et la distribution de ressources. À un moment, il faut mettre sur la balance la volonté de sortir un bon produit et le fait de ne pas vouloir parler de – ou financer – quelque chose qui pourrait servir un usage illicite. Tout compte fait, ils nous ont laissés assez libres<sup>543</sup>. »

P2P-Next se propose donc comme une arène d'expérimentation juridique à un autre niveau : selon les contextes nationaux et les parties prenantes impliquées, il s'agit de

---

<sup>540</sup> Entretien, Carlo.

<sup>541</sup> Id.

<sup>542</sup> Entretien, F. Bierbach.

<sup>543</sup> Entretien, Carlo.

comprendre ce qui qualifie, et permet de considérer, un logiciel de distribution P2P comme un instrument « légal ». Une définition « opérationnelle » est proposée, en partant des objectifs du projet : est légal un dispositif qui, bien qu'en pouvant servir des usages illicites, se concentre en priorité sur le développement d'usages spécifiques qui, d'un côté, sont encore dans une zone d'incertitude par rapport au cadre législatif existant, et de l'autre n'ont pas encore été vraiment explorés et étudiés au niveau technique. Tout en étant une stratégie de survie pour l'équipe et pour le consortium dans son ensemble (« le résultat est que, par chance, on arrive à obtenir assez d'espace pour chercher des réponses<sup>544</sup> »), cette approche est aussi une occasion d'ouvrir la voie à une nouvelle forme d'engagement avec le *peer-to-peer*, où – face aux possibles « défis » que certains usages de l'objet technique pourraient représenter – on choisit d'en préserver les contextes d'innovation actuels, ainsi que les conditions d'ouverture pour que d'autres puissent surgir.

#### UNE « VALEUR », L'ARCHITECTURE DÉCENTRALISÉE ?

Au moment où j'écris cette thèse (et décide de manière relativement arbitraire, comme c'est toujours le cas, que mon terrain doit prendre fin pour le moment) P2P-Next est en train de passer le cap du mi-projet. L'équipe Tribler, dans son rôle de coordonnateur scientifique, continue à être au premier rang de la démarche de développement de NextShare, une démarche que Friedrich n'hésite pas à définir comme un « laboratoire constamment en cours de fabrication<sup>545</sup> ».

Les activités de l'équipe Tribler sont désormais étroitement liées à celles de P2P-Next. Les collaborations dans le consortium consistent *de facto* en des cycles de décentralisation de tâches particulières, dont la recentralisation s'articule autour de la stabilisation d'une fonctionnalité spécifique :

*« For P2P Next we have specific points of time where we have to deliver something, so a few months ahead of that, we figure out what we have to change and what features we can or want to concentrate on. We ask partners to submit their code, we get all contributions together and make the next step. Also, usually the releases are based on the stabilization of a particular feature: in the last case it was remote search and before that, the new interface<sup>546</sup>. »*

Au fur et à mesure que cette démarche est devenue plus complexe, la question de la taille du projet et de l'ampleur de la base d'utilisateurs recherchée s'impose comme un choix fondamental pour la petite équipe Tribler. En effet, en accord avec la stratégie qui prévoyait de lancer un produit « *out there* » de manière relativement rapide, le but de l'équipe avait été pour ses premières années de vie de ne pas viser

---

<sup>544</sup> Id.

<sup>545</sup> Entretien, F. Bierbach.

<sup>546</sup> Entretien, Hjalmar.

une masse importante d'utilisateurs : entre 2007 et 2009, leur nombre avait varié entre cinq et dix mille, avec des oscillations importantes à l'occasion d'épisodes spécifiques<sup>547</sup>. Ce nombre relativement limité d'utilisateurs donnait la possibilité aux développeurs de compter sur leurs seules forces pour assurer une certaine qualité de service, s'occupant en personne des nombreux suivis d'erreurs et des défaillances propres à un dispositif dans ses phases initiales de développement ; mais cela les obligeait à compter exclusivement sur des simulations et des modèles pour en extraire les données statistiques à large échelle nécessaires à certains choix de développement.

La présence dans P2P-Next donne par ailleurs l'occasion à l'équipe de mettre en place des mesures et des statistiques à large échelle sur l'environnement P2P de Tribler : c'est le moment, avec des moyens financiers et une base d'utilisateurs visée beaucoup plus importants, d'implémenter ces démarches de contrôle et de recherche de consensus sur les choix de développement qui n'avaient pas pu voir le jour au moment, par exemple, du débat sur la « visibilité » de l'architecture sociale (voir Section 2). Avec les Living Labs associés à P2P-Next – dans lesquels des volontaires se proposent pour tester le logiciel et le prototype de décodeur NextShare – les mesures du trafic P2P, et des habitudes des usagers P2P, sont « institutionnalisées<sup>548</sup> ». En retour, c'est une légitimation et institutionnalisation du P2P qui se mettent en place :

*« Most recently, we have actually had the chance to put our set-top box in people's houses and looking at what happens there. If we actually do it, we are more entitled to say, we bring innovation to the people with peer-to-peer<sup>549</sup>. »*

Le financement et le contexte d'application européens donnent donc une nouvelle dimension au compromis entre la gestion d'un nombre limité d'utilisateurs, qui permet assez facilement d'obtenir des modèles tout en restant approximatifs, et l'obtention d'une base d'utilisateurs plus large qui, tout en permettant des recherches plus détaillées sur les habitudes et les profils des utilisateurs, pose à son tour des problèmes. Un passage à l'échelle du logiciel Tribler signifierait, pour l'équipe de ses développeurs, de passer l'essentiel de leur temps à résoudre les bugs et problèmes quotidiens des utilisateurs, au lieu de progresser dans la recherche. P2P-Next ouvre dans ce cas la possibilité de profiter d'une « voie moyenne » entre projet universitaire et produit sur le marché, en faisant coexister auprès de différents partenaires la dimension de recherche et celle de maintenance d'un dispositif existant et fonctionnant<sup>550</sup>.

Le défi principal pour Tribler et pour P2P-Next reste cependant la définition de

---

<sup>547</sup> Par exemple, à l'occasion du premier lancement du lecteur vidéo en P2P; la nouvelle avait été relayée par TorrentFreak, ce qui avait été suivi de près par un redoublement des téléchargements du client P2P Tribler.

<sup>548</sup> Entretien, Hjalmar.

<sup>549</sup> Id.

<sup>550</sup> Entretien, Carlo.

l'approche décentralisée et *peer-to-peer* comme valeur ajoutée par rapport au modèle sous-tendant les dispositifs existants ; et en conséquence, le fait de travailler à son acceptation par le public. Il s'agit d'une démarche d'autant plus délicate à gérer que l'approche centralisée, dans certaines de ses implémentations existantes, satisfait déjà les besoins d'une très large majorité d'utilisateurs :

*« All in all, our main challenge is that we try to build on P2P what works on centralized platforms not exceptionally well, but still fairly well. YouTube, after all, works for a lot of people and serves many customers. In that sense, the technology scales, but for some aspect, it doesn't. Many broadcasters, with the Olympic games now, were just so happy that they could get 100,000 people to see the games over the internet. An impressive number, for sure, but we are 60 million, and that's only the Dutch! P2P is one of the technologies that can do it, but there will be trouble in convincing the people that it is necessary, because centralization works pretty well from their perspective already. In this respect, having living labs is great because it gives the public a chance for a hands-on approach<sup>551</sup>. »*

Des mots d'Hjalmar ressort, une fois de plus, la volonté de rendre « visibles » les manières dont les choix d'architecture technique et l'organisation des ressources matérielles d'un réseau influencent son efficacité, et le rendent plus ou moins apte à servir des usages particuliers ; les Labs sont l'endroit où la matérialité des « tuyaux » de l'Internet peut s'explicitier et être réappropriée par les usagers.

Si pour certains usages la valeur ajoutée de la décentralisation est moins évidente aux yeux de l'utilisateur (comme souligne Friedrich, dans son habituel style caustique, « *for data storage, it is just simpler to get a freaking USB key<sup>552</sup>* »), il est positif du point de vue de P2P-Next que pour l'adaptation du P2P au streaming, l'utilité du modèle décentralisé soit non seulement plus évidente d'un point de vue technique, mais qu'elle puisse aussi se traduire dans des bienfaits visibles pour l'utilisateur et les distributeurs : une gestion plus efficace de la bande passante dans l'ensemble du système, et des changements importants dans la rapidité des téléchargements directs. Carlo conclut à ce propos que

*« en général, c'est une épée à double tranchant qu'il faut gérer. Le P2P sert les entreprises de ce secteur, c'est clair d'après les résultats techniques, et en même temps, elles en ont une peur bleue. Il faut donc leur donner un minimum de garanties de la qualité de service et de sa légalité, tout en soulignant qu'on reste *open source* et que notre logiciel est modifiable par le premier *nerd* qui passe et qui a envie de s'y attaquer. Ne jamais oublier que c'est le bon, le très bon travail de recherche qui est et sera la motivation économique déterminante de leur collaboration avec nous ; sans ça ils ne vont même pas essayer. Pour le moment, ça marche. Pour le moment<sup>553</sup>. »*

---

<sup>551</sup> Entretien, Hjalmar.

<sup>552</sup> Entretien, F. Bierbach.

<sup>553</sup> Entretien, Carlo.

C'est clairement sur ce point que l'équipe Tribler compte s'appuyer dans la suite de son expérience de développement, pour définir la valeur de son approche et intéresser à la fois un nombre suffisant d'utilisateurs pionniers et d'entreprises/partenaires à la suite de P2P-Next. Il s'agira de continuer à faire ressortir la manière dont la décentralisation du modèle technique met en valeur l'expérience utilisateur, et ouvre la porte à des modèles d'affaires novateurs, davantage centrés sur cette expérience. En même temps, il s'agit pour l'équipe de continuer à négocier auprès des institutions européennes le statut de ses recherches sur un objet non seulement « légitime » mais utile ; et à poursuivre une validation continue des résultats internes de Tribler comme ce qui peut être saisi et réapproprié par des partenaires présents et futurs, en construisant sur la même infrastructure et l'adaptant aux nouvelles exigences d'un secteur aux changements très rapides.

## EN GUISE DE SYNTHÈSE

Tribler, un projet universitaire de « communauté sociale pour le partage de contenus en peer-to-peer », est créé en 2006 par une collaboration entre les universités de Delft et d'Amsterdam, aux Pays-Bas, et soutenu par des financements publics. S'inspirant du protocole de communication BitTorrent mais cherchant à combiner les mécanismes de ce dernier avec les « aspects sociaux des communautés en ligne », Tribler se focalise ensuite sur le streaming vidéo, après que des acteurs du secteur privé ont manifesté leur intérêt pour la rencontre entre un modèle de distribution en peer-to-peer et la télévision par Internet (P2PTV). C'est là le début d'un parcours qui débouchera sur l'attribution d'un financement de quatorze millions d'euros dans le cadre du septième programme-cadre de l'Union européenne. À partir de 2009, l'équipe Tribler est le chef de projet scientifique et l'un des principaux partenaires d'une coalition d'universités et d'entreprises nommée le P2P-Next Consortium, qui se donne un objectif de taille : construire le réseau européen du futur de la télévision par Internet, basé sur une technologie de streaming vidéo en pair à pair.

En racontant l'histoire de la poursuite de cet objectif, j'ai souhaité montrer dans ce chapitre comment les négociations qui ont parsemé le développement d'un prototype de logiciel P2PTV ont contribué à la problématisation politique et technique du peer-to-peer en tant que modèle alternatif pour les services audiovisuels par Internet, ainsi qu'à sa légitimation comme « valeur » dans le contexte européen. À cette fin, j'ai suivi les manières dont le dispositif est sorti de ses frontières académiques, pour saisir des institutions ou des espaces de collaboration entre recherche publique et secteur privé ; comment l'équipe de ses développeurs a par moments capitalisé sur son étiquette P2P, et a parfois dû s'en défendre, pour proposer une relecture juridique d'une technologie aux usages politiquement controversés ; comment ses créateurs ont proposé un modèle technique et économique à la fois différenciant et risqué, dans un contexte où le principal « client » est l'Union européenne avec ses puissants moyens financiers et sa machine bureaucratique et juridique très articulée.

Le chapitre s'est concentré dans un premier temps sur les débuts du client Tribler, créé à partir d'un intérêt commun des équipes de l'université technique de Delft et de l'université Vrije à Amsterdam pour le développement de « BitTorrent *with a social twist* ». Il a suivi comment l'équipe Tribler a cherché à répondre, par l'opérationnalisation de dynamiques « sociales », à certains des problèmes et fragilités classiques du P2P. Cet ensemble de choix socio-techniques prend forme, notamment, dans les « couches inférieures » de l'algorithme qui va constituer la base du logiciel Tribler depuis le lancement de sa première version publique, au début de 2007 : Buddycast. Cet algorithme marquera définitivement la prise de distance à la fois de BitTorrent et, plus particulièrement, du client ABC qui a servi de base au développement de Tribler : c'est le dispositif qui qualifie les pairs composant le système de *buddies*, compagnons et amis, stabilisant l'ajout de la dimension sociale aux nœuds de réseau P2P que sont les clients Tribler. Dimension dont l'importance et la visibilité seront par ailleurs remises en discussion lors de la première rencontre du logiciel avec la télévision via l'Internet, qui marque le premier déplacement de Tribler vers une spécialisation dans le partage et le streaming vidéo.

La deuxième grande phase de développement de Tribler – pendant laquelle le service commence à être reconnu et saisi par des usagers/clients, bien que demeurant à plus d'un titre un travail de recherche universitaire – sera marquée par une simplification du logiciel, au moins au niveau de l'interface ; une raréfaction d'une grande partie de ces mêmes dynamiques « sociales » qui en avaient marqués les premiers pas, et par une concentration des efforts de développement sur la fonctionnalité qui semble susciter le plus d'intérêt auprès de possibles partenaires commerciaux et de recherche. Dans la deuxième section de ce chapitre, j'ai donc vu comment nombre de fonctionnalités comme le téléchargement collaboratif, bien que continuant à faire partie du système comme mécanisme techniquement capable d'assurer la stabilité du système P2P, ne sont désormais plus destinées à être visibles par l'utilisateur, à moins que celui-ci ne le décide spécifiquement. Si le paradigme « *social-based* » n'est pas abandonné, il cesse de représenter la spécificité du logiciel aux yeux de l'utilisateur pour être relégué aux seules « couches inférieures » du dispositif. Le but principal de la démarche de développement devient la collocation de la fonctionnalité de streaming vidéo, placée au centre du dispositif ; la notion de réseau pair-à-pair elle-même passe au deuxième plan par rapport à celle de lecteur vidéo, avec des implications en termes de différenciation du « produit » Tribler par rapport à ses concurrents. Caché aux yeux des utilisateurs, Buddycast continue pourtant à constituer l'ossature du système et à imprimer le « P2P social » aux interactions qui y ont lieu. Le paradigme social des origines – la démarche de développement consistant à unir les deux approches de peer-to-peer et de réseautage social comme façon d'améliorer et enrichir l'expérience utilisateur – finit donc par se stabiliser comme un des principes fondateurs du logiciel ; qui tout en restant important, se fait plus discret. L'architecture reste « sociale » derrière une interface épurée.

En même temps que le « social » dans l'interaction architecture-interface du dispositif se voit reconfiguré, se dessine un autre choix important pour le futur du système : le fait de centrer les efforts de développement, ainsi que l'attention de l'utilisateur, sur la fonctionnalité de streaming vidéo et de vidéo à la demande. La nécessité de se différencier d'autres technologies de diffusion de contenus vidéo via Internet, également considérées par l'Union européenne de radio et télévision comme ses partenaires possibles, amène les développeurs de Tribler à redéfinir la spécificité du logiciel en s'appuyant sur son architecture décentralisée ; ce qui, à son tour contribue à une définition plus précise des avantages et des difficultés d'implémenter une fonction de streaming vidéo sur pair à pair. On a vu comment, au cours de 2009, ce recentrage progressif sur le streaming vidéo amène à Tribler plusieurs propositions de partenariats R&D centrés sur le développement de la télévision via Internet, tandis que les commentateurs techniques identifient, en majorité, dans la décentralisation la caractéristique qui pourrait constituer le véritable potentiel novateur de Tribler. À la mi-2009, le projet Tribler est impliqué dans plusieurs partenariats R&D, dont le principal sera établi dans le cadre du septième programme-cadre de l'Union Européenne : un important financement est attribué à l'équipe hollandaise et à une complexe coalition de partenaires, laboratoires scientifiques et entreprises – le consortium P2P-Next.

Avec ce projet, Tribler se trouve confronté à ses différentes « casquettes ». Celle de projet principalement financé par de l'argent public, mais dans lequel des partenaires

privés voient un intérêt, comme clé d'entrée dans un marché jusque là relativement inexploré. Celle de réseau, à la fois de personnes et de machines, pouvant supporter à moyen et long terme les modifications de l'offre et de la demande dans un secteur qui est en train d'être profondément bouleversé par les nouvelles pratiques de consommation et de distribution. Celle de système qui, construit sur une technologie aux usages fortement controversés dans les arènes politiques, en propose une relecture aux connotations, fortes et également « politiques », d'ouverture, de transparence et d'interopérabilité. La troisième section suit la manière dont la mise en place de P2P-Next – une coalition complexe et hétérogène mobilisant un objet de recherche et un produit controversé – met à l'épreuve la conception du peer-to-peer comme une technologie dont il faut, d'un point de vue politique et légal, se « défendre », pour se tourner vers sa construction comme « valeur » pour l'Europe. Cette démarche se confronte à nombre de défis, que l'on a étudiés dans cette section : d'un côté, les développements techniques ultérieurs de Tribler effectués dans le cadre de P2P-Next doivent tenir compte de la réponse du public européen à un projet d'expérimentation avec une technologie, comme la presse rappelle fréquemment, « aimée des pirates ». Des questions telles que la transition entre le travail expérimental sur le pair à pair, technologie souvent considérée comme incompatible avec l'économie marchande, et les modèles d'affaires existants, ou encore l'opérationnalisation de dispositifs destinés à assurer la « légalité » du futur système et sa compatibilité avec le cadre juridique européen, sont également en jeu. Si l'opportunité politique de mobiliser le pair à pair comme « valeur » finit par influencer la démarche des développeurs, l'organisation même du travail parmi les partenaires doit à son tour tenir compte à la fois de l'avancement du travail technique, et des nécessités « de marché » des partenaires privés du consortium.

Avec l'importance croissante du rôle de coordination et gestion scientifique assumé par l'équipe Tribler, celle-ci en vient à négocier de son côté, et une fois de plus, son statut de petite équipe universitaire fondée sur des logiques *open source*. Deux stratégies coexistent désormais au sein de l'équipe, répondant à deux « identités » différentes : les exigences de recherche au niveau local et l'ajustement progressif de l'algorithme Buddycast à l'architecture globale de P2P-Next. L'élaboration d'une définition plus complexe et articulée de la « légalité » du P2P à l'œuvre dans Tribler – au delà de l'argument du « pur instrument » qui peut être mobilisé par les utilisateurs à leur discrétion et sous leur responsabilité – devient par ailleurs nécessaire : l'équipe cherche donc à expliciter l'opposition entre le discours public prévalent sur le P2P comme technologie illégale ou « innovation interdite », et les caractéristiques techniques qui font du P2P une technologie particulièrement adaptée à l'échange de lourds contenus audiovisuels. La clé du succès du projet est identifiée à deux aspects qui, comme on l'a conclu dans cette section, engagent une vision et une problématisation politiques du P2P à un niveau bien plus profond. D'un côté, l'acceptation progressive du P2P en tant que technologie « légitime » de la part des industries des contenus, dont les partenaires de P2P-Next ne constituent qu'une avant-garde avec un penchant pour l'expérimentation ; d'autre côté, la nécessité d'un cadre juridique stable, dérivant de ou au moins inspiré par l'Europe, pour encadrer cette légitimation.

La quatrième et dernière section de ce chapitre s'est attachée à montrer comment, au

fil du déroulement du projet, ces deux aspects se sont articulés dans la démarche du consortium et de l'équipe Tribler en particulier ; ce faisant, on a souhaité donner quelques pistes d'analyse de la problématisation politique du pair à pair en tant que modèle alternatif pour les services Internet, et finalement, de sa légitimation comme « valeur » dans le contexte européen. Lors de l'entrée du projet dans sa phase centrale (sa troisième année d'existence), l'innovation technique reste au centre de la démarche, pour répondre à l'exigence d' « alternative » aux systèmes existants, soit d'autres systèmes P2P de distribution de contenus, soit des systèmes de streaming vidéo concurrents. Cependant, une autre chose se dessine dans la réponse des usagers à Tribler et dans les préoccupations de ses développeurs : l'argument du « pur instrument » risque de refermer *a priori* les portes du succès pour Tribler. Les connotations politiques fortes dont le P2P est déjà doté auprès des institutions et du public européen ne peuvent être ignorées par les développeurs, à moins qu'ils ne se retranchent à nouveau derrière la réalité universitaire « pure » des débuts. On fait l'hypothèse que c'est une sorte d'agenda politique pour l'Europe et le P2P qui prend forme à travers les activités du consortium, fait d'équilibres entre les acteurs du marché, de législation en matière de droit d'auteur, de responsabilité politique européenne visant à assurer la concurrence. Autour de cet agenda, le public, les autorités nationales et internationales, les entreprises concernées ou potentiellement concernées s'engagent, se confrontent, construisent leurs propres définitions du P2P.

Une de ces confrontations se passe lors de la première présentation de NextShareTV, la plateforme de diffusion de contenus en streaming vidéo décentralisé développée par P2P-Next. À cette occasion, le prototype du décodeur qui constituera l'unité de connexion et accès au réseau européen de streaming vidéo en P2P est présenté ; il deviendra bientôt, suivant les déclarations des différents partenaires de P2P-Next, l'objet-frontière autour duquel s'articulent les formulations et reformulations des objectifs des partenaires au sein du consortium. Il en vient également à représenter le point de confrontation et croisement des différentes définitions de la « valeur » de l'approche distribué et décentralisé, engageant tour à tour la personnalisation, l'interopérabilité, la durabilité de l'infrastructure de l'Internet dans son état actuel, et de sa pertinence pour un scénario de diffusion pervasive de contenus numériques audio et vidéo. Le projet de P2P-Next se révèle par ailleurs être aussi profondément politique – au delà des déclarations « délibéré » politiques – dans le choix même d'une architecture de réseau en pair-à-pair : le consortium propose en effet que le passage d'un modèle linéaire de distribution à un modèle de plateforme centrée sur l'utilisateur ne se limite plus à la « couche supérieure » des interactions et des pratiques, mais devienne un principe véritablement intégré à la production de valeur effectuée au sein du réseau, à l'organisation de ses ressources matérielles, à l'allocation de ses coûts de distribution et de maintenance.

Au sein du consortium, et dans ses interactions extérieures, se négocie par ailleurs la proposition d'une contribution au cadre juridique pour le P2P – capable de dépasser les tentatives de régulation surplombantes qui l'ont, jusqu'alors, caractérisé. Le contexte spécifique de l'application de cette technologie au streaming vidéo qui, au contraire, bénéficie d'une régulation plus précise et moins agressive au niveau européen, devient une arène d'expérimentation pour la régulation du P2P en tant que « plomberie » sous-tendant un système non-linéaire de distribution de contenus. On a

montré comment, pour P2P-Next, il s'agit de comprendre ce qui qualifie, et fait considérer, un logiciel de distribution P2P comme un instrument légal. Une définition « opérationnelle » est proposée, en partant des objectifs du projet : serait légal un dispositif qui, bien que pouvant servir des usages illicites, se concentre en priorité sur le développement d'usages spécifiques qui, d'un côté, sont encore dans une zone d'incertitude par rapport au cadre législatif existant, et d'un autre côté, n'ont pas encore été vraiment explorés et étudiés au niveau technique. C'est occasion, conclusion, d'ouvrir la voie à une nouvelle forme d'engagement avec le pair à pair, où – face aux « défis » possibles que certains usages de l'objet technique pourraient représenter – l'on choisit de préserver les contextes d'innovation actuels, ainsi que les conditions d'ouverture pour que d'autres puissent surgir. Il me semble par ailleurs pouvoir repérer, dans les formes et modalités de participation du secteur privé et de l'UE à l'aventure de Tribler et de P2P-Next, décrites dans ce chapitre, les conditions de félicité qui attesteraient de la « valeur » du P2P comme projet industriel sérieux, et d'un possible dépassement de la collocation exclusive de ce dernier dans l'« univers sombre » des limites de la loi et du passible de justice.

Le défi principal pour Tribler et pour P2P-Next reste cependant celui de stabiliser une définition de l'approche décentralisée et pair à pair comme valeur ajoutée par rapport au modèle sous-tendant les dispositifs existants ; et en conséquence, celui de s'assurer de l'acceptation de ce modèle par le public. Il s'agit d'une démarche d'autant plus délicate à gérer que l'approche centralisée, dans certaines de ses implémentations existantes, satisfait déjà les besoins d'une très large majorité d'utilisateurs. Une fois de plus, l'enjeu central de Tribler comme dispositif distribué – au moment où le projet P2P-Next bascule vers sa phase conclusive – est la visibilité, l'explicitation, l'appropriation par les utilisateurs des manières dont les choix d'architecture technique et l'organisation des ressources matérielles d'un réseau influencent son efficacité, et le rendent plus ou moins apte à servir des usages particuliers.



## CHAPITRE 7

### CONCLUSION GÉNÉRALE

*« I am going to insist on another track of this counter-reform movement. It is particularly important because, in a way, most of us gave it an implicit consent. It is the recentralization of computing, data storage and service provision. [...] Suddenly, we woke up with centralized social networks such as Facebook, centralized group and document management such as GoogleDocs and centralized media hosting such as YouTube or Megaupload. Whether we, as individuals, as groups, as neighbours, will be able to reclaim servers, services and data is the key issue. »*

Philippe Aigrain  
Conférence PARADISO, Bruxelles, 9 septembre 2011

Qu'est-ce que dessine une architecture de réseau décentralisée du point de vue de l'articulation des acteurs et des contenus, de la répartition de responsabilités, de l'organisation du marché et de la capacité à exercer du contrôle, des formes d'existence et des rôles d'entités telles que les nœuds du réseau, les usagers, les unités centrales ? Sous quelles conditions un réseau qui répartit la responsabilité de son fonctionnement à ses marges, et suivant un modèle non hiérarchisé ou hybride, peut-il se développer dans l'Internet d'aujourd'hui ? Il s'agit des deux questions de recherche principales que cette thèse a souhaité aborder, en suivant les développeurs de trois services Internet distribués – un moteur de recherche, un service de stockage et une application pour le streaming vidéo – , les collectifs d'usagers qui se développent autour et avec ces services, et ponctuellement, les arènes politiques où on discute de l'organisation et de la gouvernance de l'Internet à moyen et long terme.

Ce travail a souhaité explorer la constitution du P2P comme opportunité et alternative pour les services Internet, dans un scénario actuel où la norme pour les services hébergés par le « réseau des réseaux » est la concentration autour de serveurs donnant accès aux contenus. La thèse a suivi les représentations que les innovateurs dans le domaine des services Internet en P2P se font de leurs usagers, quels sont les collectifs actuels et futurs, les pratiques, les formes d'organisations qu'ils envisagent et qu'ils tendent à faire exister en lien avec les technologies qu'ils élaborent. Au moyen des observations de cas et de dispositifs, des analyses des contenus stockés, échangés,

recherchés, des entretiens avec les chercheurs et les entrepreneurs responsables des projets, mais aussi avec les parties prenantes impliquées dans le développement de l'architecture technique de l'Internet, j'ai essayé de retracer et suivre les choix, les façons de faire, les transformations et modifications des services Internet distribués, qui en déterminent les formes et caractéristiques actuelles, ainsi que leur évolution dans le temps<sup>554</sup>.

Les deux questionnements principaux se sont articulés, au fil des pages et des cas d'étude qui ont constitué la partie centrale de la thèse, en plusieurs enjeux et questions plus spécifiques.

Des nouveaux partages de compétences entre fournisseurs de service, producteurs de contenus, utilisateurs et opérateurs de réseau, prennent forme avec les services P2P. En analysant les formes d'engagement et d'intéressement dans les premiers pas de ces dispositifs – des utilisateurs pionniers en premier lieu, mais aussi des autres acteurs concernés par la mise en œuvre des services – j'ai donc cherché à qualifier les manières novatrices d'aborder, avec une approche d'architecture distribuée et décentralisée, des questions classiques associées à la gestion des contenus numériques et des données informationnelles en particulier, comme celles de la sécurité ou de la *privacy*. J'ai par ailleurs cherché à tirer de cette analyse des indications sur la viabilité technique et économique des services basés sur ce modèle de réseau.

En parallèle avec ces nouveaux partages de compétences, la thèse a suivi la mise en place de dynamiques et pratiques collectives qui prennent forme avec le développement du P2P. Selon des configurations variées, les services qui constituent mes cas d'étude font apparaître des nouvelles relations entre le local et le global, qui suggèrent à leur tour la mise en place de nouvelles formes de « politique par d'autres moyens » (Latour, 1988 : 229). La thèse a cherché à explorer les questions soulevées par ces reconfigurations en termes d'articulation entre l'individu et le collectif, dans l'espace du réseau et parfois dans l'espace géographique.

Enfin, cette thèse s'est intéressée aux manières dont l'architecture technique distribuée, et ses appropriations, contribuent à façonner le droit et les droits et s'en voient à leur tour reconfigurées. Dans cette perspective, j'ai vu surgir et j'ai analysé les enjeux juridiques posés par les configurations spécifiques du P2P dans le cadre des services Internet. La thèse s'est notamment intéressée aux solutions susceptibles d'émerger pour parvenir à des formes de régulation fiables, sans pour autant entraver les innovations présentes et futures – que ce soit du point de vue des droits de propriété, de la protection de la vie privée des utilisateurs/clients, ou de leurs droits d'accès aux contenus. La redéfinition de la notion de contributeur, à la fois de ressources informatiques nécessaires au fonctionnement du dispositif, et de contenus ; la reconfiguration des droits de l'utilisateur en terme d'accès, de relation avec les autres utilisateurs, de confidentialité de ses contributions ; le déplacement de la frontière entre usages privés et usages publics : autant de questions que les suivis de

---

<sup>554</sup> La plupart des conclusions exposées dans ce chapitre bénéficient, en plus de mon propre travail de terrain et de l'analyse effectuée au cours des chapitres précédents, du travail d'exploration des expérimentations et projets en matière d'architecture distribuée, mené au sein de la tâche 2 du projet ANR ADAM (Architecture Distribuée et Applications Multimédias).

cas ont contribué à identifier et qualifier.

Le reste de ce chapitre détaille les conclusions de cette thèse, présentées autour de trois axes.

Le premier axe a trait à la construction et la production du P2P comme alternative pour les services Internet: si ce terme se réfère en premier lieu au positionnement « minoritaire » des services qui se développent sur une architecture décentralisée par rapport à un modèle dominant dans le scénario actuel, je conclus que l'alternative P2P prend également forme avec l'interaction entre la diversité des usages, les étiquettes médiatiques et les contraintes et opportunités techniques du modèle distribué ; avec la décentralisation et la partielle recentralisation de composantes et dynamiques qui assurent le fonctionnement du réseau, faisant émerger des dispositifs techniques et des modèles économiques originaux ; et, enfin, avec les nouvelles intermédiations produites par les tentatives d'élimination des intermédiaires au sein du modèle.

Le deuxième axe suit le façonnage, la fabrication et la stabilisation de collectifs d'utilisateurs et de formes de lien social sous-tendus par l'architecture P2P, et conclut en particulier que l'on peut observer l'émergence de deux tendances transversales aux trois cas. Les développeurs cherchent à introduire un « paradigme social » dans le P2P, pour mettre à profit des caractéristiques liées aux préférences, aux profils, aux relations entre utilisateurs; démarche qui contribue à un rapprochement entre les gestes d'échange ou de partage effectués par les utilisateurs, et la mise en relation directe de deux ordinateurs échangeant des paquets de données dans les « couches inférieures » du réseau. Ce rapprochement entraîne un recentrage socio-technique du dispositif sur l'utilisateur, non seulement comme personne qui utilise le service mais en tant que co-fournisseur des ressources et des équipements informatiques sur lesquels se base le fonctionnement – l'existence même – des systèmes, et la qualité du service.

Le troisième axe discute le P2P appliqué aux services Internet comme objet du droit et producteur de droits. L'écosystème de régulation et de gouvernance de l'Internet qui entoure les services P2P – la définition de la technologie P2P, et de certains de ses usages, comme plus ou moins légaux – est susceptible d'affecter l'émergence de ces services et les formes qu'ils vont prendre, d'en façonner la conception et l'architecture. En même temps, les objets, les ressources, les flux de données inscrits et produits avec le P2P contribuent à façonner des conceptions particulières de la confidentialité des données, de la sécurité d'un réseau, du contrôle sur ses informations personnelles - et peuvent de ce fait être conçus et traités, à plein titre, en tant que systèmes de définition et protection des droits de l'utilisateur des services Internet.

## SECTION 1. *NAINS, GÉANTS* ET L' « ALTERNATIVE » POUR LES SERVICES INTERNET

*Voici donc une autre vision : [...]*

*Web 2.0 : Les contenus sont produits par les internautes et hébergés par des entreprises.*

*Web 3.0 : Les contenus sont produits et hébergés par les internautes.*

*Pensez-vous que la mutualisation des ressources basée sur le volontariat se généralisera au point de rendre les grandes communautés d'internautes plus puissantes (en terme de ressources) que les géants de l'Internet d'aujourd'hui? Enfin, pensez-vous que cette approche puisse être transposée en un modèle économique et social ?*

Anh-Tuan Gai, développeur P2P, 17 février 2007  
Blog Transnets, LeMonde.fr

Le pair à pair comme « alternative » pour les services Internet : c'est ce mot aux connotations de promesse – laissant entendre qu'il existe un « arbre de possibles » (Gai, 2007) dans les futures configurations du réseau des réseaux, suggérant une vision différente et différenciante pour les services Internet du futur – qui a tout d'abord retenu mon attention pour celui qui est devenu le sujet de cette thèse. Socle commun des projets et dispositifs qui ont constitué mes cas d'étude – et de bien d'autres, à des stades de développement très différents – ce terme se réfère, en première instance, au fait que tout service Internet qui se développe sur une architecture de réseau en P2P, ou distribuée, ou décentralisée, est minoritaire par rapport au modèle dominant actuel.

Si le principe de décentralisation a gouverné, depuis les origines de l'Internet, le design des transmissions et des communications qu'il autorise, les dynamiques socio-historiques dont j'ai tracé un portrait « révisionniste<sup>555</sup> » dans le troisième chapitre de cette thèse ont largement conduit à la diffusion d'architectures client/serveur, dans lesquelles les utilisateurs demandent des services à de puissants serveurs qui stockent de l'information et gèrent le trafic du réseau. Même si le trafic sur l'Internet fonctionne sur le principe de la distribution généralisée, il a donc désormais pris la forme d'une concentration autour des serveurs qui donnent accès aux contenus. L'architecture P2P s'affirme comme alternative dans ce scénario puisqu'elle est porteuse d'un autre projet possible par rapport à un modèle dominant.

---

<sup>555</sup> J'emprunte cette formule à [Minar & Hedlund, 2001].

Pourtant, ce bref rappel historique ne saurait résumer à lui seul les manières dont on a vu l'« alternative » P2P se construire et s'expliciter avec les projets analysés. Celle-ci naît de l'interaction entre la diversité des usages, les étiquettes médiatiques et les contraintes et opportunités techniques du modèle distribué ; de la décentralisation et la partielle recentralisation de composantes et dynamiques qui assurent le fonctionnement du réseau, donnant lieu à des dispositifs techniques et des modèles économiques originaux ; et, enfin, des nouvelles intermédiations que créent les tentatives d'« élimination des intermédiaires » (Elkin-Koren, 2006), totale ou partielle. L'alternative P2P naît de la redéfinition et la mobilisation des *nains* et des *géants* qui donnent le titre à cette thèse – des pairs et des serveurs, des marges et du centre du réseau respectivement ; elle a son fondement dans la « réclamation des serveurs, des services et des données » dont parle Philippe Aigrain dans l'exergue de ce chapitre.

#### L'ALTERNATIVE P2P : UNE QUESTION D'USAGES, DE TECHNIQUE ET DE « BUZZ »

Comme le deuxième chapitre de cette thèse a souhaité montrer, le P2P est non seulement devenu l'un des termes les plus largement discutés dans le domaine des TIC depuis sa révélation au grand public en 1999. Sous cette étiquette, faisant référence à un modèle de réseau décentralisé dont la responsabilité de fonctionnement est partagée parmi ses composantes, se regroupent une variété d'usages et d'applications qui se situent à différentes couches de l'Internet, de celle des interfaces et des pratiques jusqu'aux plus profondes et « invisibles » à l'utilisateur. En proposant un cadrage de ces usages, au moyen d'un modèle à trois niveaux et d'une discussion de la littérature pertinente, j'ai souhaité mettre en lumière comment les développeurs suivis au cours de cette thèse mobilisent et agencent les dynamiques et caractéristiques spécifiques à ces réseaux par rapport au modèle client/serveur. Plusieurs problématiques sont en jeu avec l'approche distribuée et décentralisée de l'architecture des services Internet : la conception d'un réseau informatique durable ; le fait d'assurer la stabilité d'un service dans une situation structurellement et techniquement instable ; et enfin, les définitions mêmes des ressources d'un réseau informatique et de l'utilisateur qui fournit ces ressources en même temps qu'il en bénéficie.

Tout en servant une variété d'usages différents, de la recherche au calcul à large échelle, du réseautage ou stockage, c'est l'utilisation et l'allocation des ressources du réseau – un réseau d'ordinateurs, de personnes, de contenus et d'informations – qui qualifie à un premier niveau l'alternative P2P. Mes cas d'étude ont montré les différentes manières dont s'agence un principe fondamental de l'architecture distribuée : les usagers n'utilisent pas seulement des ressources (de la bande passante, de l'espace de stockage, de la capacité de calcul) pour gérer leur présence sur le réseau, mais ils en fournissent également – de sorte que si la demande à laquelle le système doit répondre augmente, la capacité totale du système augmente aussi.

Les développeurs des systèmes P2P cherchent à profiter de la nature distribuée du système en termes de stabilité et d'endurance, pour améliorer sa force et empêcher

son invalidation totale en cas d'échec d'un des nœuds. C'est la valeur ajoutée de cette « redondance distribuée », par rapport à la capacité limitée et la facilité d'invalidation du modèle centralisé, qui amène les développeurs à s'engager dans l'alternative de l'architecture distribuée. En même temps, ces développeurs effectuent le choix de ménager la constante négociation de l'incertitude socio-technique qui en découle – considérer le P2P un avantage, en dépit du fait qu'un tel système est *a priori* plus complexe à mettre en œuvre, et sa stabilité au fil du temps plus délicate à assurer. Ce choix s'accompagne, de manière transversale aux cas d'étude, à la revendication d'une démarche originale par rapport aux acteurs dominants du secteur des services Internet, qui s'appuient sur le modèle client/serveur. La proposition d'une alternative dont la spécificité n'est pas due à la nature du service proposé (et donc à l'usage qui en sera fait), mais à l'architecture du réseau qui constitue l'ossature du service, est explicite – et patiemment, graduellement construite.

Les évolutions rapides dont le P2P a fait l'objet au cours des dernières années sont pourtant le résultat, non seulement de négociations sur l'efficacité technique des dispositifs, mais aussi de contraintes extérieures et plus explicitement « politiques » liées à l'identification de la technologie P2P avec le partage non autorisé de contenus protégés par le droit d'auteur. Si on peut attribuer aux conflits techno-juridiques entre développeurs, utilisateurs et ayants droit, centrés sur cette problématique, la succession de trois générations de dispositifs de partage de fichiers en P2P (Musiani, 2011), un trait commun des systèmes P2P qui ont marqué ces différentes générations est d'avoir été considérés comme une menace pour les intérêts de l'industrie des contenus – ce qui a déterminé une très large prévalence de représentations et discours spécifiques, voir partiels, autour du P2P, à la fois dans les médias et l'opinion publique et dans les priorités de recherche sur le sujet. L'alternative P2P pour les services Internet se définit donc non seulement par rapport à leurs homologues centralisés, mais aussi par rapport à une vision dominante – et partielle – de l'architecture P2P elle-même : les développeurs élaborent, en parallèle avec leurs stratégies de développement technique – voir en se servant de ce développement technique – des positionnements politiques et parfois médiatiques par rapport au P2P tel qu'il est majoritairement construit par les perceptions et représentations de potentiels financeurs, usagers, législateurs : c'est à dire, la « technologie du partage illégal ».

## REDÉFINIR LA DÉCENTRALISATION

C'est en adoptant la décentralisation comme pierre angulaire de leur démarche que les développeurs des services Internet en P2P se trouvent, souvent, à la redéfinir – et à renégocier, entre eux et avec les utilisateurs, les rôles qu'ils attribuent à des objets tels que le client P2P, les super-nœuds, les centres de données, les serveurs. Dans un scénario global de connexions asymétriques et de concentration des ressources, le P2P apparaît comme l'alternative ouvrant la voie à une plus grande disponibilité de ressources à un coût inférieur ; à l'utilisation optimale des ressources de la part de chaque pair connecté au réseau ; à un niveau de robustesse comparable à celui du modèle centralisé mais plus durable, puisque la demande de ressources est partagée

sur l'ensemble du réseau. La réélaboration de l'histoire du P2P que j'ai proposé dans le troisième chapitre de cette thèse a donc suivi comment, en cherchant les meilleures façons de contourner les problèmes spécifiques où l'architecture de l'Internet est durement mise à l'épreuve par les scénarios de « concentration » actuels, les développeurs se retournent vers l'Internet d'il y a quinze ou vingt ans: le premier Internet qui avait été créé en tant que système de communication entre machines de statut égal, partageant des ressources l'une avec l'autre.

Pourtant, dans la pratique quotidienne des développeurs, le P2P « pur » – celui qui préconise une structure de réseau complètement décentralisée – n'est presque jamais utilisé, au profit de solutions hybrides. Des choix architecturaux extrêmes ne sont que rarement la meilleure option quand il s'agit de construire un système utilisable : le développement de services Internet « sans un centre » est, le plus souvent, une histoire d'épreuves et de compromis entre l'idéal de la décentralisation et la négociation et exploration pratique de la recentralisation.

En retraçant les différentes démarches à l'œuvre dans mes cas d'étude – allant des efforts qui visent une décentralisation radicale jusqu'à la reconversion progressive à un modèle centralisé, en passant par des hybrides « peer-to-cloud » – la définition de décentralisation que j'ai vu se construire dans ces projets et dispositifs amène, en effet, à *repenser* ce qu'est un serveur, alors même que la démarche des développeurs prétend plutôt s'en éloigner le plus possible, afin d'augmenter la spécificité des dispositifs par rapport à leur rivaux centralisés plus célèbres et répandus. Le suivi des différentes modifications qui, au cours du temps et selon les applications, concernent la présence et le rôle de points de passage obligés dans les systèmes étudiés, révèlent que souvent, en partant des tentatives de se libérer complètement des serveurs, le seul moyen est parfois d'y revenir – tout de moins partiellement, pour mieux servir certaines fonctionnalités spécifiques. Cela a été tout particulièrement le cas de Wuala : la dépendance plus ou moins importante du système de stockage vis-à-vis le réseau P2P, et l'importance du recours additionnel à des serveurs, a été l'un des aspects les plus « négociés » au cours des premières années de vie du dispositif. J'ai montré, dans ce cas, comment le compromis entre une architecture centralisée, qui profiterait de la « simplicité » de la concentration et des points de passage obligés, et une architecture distribuée, reposant sur les ressources des usagers, arrive à un point où la frontière entre le modèle client/serveur et le modèle pair à pair semble s'effacer, se brouiller, jusqu'à questionner les définitions mêmes de pair et de serveur.

Le suivi des démarches de décentralisation et recentralisation à l'œuvre dans les cas étudiés peut amener à reconsidérer, symétriquement, la classique définition du serveur comme « l'ordinateur hébergeant un ensemble de logiciels, dont le rôle est de répondre de manière automatique à des demandes envoyées par des clients via le réseau ». Le serveur, ou le groupe de serveurs, se voit saisi, mis à l'épreuve, réapproprié, transformé. Des entités prennent forme au sein de ces systèmes qui agissent en tant que principale (et temporaire) sauvegarde de la sûreté des données ; nœud « central » ou principal, un nœud plus important des autres nœuds, qui – du fait qu'il reste en ligne constamment, ou régulièrement – se voit attribuer un rôle central dans la préservation du réseau ; ou encore, garantie de la stabilité de l'application pendant sa phase d'amorçage. L'analyse de la décentralisation de ces dispositifs, et

donc la spécificité de la plateforme technologique en question, montre que ce qui est à l'œuvre serait plutôt une irrigation ou une alimentation réciproque des deux modèles P2P et client-serveur ; ce qui suggère, par ailleurs, des possibles pistes pour repenser certains des éléments dits « classiques » de ces deux architectures.

« ÉLIMINATION DES INTERMÉDIAIRES », NOUVELLES INTERMÉDIATIONS

La complexité de la démarche de décentralisation des services Internet en pair à pair que je viens d'évoquer n'influence pas seulement l'organisation technique de ces dispositifs, mais se révèle l'enjeu central de la recherche de modèles économiques pour les services basés sur des architectures décentralisées.

Statut de Gwendal Simon (gwendal) sur  
Tuesday, 07-Apr-09 12:04:26 UTC

---



**Gwendal Simon**

**vient de dire à Francesca Musiani que P2P et économie marchande sont incompatibles**

Dès lors de mon deuxième entretien exploratoire pour cette thèse, Gwendal Simon, enseignant-chercheur en informatique spécialisé dans le P2P, me disait sans hésitation : « P2P et économie marchande sont incompatibles » (et, comme montre l'image ci-dessus, il a par la suite jugé bon de partager sa conviction avec ses *followers* Identi.ca<sup>556</sup>, donc, *de facto*, avec la totalité du Web). En effet, cette thèse semble confirmer en partie l'affirmation de Gwendal : bien de services P2P, y compris certains qui sont analysés ou mentionnés dans mon travail, n'arrivent pour le moment même pas à se poser la question de la durabilité de leurs modèles économiques. Pourtant, dans les cas où ils y arrivent, on voit que la question de l'intermédiation se pose à nouveau et avec force. Cette thèse a souhaité montrer comment la recherche de l'« élimination des intermédiaires » (Elkin-Koren, 2006), que le P2P semble *naturellement*, implicitement suggérer avec sa récusation de principe du modèle client/serveur, ouvre souvent la voie à d'autres formes d'intermédiation – qui, en définissant l'intermédiaire comme quelque chose de distribué, omniprésent, pervasif et invisible, ne l'efface pas pour autant.

Le troisième chapitre de la thèse a pu donner une relecture rapide et « orientée P2P » de l'histoire de l'Internet, à la lumière de la conception égalitaire qui avait marqué ses

---

<sup>556</sup> <http://identi.ca/notice/3274641>. Identi.ca est une version alternative, distribuée (ou, comme le dit son développeur en chef Evan Prodromou, « fédérée ») de Twitter. Le concept de base est une plateforme de microblogging dans le style de Twitter, où au lieu d'avoir une entreprise qui gère la totalité des serveurs de la plateforme, les utilisateurs peuvent gérer leur propre serveur, interopérable et fiable avec tout autre serveur basé sur le protocole laconi.ca.

origines – la construction d’un réseau horizontal où la communication par paquets, le protocole TCP/IP et la symétrie des connexions préconisaient la désintermédiation (Flichy, 2001) et l’égalité entre les nœuds, les utilisateurs, les machines. Ce chapitre a également suivi la montée en puissance d’un modèle vertical dans les mesures de gestion du trafic et d’organisation des flux de données : la hiérarchisation des acteurs, les points de passage obligés, le pilotage du trafic révèlent les limites pratiques d’un Internet de masse dépourvu d’intermédiaires – mais ne cachent pas pour autant que l’architecture de l’Internet se voit durement mise à l’épreuve, d’un point de vue technique, par les scénarios de concentration mis en avant, notamment, par le Web. Dans ce contexte, les architectures de réseau distribuées reprennent du poids : la différente gestion du trafic de l’Internet qu’elles proposent, et avec celle-ci, l’alternative aux monopoles économiques des grands acteurs des services Internet, incitent beaucoup de développeurs devenus entrepreneurs à s’inspirer des dynamiques de symétrie et d’égalité des débuts.

Naissent donc des systèmes où – on l’a suivi de plusieurs manières au cours des chapitres précédents – un nœud/utilisateur est à la fois un fournisseur et un demandeur de ressources informatiques (temps CPU, espace du disque dur, bande passante), ainsi qu’un producteur et un consommateur de contenus. Mais le panorama des services Internet qui en résulte, aussi *alternatif* et riche en expérimentations soit-il, n’est pas celui d’un réseau qui « fait l’économie des intermédiaires » (Méadel, 2011). Si l’entreprise tout-puissante, la grande ferme de serveurs qui centralise la totalité du trafic ne sont pas présentes, le défi du P2P « pur » est à son tour non seulement difficilement tenable d’un point de vue technique – donnant lieu aux formes de recentralisation dont il a été question plus haut – mais s’avère peu réaliste dans la recherche d’un modèle économique.

On a un exemple particulièrement éloquent des nouvelles intermédiations qui interviennent dans le distribué avec le cas de Wuala : le cœur du modèle technique de la distribution du stockage dans ce système – un utilisateur peut échanger l’espace inutilisé du disque dur de son propre ordinateur contre de l’espace de stockage sur le « nuage d’utilisateurs » – devient aussi le cœur de son modèle économique, apportant une valeur ajoutée qui consiste en une amélioration de la qualité de l’espace de stockage. En cherchant à s’assurer une certaine durabilité dans le temps des ressources matérielles et computationnelles à sa disposition, l’entreprise contribue à réintroduire un intermédiaire puissant, cherchant activement la stabilisation d’un ensemble d’usagers/nœuds qui soient « plus présents que les autres ». Une autre intermédiation est ainsi introduite, qui tient à la fois de la stabilisation technique et de l’optimisation économique : pour participer au programme du troc de l’espace disque – et donc, être inséré à plein titre dans l’approche distribuée et décentralisée du dispositif – l’utilisateur doit rester en ligne un certain temps par jour, temps où il contribuera par sa bande passante (qui doit être d’un niveau de qualité suffisant, une contrainte supplémentaire) au bon fonctionnement et à la stabilité du système.

Du fait qu’ils s’affinent progressivement et ils se construisent de manière « répartie sur les utilisateurs », avec des interrogations à un moteur de recherche ou des échanges de musique et vidéos, les protocoles épidémiques qui sous-tendent les algorithmes de Faroo et de Tribler n’en sont pas moins des formes de médiation et

d'intermédiation importantes et pervasives. À chaque dé clic sur un hyperlien Web, à chaque déclenchement de visionnage en streaming, ils contribuent au façonnage, à la définition et à l'organisation d'un monde parmi les plusieurs mondes possibles, à travers des choix et de la mise en avant de valeurs comme la confiance et la préférence commune. Plutôt que donner la priorité aux fichiers et aux contenus, ils connectent tout d'abord des usagers, leurs préférences similaires et leurs affinités ; les algorithmes collaboratifs qui en résultent comptent parmi les nouveaux médiateurs des architectures distribuées, en filtrant les résultats moins pertinents avec la comparaison périodique des caractéristiques de téléchargement des différents usagers – et en proposant une définition du « bon usager P2P » au passage. Cette thèse a souhaité contribuer au dévoilement des nouvelles intermédiations qui surgissent avec le distribué, le décentralisé, le P2P – modèles qui semblent en être par leur nature dépourvus. Comme le souligne bien Cécile Méadel,

Même lorsque l'on semble bien être dans une situation sans intermédiaire, avec une égalité entre tous les participants, une instance fut-elle distribuée et logicielle fait fonction de médiateur, c'est-à-dire qu'elle apparaît comme un de ces relais nécessaires qui font exister les choses, qui font que, sans eux, les choses n'existeraient pas. Les choses ou les services, les applications ou les usages (Méadel, 2011 : 28).

Ces nouvelles médiations répartissent autrement la charge du collectif, des réseaux comme « bien commun » (Massit-Folléa, 2008), en redéfinissant au passage les usagers des outils collaboratifs et leur droit à partager, à stocker, à communiquer : ce que les deux prochaines sections explorent plus dans le détail.

## SECTION 2. VERS UN P2P SOCIAL : REFORMATER LES RÉSEAUX, REPENSER L'USAGER

*We built a network out of a communications architecture design for peering which we defined in client-server style, which we then defined to be the dis-empowered client at the edge and the server in the middle. We aggregated processing and storage [...] in centralized places far from the human beings who controlled or thought they controlled the operation of the computers that increasingly dominated their lives.*

*What do we need?*

*We need a really good webserver you can put in your pocket and plug in any place. [...] It should know how to bring itself up. It should know how to start its web server, how to collect all your stuff out of the social networking places where you've got it. It should know how to send an encrypted backup of everything to your friends' servers.*

*It should know how to microblog. In other words, it should know **how to be you.***

Eben Moglen

Rencontres ISOC, New York, 5 février 2010

Des volontaires dans une mission de poursuite d'égalité, des partageurs de ressources informatiques, des contributeurs à un objectif mis en commun : que sont les nœuds et groupes de nœuds qui forment les systèmes P2P ? S'agit-il de « communautés imaginaires », de processus d'identification collective de groupes abstraits où « dans l'esprit de chacun vit l'image de leur communion » (Anderson, 1996) ? De grands collectifs anonymes réunis autour d'un instrument commun, pour lesquels l'agrégation n'est qu'un moyen d'obtenir plus rapidement et efficacement ce que l'on cherche ? Ou de dynamiques, liens et relations qui se situent quelque part entre les deux extrêmes de ce *continuum* ? En suivant les types de relations entre nœuds, machines et utilisateurs que l'approche distribuée et P2P au développement des services Internet veut rendre possible – ainsi que les manières dont les utilisateurs eux-mêmes ont discuté et se sont réappropriés de cette démarche de développement – un deuxième ensemble d'enjeux s'est dessiné dans mon analyse, qui a trait au façonnage, à la fabrication ou à la stabilisation de collectifs d'utilisateurs et de formes de lien social sous-tendus par l'architecture P2P.

En observant les types de contenus construits, le formatage des relations, les types d'échange qui ont lieu avec les dispositifs analysés, la thèse s'est intéressée aux représentations que les concepteurs de services Internet en P2P se sont faites de leurs

utilisateurs pionniers – aux collectifs actuels et futurs, les usages, les formes d'organisations qu'ils envisagent et qu'ils se proposent de faire exister avec les technologies proposées – ainsi qu'aux premières appropriations des services et à leur articulation avec les développements ultérieurs de ces systèmes.

La thèse a suivi l'émergence de deux préoccupations qui, en dépit des usages différents que les dispositifs doivent servir, sont transverses aux trois cas. D'un côté, l'introduction d'un « paradigme social » dans le P2P : il s'agit pour les développeurs de mettre à profit dans leur démarche des caractéristiques liées aux préférences, aux profils, aux relations entre utilisateurs. Souvent définie par les développeurs eux-mêmes comme l'introduction d'éléments plus proprement « sociaux » dans un modèle de réseau dont les applications pratiques plus célèbres sont délibérément anonymisées et anonymisantes, cette démarche contribue à un rapprochement entre les gestes d'échange ou de partage effectués par les utilisateurs, et la mise en relation directe de deux ordinateurs échangeant des paquets de données dans les « couches inférieures » du réseau. De l'autre côté, ce rapprochement implique souvent un recentrage socio-technique du dispositif sur l'utilisateur : non seulement comme personne qui utilise le service pour rechercher des informations, stocker des données, visionner un film sur son équipement, mais comme co-fournisseur des ressources et des équipements informatiques sur lesquels se base le fonctionnement – l'existence même – des systèmes et la qualité du service.

#### L'INTRODUCTION DU « PARADIGME SOCIAL » DANS LE P2P

Cette thèse a donc exploré les façons dont les concepteurs de services Internet en P2P s'attachent à y introduire des dynamiques et des caractéristiques qu'ils définissent « sociales ». Dans la plupart des applications P2P de partage de fichiers (le P2P « classique »), les nœuds du réseau sont des entités non reliées entre elles, et souvent délibérément anonymisées, bien que le modèle de réseau informatique qui les soutient permette *de facto* la mise en relation directe entre ces nœuds, le formatage de groupes, l'établissement de relations, la mise en place de notations qui concernent la réputation des usagers ou la qualité de leurs contenus (Hales, 2006). Ces choix de développement se sont fondés, pour les applications P2P de partage de fichiers, sur un ensemble de choix politiques et légaux qui ont amené les usagers de ces applications à avoir comme premier souci l'invisibilité, ou l'impossibilité d'être retracés – et les développeurs à concentrer leurs efforts sur les manières de leur fournir cette invisibilité.<sup>557</sup>

Cette approche a par ailleurs amené à négliger, tant en phase de développement que d'utilisation de ces dispositifs, une possible valeur ajoutée que dériverait de la mise à profit de mécanismes inspirés de dynamiques d'agrégation et de lien social, comme la création de groupes basés sur les affinités ou l'établissement de « réseaux d'amis » se fondant sur une attribution de confiance réciproque et explicite. Les projets et

---

<sup>557</sup> Je développe cet argument dans [Musiani, 2011].

dispositifs analysés semblent partager le souci de conceptualiser et façonner le P2P en prenant désormais en compte, et cherchant à modéliser, ces phénomènes sociaux.

C'est cette introduction du social dans le P2P qui s'explique tout au long du développement du moteur de recherche Faroo : son algorithme de découverte et présentation des pages repose sur l'importance que chaque usager attribue à une nouvelle page, importance rehaussée par le fait que d'autres utilisateurs partagent ce même intérêt. Le social se retrouve dans la conception des usagers en tant que « éclaireurs » de la fiabilité, de la pertinence et de la sélection des résultats du moteur de recherche, et dans la manière dont se définit la qualité de ces résultats, par une exploration « *bottom-up* » à la dimension collective. La présence d'un nombre suffisant d'usagers/éclaireurs partageant de l'information sur eux-mêmes est déterminante pour assurer que les résultats seront assez variés, nombreux et complexes, un miroir fiable des priorités des utilisateurs et de leurs intérêts communs.

Le rapprochement du P2P et du social se retrouve également dans le cas de Wuala : les développeurs de ce dispositif de stockage distribué font le choix de l'équiper d'une couche destinée à permettre le partage des contenus stockés entre les utilisateurs, de manière directe – directe aux deux niveaux de l'interface et de l'application, des pratiques et de l'architecture. La mise en relation directe de deux machines échangeant des paquets de données a une contrepartie intuitive dans le geste d'échange ou de partage effectué par les utilisateurs. La notion de réseau social et celle de clé d'amitié – l'autorisation d'échange qui doit être connue seulement par qui l'attribue et qui la reçoit, et qui permet de reconstituer le fichier – contribuent à changer le dispositif dès le niveau de l'architecture, en proposant une définition d'organisation sociale des utilisateurs qui a trait à la confiance réciproque et à l'ouverture d'une partie de ses ressources informatiques à l'usage de la part d'autres pairs.

C'est finalement et tout particulièrement le cas de l'équipe Tribler, pour qui le développement d'un « *BitTorrent with a social twist* » a été depuis le début l'un des objectifs affichés. L'algorithme Buddycast, qui est à la base du système, non seulement comporte un mécanisme de recommandation décentralisée, mais qualifie les pairs composant le système de *buddies*, compagnons ou amis : différents mécanismes d'inspiration sociale y sont testés et certains écartés, et continueront à informer le dispositif même quand, par choix des développeurs, il ne sera plus visible par l'utilisateur, pour continuer à être présent au niveau de l'architecture.

Que ce soit pour construire un index de recherche de manière collaborative, pour rendre le stockage et le partage de contenus numériques plus sûr ou plus direct, ou encore pour renforcer la stabilité d'un système de streaming vidéo, cette thèse a montré qu'une tendance de développement commune aux cas étudiés est l'élaboration d'un modèle de réseau capable d'intégrer le social aux deux niveaux de l'interface et de l'application – au niveau des interactions entre usagers et des connexions entre ordinateurs – en reproduisant les liens sociaux dans les liens d'architecture.

Dans cette lignée, la thèse a également suivi comment les projets étudiés s'attachent à (re)mettre l'utilisateur au centre des processus de création, distribution et modification des contenus ayant lieu avec ces applications – mais aussi de la gestion des ressources matérielles et d'infrastructure qui garantissent leur fonctionnement. Au fil des chapitres, on a vu apparaître nombre de situations qui illustrent cette dynamique. Wuala, logiciel de stockage et partage, est façonné de manière à ce que l'utilisateur devienne l'autorité de certification de ses informations personnelles, fournissant à ses contacts un accès direct à son profil et à son disque dur, tout en gardant ses informations de profil sur son propre terminal. Faroo et ses « six niveaux de distribution » sont programmés pour que les affinités entre les utilisateurs, et entre leurs interrogations au moteur de recherche, soient détectées de façon satisfaisante et l'information puisse être captée là où elle a son origine en tout premier lieu – l'utilisateur. Et afin d'implémenter des regroupements sociaux efficaces dans Tribler, ses développeurs choisissent d'éliminer l'anonymisation des pairs – ce qui n'implique pas pour autant que l'utilisateur doit figurer au sein du réseau avec son véritable nom, mais que ce nom sera choisi une fois pour toute et demeurera stable au fil du temps vis-à-vis du système et des autres utilisateurs, en tant que base sur laquelle construire réputation, fiabilité et préférences communes.

Les développeurs étudiés dans cette thèse cherchent à créer des algorithmes distribués qui puissent permettre à des phénomènes à l'échelle globale de se dégager de décisions individuelles, établir des liens pourvus de sens entre des millions d'utilisateurs et d'objets, et trouver la manière la plus efficace de capturer et tirer parti des affinités, des habitudes, des recommandations, ainsi que de faire face à des comportements « déviants ».

En même temps, le principe du *user first* contribue à situer cet usager au centre de la « durabilité économique » des dispositifs : les équipements informatiques des utilisateurs constituent l'infrastructure sur laquelle reposent ces services et leur qualité. Chaque utilisateur est un co-fournisseur de ressources techniques et matérielles : sa coopération est nécessaire pour le système formé par les autres usagers et par leurs machines, qui a besoin d'en connaître – et de mobiliser – la fiabilité, la stabilité et la réputation. Ce qui l'expose à des risques de *free-riding*, d'exploitation et de sur-utilisation, non seulement de la part des autres usagers, mais aussi par le fournisseur de service lui-même – comme le montre l'histoire de la graduelle reconversion de Wuala : d'un système distribué s'appuyant temporairement et par nécessité sur une poignée de serveurs, à un système basé presque intégralement sur des centres de données.

#### DU SOCIAL ET DES RESSOURCES : VERS UNE CENTRALITÉ DE L'USAGER

Les innovateurs du P2P appliqué aux services Internet opèrent ici un déplacement par rapport à celles qui ont été, dans les premières années du P2P comme phénomène de masse, les priorités de recherche et développement de ces outils. Les démarches de

développement suivies dans cette thèse sont guidées par l'attribution de centralité à l'utilisateur, en tant qu'individu, acteur dans un groupe, co-fournisseur et donneur de ressources au moment même où il les exploite. Cette priorité, qui comporte le plus souvent une identification explicite et univoque de l'utilisateur au sein du système, replace ou relègue au deuxième plan un ensemble d'autres choix socio-techniques – l'anonymat en tout premier lieu – que le droit, la politique, les contraintes techniques additionnelles ont, à d'autres moments, fait privilégier.

L'union de cette priorité et de la technologie P2P, considérée comme mieux adaptée aux besoins d'un usager en réseau, plus efficace et appropriée d'un point de vue technique, semble être l'une des dynamiques centrales en train de s'explicitier dans ces projets d'innovation. Par ailleurs, cette hypothétique « prochaine frontière » de la conception des services Internet<sup>558</sup> passe par la recherche de solutions à des problèmes non négligeables d'organisation des sauvegardes, de consommation d'énergie, de gestion des ressources informatiques, de disponibilité ou pénurie de celles-ci, de durabilité de l'écologie Internet. Les développeurs/entrepreneurs des services P2P doivent notamment se confronter à la nuance subtile, la « *thin line* » qui existe entre placer les ressources de l'utilisateur au cœur socio-économique du dispositif, et les sur-utiliser ou exploiter de façon inégale ou inappropriée par rapport à la philosophie du distribué. L'avenir « nuageux » de l'Internet, la montée en puissance de systèmes qui tendent à éloigner contenus et programmes des machines de l'utilisateur, posent également problème.

L'identification d'un certain nombre d'éléments fondateurs, communs aux cas étudiés, des portraits d'usagers et de collectifs d'usagers – mutualisation de ressources, volontariat, mise en commun de préférences et affinités – permet cependant de lire cette thèse comme étude d'un possible repositionnement des utilisateurs, des marges du réseau, au centre de l'environnement socio-économique des services Internet. Ce travail a également fourni quelques éléments d'interrogation de la notion d'*empowerment* des usagers (Brousseau, Marzouki & Méadel, 2012 : 8) permise par le développement des innovations P2P. Si un premier déplacement, d'un utilisateur/consommateur à un utilisateur/producteur, a pu se produire avec les pratiques et dynamiques regroupées sous le label « Web 2.0 », serait-il désormais question d'un utilisateur gestionnaire et hébergeur des contenus numériques qu'il crée et modifie ?

---

<sup>558</sup> Le « Web 3.0 », tel qu'il est défini par Anh-Tuan Gai, le développeur qu'on a cité en exergue de la première section de ce chapitre.

### SECTION 3. LE P2P : OBJET DU DROIT, PRODUCTEUR DE DROITS

*Users of peer-to-peer networks are not only making informational works available. They also convey their judgment regarding the relative value and relevancy of informational works. Distributed networks incorporate individuals' input in determining what to distribute and when. [...] This dimension of peer-to-peer networks has political significance. By incorporating the preferences of individual users, peer-to-peer networks turn individuals into active participants in the public sphere.*

*Since the virtues of peer-to-peer networks lie in their decentralized design, maintaining a decentralized architecture is essential for the purpose of promoting public welfare [but] there is no guarantee that this decentralized design will survive the current legal challenges.*

Niva Elkin-Koren  
“Making Technology Visible”,  
*NYU Journal of Legislation and Public Policy*, 2006

Niva Elkin-Koren est parmi les rares juristes à nous le rappeler : le droit n'est pas seulement une réponse aux nouvelles technologies, la conséquence d'un besoin de réforme de régimes juridiques existants – mais il façonne ces technologies et en influence la conception<sup>559</sup>. Les applications nombreuses et variées de la technologie P2P au services Internet en sont une illustration particulièrement intéressante, comme l'a montré le fil rouge qui, au cours de cette thèse, a suivi les mobilisations juridiques des architectures distribuées et de l'« étiquette » P2P : de la part des fournisseurs, des utilisateurs, des développeurs des services P2P, ainsi que des instances politiques faisant face aux choix de les financer ou d'en réguler l'implémentation. Symétriquement, la thèse a suivi les manières dont les objets, les ressources, les flux de données inscrits et produits avec le P2P peuvent eux-mêmes être conçus et traités, à plein titre, en tant que systèmes de définition et protection des droits de l'utilisateur des services Internet – en contribuant à façonner des conceptions particulières de la confidentialité des données, de la sécurité d'un réseau, du contrôle sur ses informations personnelles.

En tant que dispositifs qui impliquent une circulation de contenus numériques à

---

<sup>559</sup> L'articulation (et la co-construction) entre le technique et le juridique sont également explorées sous cet angle par Séverine Dusollier (2005, 2012), Danièle Bourcier (2010) et Mélanie Dulong de Rosnay (2007).

grande échelle, les services Internet en P2P étudiés dans cette thèse se définissent par rapport à nombre d'autres dispositifs, acteurs et dynamiques, qui sont classifiés, pris en compte... et limités par la loi de certaines façons. L'écosystème de régulation et de gouvernance de l'Internet qui les entoure est susceptible d'affecter l'émergence de ces services et les formes qu'ils vont prendre, d'en façonner la conception et l'architecture. La définition de la technologie P2P, et de certains de ses usages, comme plus ou moins « légaux » impacte les priorités de recherche et d'innovation autour de cette technologie, la manière dont elle est communiquée et mobilisée, jusqu'aux façons mêmes dont elle est développée et elle prend forme.

D'autre part, les applications P2P étudiées suggèrent des reconfigurations dans les emplacements des données et des échanges, façonnent différemment les frontières entre l'utilisateur et le réseau, mettent à disposition des utilisateurs un outillage techno-juridique pour se protéger, échanger, construire la légitimité de connaissances communes : avec le déploiement de ces dynamiques, sont en jeu l'attribution, la reconnaissance la modification et l'équilibre des droits d'utilisateurs et fournisseurs au sein des services.

#### LA « LÉGALITÉ » DU P2P : AU DELÀ D'UN USAGE, POUR DES NOUVEAUX USAGES

« *Yes, we are legal !* » « *Does legal P2P exist ? Yes, and rapidly picking up speed.* » Les efforts d'identifier et qualifier les systèmes qu'ils conçoivent comme des services « légaux » apparaît tout au long de cette thèse comme l'une des préoccupations communes des développeurs de services P2P, indépendamment des usages que ceux-ci doivent servir. La thèse a montré comment ces efforts naissent souvent d'une exigence de positionnement par rapport au P2P tel qu'il est majoritairement représenté : pour suggérer de façon immédiate qu'il y a autre chose à cette technologie que le partage de fichiers protégés par le droit d'auteur, son usage principal et considéré comme l'unique possible dans nombre d'arènes politiques et discussions dans les médias. Cela a été tout particulièrement évident pour Tribler, qui parmi les trois dispositifs étudiés est le plus saisi par des instances « traditionnellement » politiques.

Cependant, les cas étudiés ne cherchent pas à s'insérer pleinement (et à s'« effacer ») dans la définition de légalité mobilisée par les ayants droits de l'industrie des contenus numériques, mais construisent des définitions de légalité qui mettent l'accent sur ses aspects locaux, contestés, recherchés, ambigus – pour suggérer des pistes de régulation techno-juridique alternative pour les réseaux P2P et la variété d'objets qui y circulent. La définition de la légalité du P2P dans de nouveaux contextes d'application de cette technologie ne se résume pas à une tentative de se frayer un chemin dans un marché, celui des services Internet, qui ignore ou entrave le trafic P2P. Les trois cas étudiés dans cette thèse ont pu donner une occasion d'observer comment la mobilisation du P2P dans le domaine des services Internet ouvre des voies au dépassement et à la complexification du débat, assez réducteur mais dominant, qui porte à légiférer en matière de P2P en se focalisant sur les manières de « contraster » un de ses usages. La focale se déplace sur les

caractéristiques de la technologie, les raisons techniques et économiques qui plaident en faveur de son exploitation au sein d'une variété de dispositifs.

Dans un système comme Faroo, où les contenus créés et partagés sont les interrogations des usagers et l'index qui se construit à partir d'elles, la question des échanges de fichiers protégés se pose d'emblée avec moins d'acuité. Tout en revendiquant pour leur service une fonction de « pur instrument » – où les interrogations et leurs résultats, contribution des usagers au système, n'engagent qu'eux-mêmes d'un point de vue juridique – les développeurs abordent la question de la légalité du dispositif comme occasion de négocier et définir avec plus de précision ce que le système fait ou permet de faire. Facilitateur de la mise à disposition, de la recherche et de la localisation de contenus de la part de l'utilisateur, ce moteur de recherche P2P permet d'esquisser une notion de « responsabilité techno-juridique partagée » vis-à-vis du système et des autres utilisateurs, définissant ceux-ci et leur centralité dans le dispositif au passage.

Pour Tribler, un argument de « pur instrument » très semblable – la définition du dispositif comme outil de distribution techniquement efficace, dont la responsabilité d'éventuels usages illicites retomberait entièrement sur l'utilisateur fautif – s'avère insuffisant dès que le dispositif commence à être mobilisé au delà des frontières du laboratoire de TU Delft. Cependant, cela ouvre la voie à une démarche de validation et stabilisation d'une légalité du P2P qui a trait aux types de collaborations établies avec des partenaires institutionnels, à l'argent public dépensé pour financer le projet, à la variété de contenus échangés et échangeables avec le dispositif ; en contribuant à façonner un agenda politique pour l'Europe et le P2P, autour des discussions, des engagements et des confrontations de différentes définitions du P2P par le public, les autorités nationales et internationales, les entreprises concernées ou potentiellement concernées.

Si les développeurs de Wuala se sont tout d'abord limités à ne pas considérer la question comme prioritaire, on a vu comment, une fois de plus, les utilisateurs pionniers les plus engagés apportent une contribution fondamentale à une définition plus articulée de la « légalité » d'un système de stockage en cloud distribué qui prétend faire d'eux et de leurs ressources la pierre angulaire de son fonctionnement. La définition de la légalité du dispositif met cette fois l'accent sur la nature éclatée et encryptée des contenus stockés dans le système. L'enjeu n'est pas tant la manière dont les ressources partagées par tel ou tel autre ordinateur sont exploitées, de façon automatique, dans les opérations de fragmentation et d'encryptage : les fragments ne sont pas des objets saisissables par le droit dans sa configuration actuelle. L'enjeu réside plutôt dans les opérations de réseautage et partage entre les utilisateurs qui, stockant pour la première fois des contenus dans le système et les échangeant au moyen des clés d'amitié, introduisent ou recomposent la totalité d'un fichier sur leurs ordinateurs. Quant au fournisseur de service, la responsabilité de ce qui circule dans le système ne peut que très difficilement lui être attribuée, puisqu'il s'est lui-même mis dans la position d'en savoir le moins possible sur ces contenus, jusqu'à ne pas posséder les moyens techniques de les monitorer.

Dans les cas étudiés, la légalité du P2P prend la forme de fragments distribués dans le réseau, d'encryptage des contenus, de surveillance collaborative : caractéristiques

techniques, mais aussi pratiques des utilisateurs et défis d'application. Leur suivi suggère non seulement que d'éventuelles mesures légales visant à éliminer un certain type de trafic P2P pourraient comporter la perte ou l'endommagement des possibles bénéfices économiques et politiques du modèle, mais aussi que la pluralité de normes constituant ou pouvant constituer la « légalité » du P2P est souvent inhérente aux dispositifs techniques eux-mêmes. Dispositifs qui, constamment appelés à changer au rythme rapide du progrès technologique, pourraient être finalement plus endommagés que protégés par des mesures juridiques fondées sur une photographie instantanée de la technologie à un moment donné (Dusollier, 2012 ; Elkin-Koren, 2002, 2006, 2012).

#### LA PROTECTION « PAR LA TECHNIQUE » DE LA VIE PRIVÉE, ENTRE CONTRÔLE ET RESPONSABILITÉ

Les relations entre le droit de l'utilisateur à la *privacy* informationnelle et les technologies de l'information ont récemment donné lieu à un débat complexe et articulé, se poursuivant à la fois dans des contextes académiques (Lyon, 1988 ; Etzioni, 1999 ; Katyal, 2005 ; Rodotà, 2006 ; Tavani, 2008) et dans les arènes de débat public (Cavoukian, 2006 ; Katyal, 2009). Les services basés sur Internet impliquent des types spécifiques de licences, de contrats établis avec les utilisateurs finaux, et des procédures de collecte et gestion de données. Ceux-ci posent d'importants enjeux, à la fois en termes des concessions accordées aux entreprises fournissant le service, et des procédures et emplacements physiques du stockage des données ; certaines caractéristiques des utilisateurs de ces services s'inscrivent dans les dispositifs, et sont construites à travers eux.

Le deuxième et le troisième chapitres de cette thèse ont contribué à tracer le portrait du scénario de « concentration » dans lequel les dispositifs basés sur des architectures distribuées s'insèrent et trouvent leur place. Du point de vue de la gestion des données informationnelles, des données personnelles et des contenus échangés, cela signifie que le partage, le regroupement, le stockage de ces données dans les populaires services Internet d'aujourd'hui répondent à un modèle où le trafic est redirigé vers un ensemble de machines, placées sous le contrôle direct et exclusif du fournisseur de service. Les échanges entre utilisateurs s'effectuent donc en « copiant » les données que l'on souhaite partager sur un ou plusieurs terminaux extérieurs, ou en leur donnant la permission d'indexer ses informations ; la façon dont les données s'inscrivent, circulent dans ces machines et y sont maintenues est souvent incertaine, ou les droits que le fournisseur du service acquiert sur elles sont très importants par rapport à ceux préservés par l'utilisateur, de manière souvent opaque pour ce dernier. Cette thèse a souhaité explorer trois cas où les opérations de traitement de données en question ont lieu, en priorité ou en totalité, sur et entre les ordinateurs des utilisateurs, liés directement les uns aux autres. La thèse a suivi les manières dont ce choix de modèle contribue à façonner des définitions particulières de la protection de la vie privée de l'utilisateur, en modifiant les manières dont le contrôle sur les données informationnelles et la responsabilité de leur protection sont repartis sur les utilisateurs, sur les fournisseurs de service, sur les concepteurs qui l'ont développé.

En particulier, mes trois analyses de cas suivent des démarches qui, bien qu'avec leurs spécificités et différences, peuvent se reconduire à des tentatives de « définition par l'architecture » de la *privacy* : il s'agit de démarches qui attribuent un rôle de premier plan aux contraintes et aux opportunités de protection de la vie privée inscrites dans le modèle technique de choix des développeurs (Schaar, 2010). C'est le cas du modèle de distribution à six niveaux de Faroo, qui doit permettre d'éviter la traçabilité des interrogations de la part d'une entité centrale, et de faire en soi que les informations personnelles ne quittent pas le client P2P et le terminal de l'utilisateur, à moins qu'elles ne soient encryptées avant sur ce même terminal. Cette caractéristique permet aux développeurs de résoudre la tension, *a priori* difficilement conciliable, entre la confidentialité des informations personnelles et la personnalisation des interrogations de recherche – la valeur ajoutée que le « social » doit apporter au dispositif, et qui se base sur la collecte de ces informations. Le cas de Tribler est une autre occasion de suivre cette tension, car la logique à la base du système veut que l'histoire des téléchargements faits par un usager soit partagée avec les autres utilisateurs afin de nourrir l'algorithme de recommandation. La solution envisagée a encore une fois trait à une idée de « *privacy* par l'architecture », s'appuyant sur les caractéristiques techniques du modèle distribué et décentralisé pour mitiger l'impression d'exposition et de révélation de soi que le paradigme social risque de provoquer. C'est, enfin, le cas du modèle de stockage distribué de Wuala : le cœur de ce service est le terminal de l'utilisateur, où ont lieu, grâce au client P2P, les opérations d'encryptage et de division en fragments des fichiers stockés. L'encryptage en local d'abord, et la fragmentation après – avant toute opération de partage ou téléchargement qui implique une circulation des données dans le réseau – veulent être la preuve aux yeux des utilisateurs que le fournisseur de service, indépendamment de ses intentions, ne possède pas les moyens techniques d'enfreindre la confidentialité des données des utilisateurs.

Si les développeurs considèrent qu'une protection plus articulée de la *privacy* est l'un des avantages comparatifs majeurs de leurs systèmes, et la « vendent » comme tel, les utilisateurs s'interrogent sur les implications d'une architecture P2P pour la protection de leurs données. Des questions de fiabilité de l'infrastructure distribuée s'imposent donc à l'attention des concepteurs, et les forcent parfois à revoir leurs priorités de développement : il s'agit notamment d'éclairer en quoi le fait de mettre une partie des ressources informatiques de son ordinateur à la disposition de l'ensemble du réseau pose éventuellement des menaces pour les « invisibles » données qui y sont stockées. Dans les cas de Faroo et de Wuala, où le modèle en pair à pair croise, de façon insolite mais intéressante, une logique de logiciel propriétaire, ces questionnements sont l'occasion d'explicitier la difficile articulation entre la philosophie de décentralisation à la base des dispositifs, et la fermeture du code source, vue par les utilisateurs pionniers – dans la plupart des cas, des *power users* versés en informatique, ou « usagers/innovateurs » eux-mêmes – comme un manque de transparence, voire de respect, qui les empêche d'approfondir cet aspect avec les outils à leur disposition. C'est bien d'avoir de la *privacy* par la technique – semblent dire ces utilisateurs – mais encore faut-il que l'on puisse avoir une connaissance directe de cette technique, et éventuellement, la modifier.

La thèse a suivi comment les modèles de réseau distribués remettent en question « par la technique » l'étendue et l'ampleur des droits acquis par le fournisseur du service sur les données personnelles de l'utilisateur, par rapport au contrôle que celui-ci conserve sur ces données. Avec un compromis, un « *trade-off* » : si l'utilisateur voit sa *privacy* renforcée par la possibilité d'un contrôle augmenté sur ses données et les opérations du client P2P dont elles font l'objet, en même temps et pour les mêmes raisons, sa responsabilité des actions qu'il entreprend par moyen de l'application se voit augmentée, tandis que le fournisseur renonce de son plein gré à une partie de son contrôle sur les contenus qui circulent dans et par moyen du service qu'il gère. La dimension collective de cette responsabilité se voit aussi accentuée, dans la mesure où l'infraction au comportement collaboratif a des conséquences collectives aussi bien qu'individuelles – qu'il s'agisse d'un stockage de contenus inappropriés, de l'introduction d'informations non fiables ou de spam dans un index distribué, ou d'une gestion trop « égoïste » de la bande passante partagée par le système.

En conclusion, si la légalité du P2P ne peut pas se définir sans prendre en compte la pluralité des usages possibles de la technologie et la nature dialectique de la relation entre technologie et droit, il convient aussi de noter que les (ré)compositions du P2P au sein des systèmes analysés dans cette thèse sont elles-mêmes des systèmes de reconnaissance, attribution, modification et sauvegarde des droits des utilisateurs de ces services. On a vu cette dynamique s'explicitier dans une variété de démarches, qui vont des opérations de partage, stockage et communication initiées et menées par les utilisateurs sur leurs propres terminaux, jusqu'au partage des responsabilités entre les différents acteurs impliqués dans la gestion, la transmission et la circulation des informations et des contenus.

## **OUVERTURE. DES SERVICES INTERNET EN P2P À L' « INTERNET ALTERNATIF » : UN ENJEU DE GOUVERNANCE**

### INNOVATION ET P2P, UNE ÉTUDE DE GOUVERNANCE DE L'INTERNET ?

Le deuxième chapitre de cette thèse a pu en donner un premier aperçu : on reproche souvent au champ d'étude portant l'étiquette « gouvernance de l'Internet » de réunir des sujets et des concepts qui paraissent souvent trop vastes ou trop vagues – et au pire, non ou peu pertinents. Ces critiques ne préjugent certainement pas la validité d'un champ d'études qui non seulement est en train de se construire, mais s'intéresse à un objet particulièrement dynamique (Brousseau & Marzouki, 2012 : 368-371) ; toutefois, force est de constater que la littérature spécialisée dans le domaine de la gouvernance de l'Internet tend souvent à se concentrer exclusivement sur un nombre limité d'institutions internationales impliquées dans d'explicités discussions et débats sur la gouvernance globale de l'Internet. La qualification de « gouvernance de l'Internet » ne s'applique normalement pas à l'étude d'un grand nombre d'activités et pratiques, sur et par moyen de l'Internet, qui jouent pourtant un rôle de premier plan dans le formatage et la régulation du « réseau des réseaux » (van Eeten, 2009).

Les approches STS, comme celles adoptées dans cette thèse, peuvent aider dans ce contexte à se désengager d'une conception de l'Internet comme espace identifiable *a priori*, étranger aux forces institutionnelles de la « réalité » hors ligne, ou au contraire, entièrement retranché derrière les espaces codifiés de la politique traditionnelle (Cheniti, 2009). Ces approches permettent de mettre l'accent sur les mécanismes qui amènent les différents participants à la gestion technique, politique et économique de l'Internet à construire des connaissances communes, reconnaître certaines d'entre elles et pas d'autres comme « faits » de l'Internet, tracer des limites qui puissent réconcilier les préoccupations des experts et des usagers.

Dans cette lignée, je souhaite suggérer que cette thèse peut être comprise comme une contribution à l'étude de la gouvernance de l'Internet. Contribution qui cherche à aborder et faire croiser deux directions peu explorées dans ce champ : une qui s'attache à étudier la gouvernance de l'Internet « telle qu'elle se fait », celle qui, au sens de Michel Callon et Bruno Latour (e.g. 1990), agit, produit et s'organise ; et une qui ouvre l'enquête empirique à des lieux et des dispositifs « autres » par rapport à ceux sur lesquels le champ se concentre normalement, un nombre réduit d'institutions politiques internationales et supra-nationales. Cette thèse souhaite contribuer – ou plus précisément, être un possible prélude – à une ré-conceptualisation pragmatique du domaine « gouvernance de l'Internet », en tant que champ d'étude qui puisse englober avec (plus de ?) précision les procédures, dynamiques, actants et acteurs clés pour l'évolution et la mise en forme technique, économique et politique de l'Internet.

## POUR UNE ÉTUDE DE L' « INTERNET ALTERNATIF »

Pour la plupart des spécialistes de communication en réseau, l'étiquette « Internet alternatif » se réfère à l'appropriation par les mouvements sociaux des instruments de communication et des médias sociaux basés sur Internet, avec des buts de réforme et changement « par le bas » de l'ordre politique et social (Atton, 2005). Le mot « alternatif » est donc utilisé dans ce contexte pour indiquer ces pratiques qui, au moyen des nouveaux médias, contribuent à l'*empowerment* d'individus, groupes et organisations en proposant des canaux alternatifs pour la communication et la discussion de leur position dans les politiques locales et globales (Brousseau, Marzouki & Méadel, 2012 : 7-9). Très récemment, ce mot a été tout particulièrement mobilisé par les spécialistes du « printemps arabe » qui ont exploré dynamiques et instruments de communication dans ce contexte comme exemples de comment un Internet ouvert et libre est à la fois un instrument de pouvoir et une entité dont le futur est incertain.

Pourtant, encore plus récemment, l'appel pour un « Internet alternatif » est en train d'acquérir une signification plus profonde. C'est délibérément que je m'appuie ici sur le double sens de cet adjectif, car il est question dans ce cas d'une architecture différente pour le « réseau des réseaux », à partir de ses couches inférieures. Il s'agit des projets qui ont trait à une infrastructure alternative pour l'Internet, qui comportent des implications considérables pour l'étendue et la qualité du contrôle des utilisateurs sur leurs ordinateurs, leurs données et leurs échanges – finalement, pour les valeurs qui sous-tendent l'Internet dans sa globalité. Les instruments analytiques et l'approche mise en œuvre dans cette thèse peuvent fournir une base pour l'étude de cet objet complexe qui, à bien des égards, est le fils naturel des dispositifs analysés au fil de ces chapitres – et qui tient à la fois de l'imaginaire, du principe organisationnel et de l'artefact socio-technique.

Les discussions initiées par le « techno-anarchiste » Peter Sunde autour d'un serveur racine alternatif, capable de s'ériger à rival de l'Internet Corporation for Assigned Names and Numbers (ICANN) en hébergeant le registre des noms de domaine, de manière décentralisée, sur les ordinateurs d'utilisateurs volontaires ; le « Projet Kleinrock », un réseau coopératif de routeurs wifi domestiques qui puisse opérer sans l'intervention des fournisseurs d'accès à Internet (FAI) traditionnels ; Tor, le réseau mondial décentralisé de routeurs qui anonymise la transmission des flux TCP ; ou encore, Hackerspace Global Grid, un projet d'infrastructure de communication alternative, via satellites, destinée à un libre flux d'information, qui figure parmi les réponses plus exubérantes et originales au contesté Stop Online Piracy Act (SOPA) américain. Autant d'arènes où le projet de développement d'une « deuxième couche » de l'Internet, autonome et décentralisée, peut être saisi et ses implications approfondies.

L'histoire nous montre – par exemple, les discussions autour d'AlterNic, en 1997, révèlent que la proposition de Peter Sunde n'est pas le premier appel à un Domain Name System alternatif – qu'un facteur crucial pour le développement d'une infrastructure alternative pour l'Internet est l'appropriation par les utilisateurs, au

moment même où ils doivent fournir les ressources informatiques nécessaires à son fonctionnement. Comme cela a été le cas pour les dispositifs observés et analysés tout au long de cette thèse, non seulement les utilisateurs de l'Internet devraient, dans ce modèle d'infrastructure alternative, faire confiance au reste du réseau pour gérer une portion de leurs ressources logicielles et matérielles. Ils devraient aussi, dans les cas d'un DNS alternatif, dépendre des autres utilisateurs et de leurs ordinateurs pour les rediriger vers le nom de domaine approprié. Dès lors, des questions de recherche qui sembleront familières aux lecteurs de cette thèse, se posent : si les utilisateurs sont habitués à avoir confiance dans les serveurs DNS classiques, comme OpenDNS ou GoogleDNS, pour leur indiquer la bonne direction quand ils veulent accéder à un site web, qu'est-ce qui change quand ils doivent faire de même avec un ordinateur domestique parmi d'autres ? Sur quelles valeurs et caractéristiques techniques sera conçu et implémenté un réseau où les usagers accepteront de transformer leur équipement de connexion Internet en un routeur wifi parmi d'autres, pour le « bien commun » d'un Internet alternatif global ? Comment les acteurs politiques de la gouvernance globale de l'Internet prennent-ils en compte ces projets d'Internet alternatif et leur « potentiel de changement perturbateur et de surprise » (Rejeski, 2003) ?

À la fois pour son approche méthodologique et pour ses objets d'étude, cette thèse a commencé à proposer des pistes pour répondre à ces questions. Si la diversité et la multiplicité des enjeux qu'elle a dessinés ne plaident pas en faveur de la spécialisation de ce travail – ce qui en constitue d'ailleurs, peut-être, une limite – j'espère avoir apporté des réponses articulées, motivées et originales aux questions qui l'ont soutenu. Avec l'idée de contribuer à l'exploration de cette multitude de systèmes naissants qui – axés sur l'utilisateur et sur l'auto-organisation, principalement caractérisés par un développement et un système de contrôle décentralisés – proposent des alternatives aux infrastructures centralisées dominantes aujourd'hui<sup>560</sup>, sans pour autant exclure toute hiérarchie. Tel est, probablement, l'un des principaux apports de cette thèse : l'avoir montré comment le P2P répartit autrement la responsabilité entre les nains et les géants du réseau, tout en montrant les limites d'une partition trop nette entre le *top-down* et le *bottom-up*, les serveurs et les clients – le centre et la périphérie.

---

<sup>560</sup> Les travaux conduits récemment par l'université technique de Delft et la Next Generation Infrastructures Foundation ont amené à définir ces systèmes en termes de « *inverse infrastructures* », dispositifs qui « *display general patterns of emergence and development that are opposite in nature from those of large-scale infrastructures familiar to us today* » (Egyedi & Mehos, 2012 : 4). Voir aussi [Egyedi et al., 2009].

## BIBLIOGRAPHIE

- [Abbate, 2012] Abbate, J. (2012). L'histoire de l'Internet au prisme des STS. *Le temps des médias*, 18: 170-180.
- [Abbate, 1999] Abbate, J. (1999). *Inventing the Internet*. Cambridge, MA, The MIT Press.
- [Acquisti & Gross, 2006] Acquisti, A. et R. Gross (2006). « Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook », en P. Golle et G. Danezis (dir.), *Proceedings of 6th Workshop on Privacy Enhancing Technologies*, Cambridge, U.K: Robinson College, 36-58.
- [Adar & Huberman, 2000] Adar, E., et Huberman, B.A. (2000). Free riding on Gnutella. *First Monday*, 5 (10).  
(<http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/792/701>).
- [Agre, 2003] Agre, P. (2003). Peer-to-Peer and the Promise of Internet Equality. *Communications of the ACM*, 46 (2) : 39-42.
- [Aguillon & Cardon, 2007] Aguitton, C. & D. Cardon (2007). The Strength of Weak Cooperation: An Attempt to Understand the Meaning of Web 2.0. *Communications et Stratégies*, 65: 51-65.
- [Aigrain, 2010] Aigrain, P. (2010). Declouding Freedom: Reclaiming Servers, Services and Data. In 2020 FLOSS Roadmap (2010 Version/3<sup>rd</sup> Edition),  
<https://flossroadmap.co-ment.com/text/NUFVxf6wwK2/view/>
- [Aigrain, 2011] Aigrain, P. (2011). Another Narrative. Addressing Research Challenges and Other Open Issues session, *PARADISO Conference*, Brussels, 7–9 Sept. 2011.
- [Akrich, 1989] Akrich, M. (1989). De la position relative des localités. Systèmes électriques et réseaux socio-politiques. *Cahiers du Centre d'Études pour l'Emploi*, 32 : 117-166.
- [Akrich, 1992] Akrich, M. (1992). The De-description of Technical Objects. In Bijker, W. & J. Law (eds.), *Shaping Technology/Building Society. Studies in Sociotechnical Change*, Cambridge, MA : MIT Press, 205-224.
- [Akrich, 1993] Akrich, M. (1993). Les objets techniques et leurs utilisateurs : de la conception à l'action. In Bernard Conein, Nicolas Dodier et Laurent Thévenot (eds.), *Les Objets dans l'action*, Paris : Editions de l'EHESS, 35-57.
- [Akrich, 1998] Akrich, M. (1998). Les utilisateurs, acteurs de l'innovation, *Education permanente*, 134 : 78-89.

[Akrich, Callon & Latour, 1988] Akrich, M., Callon, M. & Latour, B. (1988). A quoi tient le succès des innovations ? Deuxième épisode : L'art de choisir les bons porte-parole. *Gérer et comprendre*, 12 : 14-29.

[Akrich & Latour, 1992] Akrich, M. & Latour, B. (1992). A Summary of a Convenient Vocabulary for the Semiotics of Human and Nonhuman Assemblies. In Bijker, W. & J. Law (eds.), *Shaping Technology/Building Society. Studies in Sociotechnical Change*, Cambridge, MA : MIT Press, 259-164.

[Akrich & Méadel, 2012] Akrich, M. & C. Méadel (2012). Policing exchanges as self-description in Internet groups. In E. Brousseau, M. Marzouki, C. Méadel (Eds.), *Governance, Regulations and Powers on the Internet*, Cambridge : Cambridge University Press.

[Anderson, 2001] Anderson, D. (2001). SETI@home. In A. Oram (Ed.), *Peer-to-peer: Harnessing the Power of Disruptive Technologies*, 67-76. Sebastopol, CA: O'Reilly.

[Anderson et al., 2002] Anderson, D.P., Cobb, J., Korpela, E., Lebofsky, M., & Werthimer, D. (2002). SETI@home: An experiment in public-resource computing. *Communications of the ACM*, 45 (11), 56-61.

[Anderson, 1996] Anderson, B. (1996). *L'imaginaire national. Réflexions sur l'origine et l'essor du nationalisme*, Paris : La Découverte.

[Appel, Boulanger et Massou, 2010] Appel, V., Boulanger, H., Massou, L. (eds., 2010). *Les dispositifs d'information et de communication. Concepts, usages et objets*. Bruxelles : De Boeck.

[Argenton & Prüfer, 2011] Argenton, C. & J. Prüfer (2011). Search Engine Competition with Network Externalities. *TILEC Discussion Paper*.

[Armstrong & Hagel, 1996] Armstrong, A., et Hagel, J. (1996). The real value of on-line communities. *Harvard Business Review*, 74, 134-141.

[Assadi & Beaudouin, 2002] Assadi, H. & Beaudouin, V. (2002). Comment utilise-t-on les moteurs de recherche sur Internet ? *Réseaux*, 6 (116) : 171-198.

[Atton, 2005] Atton, C. (2005). *An Alternative Internet*. Edinburgh, UK: Edinburgh University Press.

[Auray, 2012] Auray, N. (2012). Information communities and open governance: boundaries, statuses and conflicts. In E. Brousseau, M. Marzouki, C. Méadel (Eds.), *Governance, Regulations and Powers on the Internet*, Cambridge : Cambridge University Press.

[Auray, 2007] Auray, N. (2007). Le modèle souverainiste des communautés en ligne : impératif participatif et désacralisation du vote, *Hermès*, 47 : 137-144.

[Bacache-Beauvallet & Meinzel, 2012] Bacache-Beauvallet, M. & P. Meinzel (2012).

L'économie des réseaux pair-à-pair. *Working Paper*, présenté au séminaire ADAM, MINES ParisTech, 12 janvier 2012.

[Baccelli, 2005] Baccelli, F. (2005). Internet : modéliser le trafic pour mieux le gérer, *Interstices*, [http://interstices.info/jcms/c\\_12842/internet-modeliser-le-traffic-pour-mieux-le-gerer](http://interstices.info/jcms/c_12842/internet-modeliser-le-traffic-pour-mieux-le-gerer)

[Baird, 2002] Baird, Z. (2002). Governing the Internet : Engaging Government, Business, and Nonprofits. *Foreign Affairs*, 81 (6) : 15-20.

[Baker et al., 2002] Baker, M., Buyya, R., & Laforenza, D. (2002). Grids and grid technologies for wide-area distributed computing. *International Journal on Software: Practice & Experience (SPE)*, 32 (15), 1437-1466.

[Bal, Lohr & Reinefeld, 2002] Bal, H.E., Löhr, K.-P., & Reinefeld, A. (Eds.) (2002). *Proceedings of the Second IEEE/ACM International Symposium on Cluster Computing and the Grid*. Washington, DC.

[Balakrishnan et al., 2003] Balakrishnan, H., Kaashoek M.F., Karger, D., Morris, R., et Stoica, I. (2003). Looking up data in P2P systems. *Communications of the ACM*, 46 (2), 43-48.

[Bardini, 1996] Bardini, T. (1996). Changement et réseaux sociotechniques: de l'inscription à l'affordance. *Réseaux*, 76 : 125-153.

[Bardini, 2000] Bardini, T. (2000). *Bootstrapping: Douglas Engelbart, Coevolution, and the Origins of Personal Computing*. Stanford, CA : Stanford University Press.

[Barkai, 2001] Barkai, D. (2001). *Peer-to-peer computing. Technologies for sharing and collaboration on the net*. Hillsboro, OR: Intel Press.

[Bauwens, 2005a] Bauwens, M. (2005). P2P and Human Evolution: Placing Peer to Peer Theory in an Integral Framework. *Integral Visioning*. <http://www.integralworld.net/bauwens2.html>.

[Bauwens, 2005b] Bauwens, M. (2005). The Political Economy of Peer Production. *CTheory*. [www.ctheory.net/articles.aspx?id=499](http://www.ctheory.net/articles.aspx?id=499)

[Beaudouin & Licoppe, 2002] Beaudouin, V. & Licoppe, C. (2002). Présentation (Dossier « Parcours sur Internet »). *Réseaux*, 6 (116) : 9-15.

[Bender et al., 2005] Bender, M., S. Michel, P. Triantafillou, G. Weikum, C. Zimmer (2005). Improving Collection Selection with Overlap Awareness in P2P Search Engines. *SIGIR'05*, August 15–19, 2005, Salvador, Brazil.

[Benkler, 2004] Benkler, Y. (2004). Sharing Nicely: On Shareable Goods and the Emergence of Sharing as a Modality of Economic Production. *The Yale Law Journal*, 114 (2), 273-358.

[Bennett & Raab, 2003] Bennett, C. J. & Raab, C. D. (2003). *The Governance of*

*Privacy: Policy Instruments in Global Perspective*. Aldershot, UK: Ashgate.

[Bertier et al., 2010] Bertier, M., Y. Busnel & A. M. Kermarrec (2010). « Rumeurs, populations et communautés : équivalence uniquement sociologique ? Protocole de population versus protocoles épidémiques. » *12èmes Rencontres Francophones sur les Aspects Algorithmiques des Télécommunications* (Algotel 2010).

[Beuscart, 2002] Beuscart, J.-S. (2002). Les usagers de Napster, entre communauté et clientèle. Construction et régulation d'un collectif sociotechnique, *Sociologie du Travail*, 44 (4): 461-480.

[Beuscart & Peerbaye, 2006] Beuscart, J. S. & Peerbaye, A. (2006). Histoires de dispositifs (Introduction). *Terrains & Travaux*, 2 (11) : 3-15.

[Berners-Lee, 2010] Berners-Lee, T. (2010). Long Live the Web: A Call for Continued Open Standards and Neutrality, *Scientific American*, novembre 2010, <http://www.scientificamerican.com/article.cfm?id=long-live-the-web>

[Berners-Lee, 2000] Berners-Lee, T. (2000). *Weaving the Web*. New York: Harper-Business.

[Boczkowski & Lievrouw, 2008] Boczkowski P. & Lievrouw L. (2008). Bridging STS and communication studies: Scholarship on media and information technologies. In Hackett, E., Amsterdamska, O., Lynch, M., and Wajcman, J. (eds.) *The Handbook of Science and Technology Studies*. Cambridge, MA: The MIT Press.

[Bortzmeyer, 2008] Bortzmeyer, S. (2008). Le Pakistan pirate YouTube. *Blog de Stéphane Bortzmeyer*, février 2008, <http://www.bortzmeyer.org/pakistan-pirate-youtube.html>.

[Boyd, 2008] Boyd, D. (2008). Facebook's Privacy Trainwreck: Exposure, Invasion, and Social Convergence. *Convergence*, 14 (1).

[Boyd, 2004] Boyd, D. (2004). Friendster and Publicly Articulated Social Networks. *Conference on Human Factors and Computing Systems*. Vienna: ACM, April 24-29, 2004.

[Boyd & Ellison, 2007] Boyd, D. & N. Ellison (2007). « Social Network Sites: Definition, History, and Scholarship », *Journal of Computer-Mediated Communication*, 13 (1).

[Boyle, 2003] Boyle, J. (2003). The second enclosure movement and the construction of the public domain. *Law and Contemporary Problems*, 33, <http://www.law.duke.edu/journals/66LCPBoyle>

[Bourcier, 2010] Bourcier, D. (2010). Commons Digital Works : Thinking Governance. In Bourcier, D., Casanovas, P., Dulong de Rosnay, M. & Maracke, C. (eds.), *Intelligent Multimedia. Managing Creative Works in a Digital World*.

Florence : European Press Academic Publishing, 23-46.

[Bowker & Star, 1999] Bowker, G. C. & Star, S. L. (1999). *Sorting Things Out: Classification and Its Consequences*. Cambridge (MA), The MIT Press.

[Braman, 2011] Braman, S. (2011). Designing for Instability: Internet Architecture and Constant Change. *Media In Transition 7 (MIT7) Unstable Platforms: the Promise and Peril of Transition*, Cambridge, MA, May 13-15, 2011.

[Bredemayer & Malan, 2001] Bredemayer, D. & R. Malan (2001). Architecture Definitions, [http://www.bredemeyer.com/pdf\\_files/Definitions.pdf](http://www.bredemeyer.com/pdf_files/Definitions.pdf).

[Bricklin, 2001] Bricklin, D. (2001). The Cornucopia of the Commons. In A. Oram (Ed.), *Peer-to-Peer: Harnessing the Power of Disruptive Technologies* (pp. 59-63). Sebastopol, CA : O'Reilly.

[Brousseau & Marzouki, 2012] Brousseau, E. & Marzouki, M. (2012). Internet Governance: Old Issues, New Framings, Uncertain Implications. In Brousseau, E., M. Marzouki & C. Méadel (eds.), *Governance, Regulation and Powers on the Internet*. Cambridge : Cambridge University Press, 368-397.

[Brousseau, Marzouki & Méadel, 2012] Brousseau, E., M. Marzouki & C. Méadel (eds., 2012). *Governance, Regulation and Powers on the Internet*. Cambridge : Cambridge University Press.

[Bucchi, 2006] Bucchi, M. (2006). *Scegliere il mondo che vogliamo. Cittadini, politica, tecnoscienza*. Bologna : Il Mulino.

[Bursell, 2005] Bursell, M. (2005). Security and Trust in P2P Systems. In R. Subramanian et B. D. Goodman (Eds.) *Peer-to-peer computing: The Evolution of a Disruptive Technology*, 145-165. Hershey : Idea Group Publishing.

[Caillaud & Jullien, 2002] Caillaud, B. & B. Jullien (2002). Chicken & Egg: Competition Among Intermediation Service Providers. Working Paper, University of Toulouse.

[Caliskan & Callon, 2010] Çalışkan, K. & M. Callon (2010). Economization, Part 2 : A Research Programme for the Study of Markets, *Economy and Society*, 39 (1) : 1 - 32.

[Callon, 1986a] Callon, M. (1986). Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuç Bay. In J. Law (ed.) *Power, Action and Belief: A New Sociology of Knowledge?*, London : Routledge.

[Callon, 1986b] Callon, M. (1986). The Sociology of an Actor-Network: The Case of the Electric Vehicle. In Callon, M., Law, J. & Rip, A. (eds.) *Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World*, London: Macmillan Press, 19-34.

[Callon, 1991] Callon, M. (1991). Techno-economic Networks and Irreversibility. In Law, J. (eds.) *A Sociology of Monsters: Essays on Power, Technology, and Domination*, London: Routledge, 132-161.

[Callon, 2006] Callon, M. (2006). Sociologie de l'acteur-réseau. In Akrich, M., Callon, M. & Latour, B. *Sociologie de la traduction. Textes fondateurs*. Paris : Presses des Mines, 267-276.

[Callon, Lascoumes & Barthe, 2009] Callon, M., Lascoumes, P. & Barthe, Y. (2001). *Agir dans un monde incertain. Essai sur la démocratie technique*, Paris: Seuil.

[Callon & Latour, 1990] Callon, M. & Latour, B. (eds., 1990). *La science telle qu'elle se fait. Anthologie de la sociologie des sciences de langue anglaise*. Paris : La Découverte.

[Callon & Latour, 1986] Callon, M. & Latour, B. (1986). Les paradoxes de la modernité: comment concevoir les innovations? *Prospective et santé*, 36 : 13-25.

[Callon, Méadel & Rabeharisoa, 2002] Callon, M., C. Méadel & V. Rabeharisoa (2002). The Economy of Qualities. *Economy and Society*, 31 (2):194–217.

[Callon, Millo & Muniesa, 2007] Callon, M., Y. Millo & F. Muniesa (eds., 2007). *Market Devices*. Wiley-Blackwell.

[Cardon, 2010] Cardon, D. (2010). *La démocratie Internet. Promesses et limites*. Paris : Le Seuil.

[Cardon, 2008] Cardon, D. (2008). Le design de la visibilité. Un essai de cartographie du web 2.0. *Réseaux*, 26 (152) : 93-137.

[Cardon, 2006] Cardon, D. (2006). La trajectoire des innovations ascendantes : inventivité, coproduction et collectifs sur Internet. *Actes du colloque « Innovations, Usages, Réseaux »*, 17 et 18 novembre 2006.

[Castells, 2000] Castells, M. (2000). Toward a Sociology of the Networked Society. *Contemporary Sociology*, 29 (5), 693-699.

[Castro, 2001] Castro, M. (2001). Practical Byzantine fault tolerance. (<http://www.lcs.mit.edu/publications/pubs/pdf/MIT-LCS-TR-817.pdf>).

[Cavoukian, 2006] Cavoukian, A. (2006). Privacy by Design : The 7 Foundational Principles. Implementation and Mapping of Fair Information Practices. <https://www.privacyassociation.org/media/presentations/11Summit/RealitiesHO1.pdf>

[Cavoukian, 2009] Cavoukian, A. (2009). "Privacy by Design". The Answer to Overcoming Negative Externalities Arising From Poor Management of Personal Data. Trust Economics Workshop, Londres, 23 juin 2009.

[Cavoukian, 2010] Cavoukian, A. (eds., 2010) Special Issue: Privacy by Design: The Next Generation in the Evolution of Privacy. *Identity in the Information Society*, 3(2).

[Cheniti, 2009] Cheniti, T. (2009). *Global Internet Governance in Practice. Mundane Encounters and Multiple Enactments*. Unpublished DPhil Thesis, University of Oxford.

[Chernov et al., 2007] Chernov, S., P. Serdyukov, M. Bender, S. Michel, G. Weikum & C. Zimmer (2007). Database Selection and Result Merging in P2P Web Search. *DBISP2P 2005/2006, LNCS 4125* : 26–37.

[Christensen et al., 2006] Christensen, C. M., H. Baumann, R. Ruggles, & T. M. Sadtler (2006). Disruptive Innovation for Social Change, *Harvard Business Review*, December 2006.

[Christensen & Overdorf, 2000] Christensen, C. M. & M. Overdorf (2000). Meeting the Challenge of Disruptive Change, *Harvard Business Review*, March–April 2000.

[Clarke, 1991] Clarke, A. E. (1991). Social Worlds/Arenas Theory as Organizational Theory. In Maines, D. (eds.) *Social Organization and Social Process: Essays in Honor of Anselm Strauss*, Edison, NJ: Aldine Transaction, 119-158.

[Clarke, 2003] Clarke, I. (2003). Freenet's Next Generation Routing Protocol. (<http://freenetproject.org/>).

[Clarke et al., 2002] Clarke, I., Miller, S.G., Hong, T.W., Sandberg, O., & Wiley, B. (2002). Protecting free expression online with Freenet. *IEEE Internet Computing*, 6 (1), 40-49.

[Clarke & Star, 2008] Clarke, A. E. & Star, S. L. (2008). The Social Worlds Framework: A Theory/Method Package. In Hackett, E., Amsterdamska, O., Lynch, M., and Wajcman, J. (eds.) *The Handbook of Science and Technology Studies*. Cambridge, MA: The MIT Press, 113-137.

[Cohen, 2003] Cohen, B. (2003). Incentives Build Robustness in BitTorrent (<http://www.bitconjuror.org/BitTorrent/bittorrentecon.pdf>).

[Coleman, 2009] Coleman, Gabriella (2009). “Code Is Speech: Legal Tinkering, Expertise, and Protest Among Free and Open Source Software Developers”. *Cultural Anthropology*, 24(3): 420-454.

[Cotte & Després-Lonnet, 2004] Cotte, D. et Després-Lonnet, M., (2004), Le document numérique comme ‘lego®’ ou La dialectique peut-elle casser des briques ?, *Revue I3*, 4: 159-171.

[Crowcroft & Pratt, 2002] Crowcroft, J., & Pratt, I. (2002). Peer to peer: Peering into the future. *Proceedings of the IFIP-TC6 Networks 2002 Conference*, 1-19.

[Curien & Maxwell, 2011] Curien, N. & W. Maxwell (2011). *La neutralité d'Internet*. Paris: La Découverte.

- [Czarniawska, 2008] Czarniawska, B. (2008). *A Theory Of Organizing*. Cheltenham, UK: Edward Elgar Publishing.
- [Dauphin & Dagiral, 2005] Dauphin, F. & E. Dagiral (2005). P2P: From File Sharing to Meta-information Pooling?, *Communications & Strategies*, 59 (3) : 35-51.
- [de Certeau, 1980] de Certeau, M. (1980) *L'invention du quotidien (1, Arts de faire)*. Paris: Gallimard.
- [de Fornel, 1994] de Fornel, M. (1994). Le cadre interactionnel de l'échange visiophonique. *Réseaux*, 64 : 107-132.
- [DeNardis, 2009] DeNardis, L. (2009). *Protocol Politics : The Globalization of Internet Governance*. Cambridge, MA : The MIT Press.
- [Deleuze et Guattari, 1972] Deleuze, G. & Guattari, F. (1972). *Capitalisme et Schizophrénie 1. L'Anti-Œdipe*. Paris : Minuit.
- [Deleuze et Guattari, 1980] Deleuze, G. & Guattari, F. (1980). *Capitalisme et Schizophrénie 1. Mille Plateaux*. Paris : Minuit.
- [Delmas-Marty, 2012] Delmas-Marty, M. (2012). The Internet : Disrupting, Revealing and Producing Rules. In Massit-Folléa, F., C. Méadel & L. Monnoyer-Smith, (2012, eds.) *Normative Experience in Internet Politics*, Paris, Presses des Mines.
- [Denouël & Granjon, 2011] Denouël, J. & Granjon, F. (2011). *Communiquer à l'ère numérique : Regards croisés sur la sociologie des usages*. Paris : Presses de l'Ecole des Mines.
- [Diffie & Landau, 2007] Diffie, W., & Landau, S. (2007). *Privacy on the Line: The Politics of Wiretapping and Encryption, Updated and Expanded Edition*. Cambridge, MA: The MIT Press.
- [DiMaggio et al., 2001] Di Maggio, P., Hargittai, E., Neuman, W. R. & Robinson, J. P. (2001). Social Implications of the Internet. *Annual Review of Sociology*, 27, 307-336.
- [Dodier, 1995] Dodier, N. (1995). *Les Hommes et les Machines. La conscience collective dans les sociétés technicisées*. Paris: Métailié.
- [Douthwaite, 2002] Douthwaite, B. (2002). *Enabling Innovation: A Practical Guide to Understanding and Fostering Technological Change*. Londres: Zed Books.
- [Dulong de Rosnay, 2007] Dulong de Rosnay, M. (2007). *La mise à disposition des œuvres et des informations sur les réseaux: régulation juridique et régulation technique*. Thèse de doctorat en droit, Université Panthéon-Assas, Paris, France.
- [Dusollier, 2012] Dusollier, S. (2012). DRM at the intersection of compyright law

and technology : a case study for regulation. In Brousseau, E., M. Marzouki & C. Méadel (eds). *Governance, Regulation and Powers on the Internet*. Cambridge : Cambridge University Press, 297-317.

[Dusollier, 2005] Dusollier, S. (2005). *Droit d'auteur et protection des œuvres dans l'univers numérique : Droits et exceptions à la lumière des dispositifs de verrouillage des œuvres*. Bruxelles : Larcier.

[Edwards, 2002] Edwards, J. (2002). *Peer-to-peer programming on Groove*. Indianapolis, IN: Addison-Wesley.

[Egyedi & Mehos, 2012] Egyedi, T. M. & D. C. Mehos (eds., 2012). *Inverse Infrastructures : Disrupting Networks From Below*. Cheltenham, UK : Edward Elgar Publishing.

[Egyedi et al., 2009] Egyedi, T. M., D. C. Mehos & W. G. Vree (2009). New Perspectives on Inverse Infrastructures. *2009 Second Conference on Infrastructure Systems and Services : Developing 21st Century Infrastructure Networks (INFRA)*, 1-7.

[Elkin-Koren, 2002] Elkin-Koren, N. (2002). It's All About Control: Rethinking Copyright in the New Information Landscape. In Elkin-Koren, N. & Netanel, N. W. (eds.) *The Commodification of Information*, Kluwer Law International, The Hague.

[Elkin-Koren, 2006] Elkin-Koren, N. (2006). Making Technology Visible: Liability of Internet Service Providers for Peer-to-Peer Traffic. *New York University Journal of Legislation & Public Policy*, 9 (15), 15-76.

[Elkin-Koren, 2012] Elkin-Koren, N. (2012). Governing Access to User-Generated Content : The Changing Nature of Private Ordering in Digital Networks. In Brousseau, E., Marzouki, M., Méadel, C. (eds.), *Governance, Regulations and Powers on the Internet*, Cambridge : Cambridge University Press.

[Elkin-Koren & Salzberger, 2004] Elkin-Koren, N. & E. M. Salzberger (2004). *Law, Economics and Cyberspace*. Cheltenham, UK : Edward Elgar Publishing.

[Ertzscheid, 2007] Ertzscheid, O. (2007). « Le web implicite ». *Affordance.info*. [http://affordance.typepad.com/mon\\_weblog/2007/06/le-web-implicit.html](http://affordance.typepad.com/mon_weblog/2007/06/le-web-implicit.html)

[Etzioni, 1999] Etzioni, A. (1999). *The Limits of Privacy*. New York, NY: Basic.

[Evans, 2007] Evans, D. G. (2007). Human Rights: Four Generations of Practice and Development. In A. Abdi & L. Shultz (Eds.), *Educating for Human Rights and Global Citizenship* (pp. 25-38). Albany, NY: SUNY Press.

[Everett-Church, 1999] Everett-Church, R. (1999) The Spam That Started It All, *Wired*, <http://www.wired.com/politics/law/news/1999/04/19098>

[Farchy, 2003] Farchy, J. (2003), *Internet et le droit d'auteur : la culture Napster*. Paris : CNRS Editions.

- [Figueiredo *et al.*, 2008] Figueiredo, R. J., Boykin, P. O., St. Juste, P. and Wolinsky, D. (2008). Social VPNs: Integrating Overlay and Social Networks for Seamless P2P Networking. *Proceedings of the 2008 IEEE 17th Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises*. Washington, DC: IEEE Computer Society.
- [Flichy, 2001] Flichy, P. (2001). *L'imaginaire d'Internet*. Paris: La Découverte.
- [Flichy, 1995] Flichy, P. (1995). *L'innovation technique. Récents développements en sciences sociales : vers une nouvelle théorie de l'innovation*. Paris: La Découverte.
- [Floridi, 2005] Floridi, L. (2005). The Ontological Interpretation of Informational Privacy. *Ethics and Information Technology*, 7, 185-200.
- [Flyverbom, 2011] Flyverbom, M. (2011). *The Power of Networks : Organizing the Global Politics of the Internet*. Cheltenham, UK : Edward Elgar Publishing.
- [Foster, 2002] Foster, I. (2002). Grid: A New Infrastructure For 21st Century Science, *Physics Today*, 2002.
- [Foster & Iamnitchi, 2003] Foster, I. & A. Iamnitchi (2003). On Death, Taxes, and the Convergence of Peer-to-peer and Grid Computing, *Proceedings of IPTPS'03*.
- [Foster & Kesselman, 2004] Foster, I., & C. Kesselman (Eds., 2004). *The grid: Blueprint for a new computing infrastructure* (2nd edition). San Francisco: Morgan Kaufmann.
- [Foucault, 1994 (1977)] Foucault, M. (1994, or. 1977). Le jeu de Michel Foucault. In Defert, D., Ewald, F. & Lagrange, J. (eds.) *Dits et écrits 1954-1988. Vol. III : 1976-1979*, Paris : Gallimard, 298-329.
- [Franceschet, 2010] Franceschet, Massimo (2010). "PageRank: Standing on the shoulders of giants". <http://arxiv.org/abs/1002.2858>
- [Frickel, 2004] Frickel, S. (2004). Just Science?: Organizing Scientist Activism in the U.S. Environmental Justice Movement, *Science as Culture*, 13 (4) : 449-469.
- [Frickel & Gross, 2005] Frickel, S. & Gross, N. (2005). A General Theory of Scientific/Intellectual Movements, *American Sociological Review*, 70 : 204-232.
- [Frosini, 1998] Frosini, V. (1998). *Informática y Derecho*, Bogotá, Editorial Temis S.A.
- [Fujimura, 1992] Fujimura, J. (1992). Crafting Science: Standardized Packages, Boundary Objects, and 'Translation'. In Pickering, A. (eds.) *Science as Practice and Culture*, Chicago: University of Chicago Press, 168-211.
- [Fuller, 2008] Fuller, M. (2008, Eds.). *Software Studies: A Lexicon*. Cambridge, MA:

The MIT Press.

[Gai, 2007] Gai, A.-T. (2007). Web 3.0: une autre branche pour l'arbre des possibles, *Transnets*.

[Ganascia, 2009] Ganascia, J.-G. (2009). *Voir et pouvoir : qui nous surveille ? Un essai sur la sousveillance et la surveillance à l'ère de l'infosphère*, Paris, Editions du Pommier.

[Gensollen, 2009] Gensollen, M. (2009). Le web relationnel : vers une économie plus sociale ? In F. Millerand, S. Proulx & J. Rueff (eds.), *Le Web relationnel, mutation de la communication ?* Québec : Presses de l'Université du Québec.

[Goldhaber, 1997] Goldhaber, Michael H. (1997). "The attention economy and the Net". *First Monday*, 2(4), <http://firstmonday.org/article/view/519/440>.

[Goldstein, 2005] Goldstein, Seth (2005). "Myware".  
<http://majestic.typepad.com/seth/archives.html>

[Gramoli et al., n.d.] Gramoli, V., E. Le Merrer & A. M. Kermarrec (n.d.). « GossiPeer : vers l'uniformisation du développement d'algorithmes épidémiques », Document de travail INRIA.  
<https://gforge.inria.fr/docman/view.php/904/6066/gossipeer.pdf>

[Grint & Woolgar, 1997] Grint, K. & Woolgar, S. (1997). *The Machine at Work: Technology, Work, and Organization*. Oxford, UK : Polity Press.

[Grodzinsky & Tavani, 2005] Grodzinsky, F. S. & Tavani, H. T. (2005). P2P Networks and the Verizon v. RIAA Case: Implications for Personal Privacy and Intellectual Property. *Ethics and Information Technology*, 7, 243-250.

[Grolimund et al., 2006] Grolimund, D., L. Meisser, S. Schmid et R. Wattenhofer (2006). « Cryptree: A Folder Tree Structure for Cryptographic File Systems », *Symposium on Reliable Distributed Systems (SRDS)'06*.

[Guillaud, 2007] Guillaud, Hubert (2007). "Vers le web implicite". *InternetActu*,  
<http://www.internetactu.net/2007/07/06/vers-le-web-implicite/>

[Gunther, 2002] Gunther, N. (2002). Hypernets -- Good (G)news for Gnutella. *Proceedings of CoRR'02*.

[Hagel & Armstrong, 1997] Hagel, J., et Armstrong, A. (1997). *Net gain. Expanding markets through virtual communities*. Boston: Harvard Business School Press.

[Haggerty and Erickson, 2000] Haggerty, K. D., & Ericson, R. V. (2000). The Surveillant Assemblage. *British Journal of Sociology*, 51(4), 605-622.

[Hales, 2006] Hales, D. (2006). Emergent Group-Level Selection in a Peer-to-Peer Network. *Complexus*, 2006 (3): 108-118.

- [Hales, Arteconi, Marcozzi & Chao, 2008] Hales, D., Arteconi, S., Marcozzi, A. & Chao, I. (2008). Towards a Group Selection Design Pattern. In F. Meyer (Eds.), *The European Integrated Project “Dynamically Evolving, Large Scale Information Systems (DELIS)” Proceedings of the final workshop*. Barcelona, February 27-28, 2008.
- [Hamelink, 2000] Hamelink, C. J. (2000). *The Ethics of Cyberspace*. Thousand Oaks/London/New Delhi: Sage Publications.
- [Hancock, 2001] Hancock, D. (2001). An Assessment of ICANN’s Mandatory Uniform Dispute Resolution Policy in Resolving Disputes Over Domain Names, *The Journal of Information, Law and Technology*, 3/2001, [http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2001\\_3/hancock/](http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2001_3/hancock/).
- [Hasan et al., 2005] Hasan, R., Anwar, Z., Yurcik, W., Brumbaugh, L. & Campbell, R. (2005). A Survey of Peer-to-Peer Storage Techniques for Distributed File Systems. Working paper.
- [Hauben & Hauben, 1996] Hauben M. & Hauben R. (1996). *Netizens: On the History and Impact of Usenet and the Internet*, Wiley-IEEE Computer Society Press.
- [Hess, 2005] Hess, D. J. (2005). Technology- and Product-Oriented Movements: Approximating Social Movement Studies and Science and Technology Studies, *Science, Technology & Human Values*, 30 (4) : 515-535.
- [Hine, 2000] Hine, C. (2000). *Virtual Ethnography*. London : Sage.
- [Humble & Farley, 2010] Humble, J. & D. Farley (2010). *Continuous Delivery : Reliable Software Releases through Build, Test, and Deployment Automation*. Boston: Addison-Wesley.
- [Hummel, 2002] Hummel, T. (2002). Instant messaging. In D. Schoder, K. Fischbach, & R. Teichmann (Eds.), *Peer-to-peer – Ökonomische, technologische und juristische Perspektiven*, 59–70. Berlin: Springer.
- [Hummel & Lechner, 2001] Hummel, J., et Lechner, U. (2001). The community model of content management, a case study of the music industry. *Journal of Media Management*, 3(1), 4-14.
- [Hummel & Lechner, 2002] Hummel, J., et Lechner, U. (2002). Business models and system architectures of virtual communities. From a sociological phenomenon to peer-to-peer architectures. *International Journal of Electronic Commerce*, 6(3), 41-53.
- [Irwin, 2008] Irwin, A. (2008). STS perspectives on scientific governance. In Hackett E., Amsterdamska O., Lynch M., and Wajcman J. (eds) *The Handbook of Science and Technology Studies*. Cambridge, MA: The MIT Press, 583–607.
- [James, 2011] James, P. (last modified Mar 2011). Living Without Sessions.

<http://www.peej.co.uk/articles/no-sessions.html>

[Johnson, Crawford, & Palfrey, 2004] Johnson, D. R., Crawford, S. P. & Palfrey, Jr., J. (2004). The Accountable Internet: Peer Production of Internet Governance, *Virginia Journal of Law and Technology*, 9 (9).

[Jouët, 2000] Jouët, J. (2000). Retour critique sur la sociologie des usages. *Réseaux*, 18 (100) : 487-522.

[Jouët, 2011] Jouët, J. (2011). Des usages de la télématique aux *Internet Studies*. In Denouël, J. & Granjon, F. (eds.) *Communiquer à l'ère numérique : Regards croisés sur la sociologie des usages*. Paris : Presses de l'Ecole des Mines, 45-90.

[Kaplan, 2007] Kaplan, Daniel (2007). « Musique et numérique: l'économie MySpace est elle favorable à la creation? ». *InternetActu*, <http://www.internetactu.net/2007/04/12/musique-et-numerique-leconomie-myspace-est-elle-favorable-a-la-creation/>

[Karagiannis et al., 2004] Karagiannis, T., Broido, A., Brownlee, N., Claffy, KC., Faloutsos, M. (2004). Is P2P Dying or Just Hiding ? *Globecom Working Paper*.

[Katyal, 2004] Katyal, S. (2004). The New Surveillance. *Case Western Reserve Law Review*, 54, 297-385.

[Katyal, 2005] Katyal, S. (2005). Privacy Vs. Piracy. *Yale Journal of Law and Technology*, 7, 222-345.

[Katyal, 2009] Katyal, S. (2009). Filtering, Piracy Surveillance, and Disobedience. *Columbia Journal of Law & the Arts*, 32 (4) : 401-426.

[Kessous et al., 2010] Kessous, Emmanuel, Kevin Mellet et Moustafa Zouinar (2010). « L'économie de l'attention : Entre protection des ressources cognitives et extraction de la valeur », *Sociologie du travail*, 52(3) : 359-373.

[Kirschenbaum, 2003] Kirschenbaum, M. (2003). Virtuality and VRML: Software Studies after Manovich. *Electronic Book Review*. Retrieved November 28<sup>th</sup>, 2011, from <http://www.electronicbookreview.com/thread/technocapitalism/morememory>

[Klein, 2002] Klein, B. (2002). When Do Users Detect Information Quality Problems on the World Wide Web ? *Eighth Americas Conference on Information Systems*.

[Kubiatowicz et al., 2005] Kubiatowicz et al. (2005). OceanStore : An Architecture for Global-Scale Persistent Storage, *ACM ASPLOS*.

[Kulkarni, 2009] Kulkarni, S. (2009). Badumna Network Suite: A Decentralized Network Engine for Massively Multiplayer Online Applications. *Working Paper NICTA*.

[Laflaquière, 2005] Laflaquière, J. (2005). Les « autres » applications des technologies peer-to-peer. *Multitudes*, 2 (21) : 59-68.

- [Langley, 2001] Langley, A. (2001). Freenet. In A. Oram (Ed.), *Peer-to-peer: Peer-to-peer: Harnessing the Power of Disruptive Technologies*, 123–132. Sebastopol, CA: O'Reilly.
- [Latour, 2006] Latour, B. (2006). *Changer de société : refaire de la sociologie*. Paris: La Découverte.
- [Latour, 2005] Latour, B. (2005). *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford, UK: Oxford University Press.
- [Latour, 1988] Latour, B. (1988). *The Pasteurization of France*. Cambridge, MA : Harvard University Press.
- [Latour, 1987] Latour, B. (1987). *Science in Action : How to follow scientists and engineers through society*. Cambridge, MA : Harvard University Press.
- [Latzko-Toth, 2010] Latzko-Toth, G. (2010). *La co-construction d'un dispositif socio-technique de communication : le cas de l'Internet Relay Chat*. Thèse de doctorat, Université du Québec à Montréal.
- [Law, 1987] Law, J. (1987). Technology and Heterogeneous Engineering: The Case of Portuguese Expansion. In Bijker, W. E., Hughes, T. P. & Pinch, T. (eds.) *The Social Construction of Technological Systems. New Directions in the Sociology and History of Technology*, Cambridge, MA: The MIT Press, 111-134.
- [Law, 1992] Law, J. (1992). Notes on the Theory of Actor-Network: Ordering, Strategy and Heterogeneity. *Systems Practice*, 5 (4) : 379–393.
- [Le Fessant, 2009] Le Fessant, F. (2009). « Les réseaux sociaux au secours des réseaux pair-à-pair », *Défense nationale et sécurité collective*, 3 : 29-35.
- [Lessig, 2002] Lessig, L. (2002). *The Future of Ideas*. New York: Vintage Books.
- [Lessig, 2000] Lessig, L. (2000). The Law in Code, and the Code in Law. Communication présentée au pcForum 2000, Phoenix, AZ.  
[www.lessig.org/content/articles/works/pcforum.pdf](http://www.lessig.org/content/articles/works/pcforum.pdf)
- [Lessig, 1999] Lessig, L. (1999). *Code and Other Laws of Cyberspace*. New York : Basic Books.
- [Lethin, 2001] Lethin, R. (2001). Reputation. In A. Oram (Ed.), *Peer-to-peer: Harnessing the Power of Disruptive Technologies*, 341–353. Sebastopol, CA: O'Reilly.
- [Leuf, 2002] Leuf, B. (2002). *Peer-to-peer. Collaboration and sharing over the Internet*. Boston: Addison-Wesley.
- [Li, 2007] Li, J. (2007). *A Survey of Peer-to-Peer Network Security Issues*.  
<http://www.cse.wustl.edu/~jain/>
- [Li et al., 2011] Li, Z., J. Wu, J. Xie, T. Zhang, G. Chen & Y. Dai (2011). Stability-

Optimal Grouping Strategy of Peer-to-Peer Systems, *IEEE Trans. Parallel Distributed Systems*, 2079-2087.

[Licklider & Taylor, 1968] Licklider, J.C.R., et Taylor, W. (1968). The computer as a communication device. *Science and Technology*, 21-40. Ré-imprimé en <http://memex.org/licklider.pdf>

[Lindsay, 2003] Lindsay, C. (2003). From the Shadows: Users as Designers, Producers, Marketers, Distributers and Technical Support. In Oudshoorn, N. & Pinch, T. (eds.) *How Users Matter. The Co-Construction of Users and Technologies*, Cambridge, MA : The MIT Press, 29-50.

[Lowe, 2009] Lowe, Janet (2009). *Google Speaks: Secrets of the World's Greatest Billionaire Entrepreneurs, Sergey Brin and Larry Page*. Hoboken, NJ: John Wiley & Sons.

[Lyon, 1988] Lyon, D. (1988). *The Information Society: Issue and Illusion*. Oxford: Polity Press.

[Malcolm, 2008] Malcolm, J. (2008). *Multi-Stakeholder Governance and the Internet Governance Forum*. Wembley, WA : Terminus Press.

[Mallard, 2011] Mallard, A. (2011). « Explorer les usages: un enjeu renouvelé pour l'innovation des TIC », en J. Denouel & F. Granjon (dir.) *Communiquer à l'ère numérique. Regards croisés sur la sociologie des usages*, Paris : Presses de l'Ecole des Mines.

[Manovich, 2001] Manovich, L. (2001). *The Language of New Media*. Cambridge, MA: The MIT Press.

[Marcozzi & Hales, 2008] Marcozzi, A. & Hales, D. (2008). Emergent Social Rationality in a Peer-to-Peer System. *Advances in Complex Systems (ACS)*, 11 (4), 581-595.

[Marshall, 2002] Marshall, L. (2002). Metallica and Morality : The Rhetorical Battleground of the Napster Wars. *Entertainment Law*, 1(1) : 1-19.

[Marino, 2006] Marino, M. C. (2006). Critical Code Studies. *Electronic Book Review*. Retrieved November 28<sup>th</sup>, 2011, from <http://www.electronicbookreview.com/thread/electropoetics/codology>

[Marsden, 2010] Marsden, C. (2010). *Net Neutrality: Towards a Co-Regulatory Solution*. Bloomsbury USA.

[Martin, van Brakel & Bernhard, 2009] Martin, A. K., Van Brakel, R. & Bernhard, D. (2009). Understanding Resistance to Digital Surveillance: Towards a Multi-Disciplinary, Multi-Actor Framework. *Surveillance & Society*, 6(3), 213-232.

[Massit-Folléa, 2008] Massit-Folléa, F. (2008). Gouverner l'Internet comme un bien

commun mondial ?, EuroDIG. En ligne sur  
<http://www.voxinternet.org/spip.php?article251>.

[Massit-Folléa, Méadel & Monnoyer-Smith, 2012] Massit-Folléa, F., C. Méadel & L. Monnoyer-Smith, (2012, eds.) *Normative Experience in Internet Politics*, Paris, Presses des Mines.

[Méadel & Marzouki, 2005] Méadel, C. & M. Marzouki (2005). Quelle justice pour Internet ? L'arbitrage sur les noms de domaine. In *Internet, nouvel espace public mondialisé ?*, Actes du colloque franco-québécois du 28 novembre 2003, Paris.

[Méadel, 2011] Méadel, C. (2011). « Intermédiation, médiation, désintermédiation... », *Au fil des Labs #1*, Labs Hadopi.

[Méadel, 1994] Méadel, C. (1994). *Histoire de la radio des années trente*. Paris : Anthropos/INA.

[Millerand, 1998] Millerand, F. (1998). Usages des NTIC : les approches de la diffusion, de l'innovation et de l'appropriation (1<sup>ère</sup> partie). *COMPOSITE*, v98.1, <http://www.composite.org/index.php/revue/article/viewPDFInterstitial/102/80>

[Milojicic et al., 2002] Milojicic, D.S., Kalogeraki, V., Lukose, R., Nagaraja, K., Pruyne, J., Richard, B., et al. (2002). Peer-to-peer computing. *Hewlett Packard Tech Report*, 2002.

[Minar, 2001] Minar, N. (2001). Distributed systems topologies: Part 1 ([http://www.openp2p.com/pub/a/p2p/2001/12/14/topologies\\_one.html](http://www.openp2p.com/pub/a/p2p/2001/12/14/topologies_one.html))

[Minar, 2002] Minar, N. (2002). Distributed systems topologies: Part 2 ([http://www.openp2p.com/pub/a/p2p/2002/01/08/p2p\\_topologies\\_pt2.html](http://www.openp2p.com/pub/a/p2p/2002/01/08/p2p_topologies_pt2.html))

[Minar & Hedlund, 2001] Minar, N. et Hedlund, M. (2001). A network of peers – Peer-to-peer models through the history of the Internet. In A. Oram (Ed.), *Peer-to-peer: Harnessing the Power of Disruptive Technologies*, 9-20. Sebastopol, CA: O'Reilly.

[Missika, 2006] Missika, J.-L. (2006). *La fin de la télévision*. Paris : Le Seuil/La république des idées.

[Moglen, 2010] Moglen, E. (2010). Freedom In The Cloud : Software Freedom, Privacy and Security for Web 2.0 and Cloud Computing. ISOC Meeting, New York Branch, 5 February 2010.

[Monberg, 2005] Monberg, J. (2005). Science and Technology Studies Approaches to Internet Research. *The Information Society*, 21 (4) : 281-284.

[Monnoyer-Smith, 2008] Monnoyer-Smith, L. (2008). Postface. La notion d'imaginaire : mauvaise réponse à une bonne question ? In Lakel, A., Massit-Folléa, F. & Robert, P. (eds.) *Imaginaire(s) des technologies d'information et de*

*communication*, Paris : Editions de la Maison des Sciences de l'Homme/PraTICs, 103-108.

[Morbidelli, Pegoraro, Reposo, & Volpi, 2004] Morbidelli, G., Pegoraro, L., Reposo, A. & Volpi, M. (2004). *Diritto Pubblico Comparato*, Torino: G. Giappichelli Editore.

[Mounier, 2002] Mounier, P. (2002). *Les maîtres du réseau. Les enjeux politiques d'Internet*. Paris : La Découverte.

[Mowbray, 2009] Mowbray, M. (2009). The Fog over the Grimpen Mire : Cloud Computing and the Law. *SCRIPTed*, 6(1) : 132-146

[Mueller, 2012] Mueller, M. (2012). Property and Commons in Internet Governance. In Brousseau, E., M. Marzouki & C. Méadel (eds., 2012). *Governance, Regulation and Powers on the Internet*. Cambridge : Cambridge University Press.

[Mueller, 2002a] Mueller, M. (2002). *Ruling the Root : Internet Governance and the Taming of Cyberspace*. Cambridge, MA : The MIT Press.

[Mueller, 2002b] Mueller, M. (2002). Success by Default: A New Profile of Domain Name Trademark Disputes Under ICANN's UDRP, Report, <http://dcc.syr.edu/markle/markle-report-final.pdf>

[Mueller, 2001] Mueller, M. (2001). Rough Justice: A Statistical Assessment of ICANN's Uniform Dispute Resolution Policy, *The Information Society*, 17(3) : 153-163.

[Muniesa, Millo & Callon, 2007] Muniesa, F., Millo, Y. & Callon, M. (2007). An introduction to market devices. *Sociological Review*, 55 (2) : 1-12.

[Musiani, à paraître 2012] Musiani, F. (à paraître 2012). Caring About the Plumbing : On the Importance of Architectures in Social Studies of (Peer-to-Peer) Technology. *Journal of Peer Production*, 1.

[Musiani, 2011] Musiani, F. (2011). Privacy as Invisibility : Pervasive Surveillance and the Privatization of Peer-to-Peer Systems. *tripleC*, 9(2): 126-140.

[Musiani, 2010a] Musiani, F. (2010). « Ménager le droit à la vie privée, entre anonymat et connaissance de l'identité: les débuts des réseaux sociaux en pair-à-pair », *Terminal*, 105: 107-116.

[Musiani, 2010b] Musiani, F. (2010). « When Social Links Are Network Links: the Dawn of Peer-to-Peer Social Networks and Its Implications for Privacy », *Observatorio*, 4(3), 185-207.

[Musiani, 2009] Musiani, F. (2009). *Cyberhandshakes : How the Internet Challenges Dispute Resolution (...And Simplifies It)*. Madrid : EuroEditions.

[Musiani, 2007] Musiani, F. (2007). *New Voices and Power Balances : The Potential of the Internet Governance Forum for Collaborative Decision-Making on the Future of the Internet*. Thèse de Master, Université de Padoue, Italie.

- [Musiani & Schafer, 2011] Musiani, F. & Schafer, V. (2011). Le modèle Internet en question (années 1970-2010). *Flux*, 85-86 (3-4): 62-71.
- [Neff & Stark, 2003] Neff, G. & Stark, D. (2003). Permanently Beta: Responsive Organization in the Internet Era. In Howard, P. & Jones, S. G. (eds.) *Society Online: The Internet in Context*, Thousand Oaks, CA : Sage Publications, 173-188.
- [Neumann & Star, 1996] Neumann, L. & Star, S. L. (1996). Making Infrastructure: the Dream of a Common Language. In J. Blomberg, F. Kensing, & E. Dykstra-Erickson (Eds.), *Proceedings of the PDC '96* (pp. 231-240). Palo Alto, CA: Computer Professionals for Social Responsibility.
- [Noll & Meinel, 2007] Noll, M. G. & C. Meinel (2007). Web Search Personalization Via Social Bookmarking and Tagging. *Lecture Notes in Computer Science*, 4825/2007: 367-380
- [Oram, 2004] Oram, A. (2004). From P2P to Web Services: Addressing and Coordination, *O'Reilly XML.com*, <http://www.xml.com/pub/a/2004/04/07/p2p-ws.html>
- [Oram, 2001] Oram, A. (Ed.) (2001). *Peer-to-peer: Harnessing the Power of Disruptive Technologies*. Sebastopol, CA: O'Reilly.
- [O'Reilly, 2000] O'Reilly, T. (2000). Remaking the Peer-to-Peer Meme. In A. Oram (Ed.), *Peer-to-peer: Peer-to-peer: Harnessing the Power of Disruptive Technologies*. Sebastopol, CA: O'Reilly.
- [Ostrom, 1990] Ostrom, E. (1990). *Governing the Commons : The Evolution of Institutions for Collective Action*, Cambridge : Cambridge University Press.
- [Oudshoorn & Pinch, 2003] Oudshoorn, N. & Pinch, T. (eds., 2003). *How Users Matter. The Co-Construction of Users and Technologies*, Cambridge, MA : The MIT Press.
- [Paloque-Bergès, 2012] Paloque-Bergès, C. (2012). La mémoire culturelle d'Internet : le folklore de Usenet. *Le Temps des Médias*, 18.
- [Parziale et al., 2006] Parziale, L., D. Britt, C. Davis, J. Forrester, W. Liu, C. Matthews & N. Rosselot (2006). *TCP/IP Tutorial and Technical Overview*. Eighth Edition. IBM Redbooks.
- [Patterson, 2004] Patterson, A. (2004). Why Writing Your Search Engine Is Hard, *Queue*, [www.acmqueue.com](http://www.acmqueue.com)
- [Pfaffenberger, 1996] Pfaffenberger, B. (1996). If I Want It, It's OK: Usenet and the (Outer) Limits of Free Speech. *The Information Society*, 4(12) : 365-373. <http://pfaff.sts.virginia.edu/bphome/docs/pdf/usenet.pdf>
- [Pollock, 2008] Pollock, R. (2008). Is Google the Next Microsoft? Competition, Welfare and Regulation in Internet Search. Working Paper, University of Cambridge,

UK. [http://rufuspollock.org/economics/papers/search\\_engines.pdf](http://rufuspollock.org/economics/papers/search_engines.pdf)

[Pouwelse et al., 2006] Pouwelse, J. A., P. Garbacki, J. Wang, A. Bakker, J. Yang, A. Iosup, D.H.J. Epema, M. Reinders, M. van Steen, H.J. Sips (2006). Tribler: a Social-Based Peer-to-Peer System. *Concurrency and Computation: Practice & Experience*, 20 (2), 127-138.

[Preibusch et al., 2007] Preibusch, S., B. Hoser, S. Guerses, & B. Berendt (2007). Ubiquitous social networks – opportunities and challenges for privacy-aware user modeling. *Proceedings of the Data Mining for User Modelling Workshop*, Corfu, June 2007. <http://ideas.repec.org/p/diw/diwwpp/dp698.html>.

[Proulx, 2009] Proulx, S. (2009). L'intelligence du grand nombre : la puissance d'agir des contributeurs sur Internet – limites et possibilités. 7<sup>ème</sup> colloque du chapitre français de l'ISKO, *Intelligence collective et organisation des connaissances*, Lyon, 24-26 juin 2009. <http://pro.ovh.net/~iskofran/pdf/isko2009/PROULX.pdf>

[Raymond, 1997] Raymond, E. (1997). *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*. O'Reilly Media.

[Rejeski, 2003] Rejeski, D. (2003). Making Policy in a Moore's Law World. *Ubiquity*, December 2003. <http://ubiquity.acm.org/article.cfm?id=962076>.

[Ribes & Lee, 2010] Ribes, D. & Lee, C. P. (2010). Sociotechnical Studies of Cyberinfrastructure and e-Research: Current Themes and Future Trajectories. *Computer Supported Cooperative Work*, 19, 231-244.

[Rieder, 2008] Rieder, Bernhard (2008). "Democratizing Search". *World-Information Institute Conference "Deep Search"*, [http://world-information.org/wii/deep\\_search/en/program](http://world-information.org/wii/deep_search/en/program)

[Rieder & Sire, 2012] Rieder, B. & G. Sire (à paraître 2012). Strange Commodities, Skewed Incentives: A Microeconomic Critique of How Google Conceives the Web. *New Media & Society*.

[Ripeanu, 2001] Ripeanu, M. (2001). Peer-to-peer Architecture Case Study : Gnutella Network. University of Chicago Technical Report TR-2001-26.

[Rivière, 2008] Rivière, E. (2008). « Protocoles épidémiques principes et utilisations pour les systèmes répartis à grande échelle », *Séminaire GREYC*, Université de Caen, <https://www.greyc.fr/node/264>

[Rodotà, 1997] Rodotà, S. (1997). *Tecnopolitica. La democrazia e le nuove tecnologie della comunicazione*, Roma-Bari, Laterza.

[Rodotà, 2006] Rodotà, S. (2006). *La vita e le regole. Tra diritto e non diritto*, Milano, Feltrinelli.

[Rozo-Acuna, 2002] Rozo-Acuna, E. (2002). *Habeas Data costituzionale: nuova*

garanzia giurisdizionale del diritto pubblico latinoamericano, *Diritto pubblico comparato ed europeo*, 4 : 1829-1872.

[Salaün, 2012] Salaün, J.-M. (2012). *Vu, lu, su. Les architectes de l'information face à l'oligopole du Web*. Paris: La Découverte.

[Schaar, 2010] Schaar, P. (2010). Privacy by Design. *Identity in the Information Society*, 3(2) : 267-274.

[Schafer, Le Crosnier & Musiani, 2011] Schafer, V., H. Le Crosnier & F. Musiani (2011). *La neutralité de l'Internet, un enjeu de communication*. Paris: CNRS Editions/Les Essentiels d'Hermès.

[Schafer & Thierry, 2012] Schafer, V. & B. G. Thierry (2012). *Le Minitel. L'enfance numérique de la France*. Paris : Nuvis.

[Shaker, 2006] Shaker, L. (2006) In Google We Trust: Information Integrity in the Digital Age. *First Monday*, 11 (4).  
<http://www.firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1320/1240>

[Schoder & Fischbach, 2003] Schoder, D. et Fischbach, K. (2003). Peer-to-peer prospects. *Communications of the ACM*, 46 (2), 27–29.

[Schoder, Fischbach & Schmitt, 2005] Schoder, D., Fischbach, K. et Schmitt, C. (2005). Core concepts in peer-to-peer networking. In R. Subramanian et B. D. Goodman (Eds.) *Peer-to-peer computing : The Evolution of a Disruptive Technology*, 1-27. Hershey : Idea Group Publishing.

[Schoder, Fischbach & Teichmann, 2002] Schoder, D., Fischbach, K. et Teichmann, R. (Eds., 2002). *Peer-to-peer - Ökonomische, technologische und juristische Perspektiven*. Berlin: Springer.

[Schollmeier, 2001] Schollmeier, R. (2002). A definition of peer-to-peer networking for the classification of peer-to-peer architectures and applications. *Proceedings of the First International Conference on Peer-to-Peer Computing*, 27–29.

[Shields, 1996] Shields, R. (eds., 1996). *Cultures of Internet. Virtual Spaces, Real Histories, Living Bodies*. London : Sage.

[Shirky, 2001] Shirky, C. (2001). Listening to Napster. In A. Oram (Ed.), *Peer-to-peer: Harnessing the Power of Disruptive Technologies*, 21-37. Sebastopol, CA: O'Reilly.

[Shirky, 2000] Shirky, C. (2000). Modern P2P Definition, *OpenP2P.com*,  
<http://www.openp2p.com/pub/a/p2p//11/24/shirky1-whatisp2p.html>

[Shirky et al., 2001] Shirky, C., Truelove, K., Dornfest, R., Gonze, L., & Dougherty, D. (Eds., 2001). *2001 P2P networking overview*. Sebastopol, CA: O'Reilly.

[Simon, 1971] Simon, Herbert A. (1971). "Designing Organizations for an

Information-Rich World”, in Martin Greenberger, *Computers, Communication, and the Public Interest*, Baltimore, MD: The Johns Hopkins Press.

[Simonnot, 2008] Simonnot, B. (eds., 2008). Moteurs de recherche. Usages et enjeux. Dossier, *Questions de communication*, 18.

[Smith, Clippinger & Konsynski, 2003] Smith, H., Clippinger, J. et Konsynski, B. (2003). Riding the wave: Discovering the value of P2P technologies. *Communications of the Association for Information Systems*, 11, 94-107.

[Smyth et al., 2009] Smyth, B., P. Briggs, M. Coyle & M. P. O’Mahony (2009). Google Shared : A Case-Based Perspective on Social Web Search. *Lecture Notes in Computer Science*, 5650/2009 : 494-508.

[Solove, 2004] Solove, D. J. (2004). *The Digital Person: Technology and Privacy in the Information Age*. New York, NY: New York University Press.

[Solove, 2006] Solove, D. J. (2006). A Taxonomy of Privacy. *University of Pennsylvania Law Review*, 154 (3) : 477-560.

[Star, 1999] Star, S. L. (1999). The Ethnography of Infrastructure. *American Behavioral Scientist*, 43 (3), 377-391.

[Star, 2002] Star, S. L. (2002). Infrastructure and ethnographic practice: Working on the Fringes. *Scandinavian Journal of Information Systems*, 14 (2), 107-122.

[Star & Bowker, 2006] Star, S. L. & Bowker, G. (2006). How To Infrastructure. In Lievrouw, L. A. (Ed.), *Handbook of New Media* (pp. 151-162), London: Sage.

[Star & Griesemer, 1989] Star, S. L. & Griesemer, J. (1989). Institutional Ecology, ‘Translations’ and Boundary Objects: Amateurs and Professionals in Berkeley’s Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19 (3) : 387-420.

[Star & Ruhleder, 1996] Star, S. L., and Ruhleder, K. (1996). Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces. *Information Systems Research*, 7, 111-133.

[Strauss, 1978] Strauss, A. (1978). A Social Worlds Perspective. *Studies in Symbolic Interaction*, 1 : 119-128.

[Stump, 2002] Stump, M. (2002). Peer-to-peer tracking can save cash: Ellacoya. ([http://www.ellacoya.com/news/pdf/10\\_07\\_02\\_mcn.pdf](http://www.ellacoya.com/news/pdf/10_07_02_mcn.pdf)).

[Subramanian & Goodman, 2005] Subramanian, R. & B. D. Goodman (Eds., 2005). *Peer-to-Peer Computing: The Evolution of a Disruptive Technology*. Hershey : Idea Group Publishing.

[Suvanto, 2005] Suvanto, M. (2005). Privacy in Peer-to-Peer Networks. *Helsinki*

University of Technology T-110.551 Seminar on Internetworking. Retrieved November 29, 2010 from [www.tml.tkk.fi/Publications/C/18/suvanto.pdf](http://www.tml.tkk.fi/Publications/C/18/suvanto.pdf).

[Taylor & Harrison, 2009] Taylor, I. & Harrison, A. (2009). *From P2P to Web Services and Grids: Evolving Distributed Communities. Second and Expanded Edition*. London: Springer-Verlag.

[Tavani, 2008] Tavani, H. T. (2008). Informational Privacy. Concepts, Theories and Controversies. In K. E. Himma & H. T. Tavani (Eds.), *The Handbook of Information and Computer Ethics* (pp. 131-164). Hoboken, NJ: John Wiley & Sons.

[Trompette et Vinck, 2009] Trompette, P. & Vinck, D. (Eds., 2009) Retour sur la notion d'objet-frontière. *Revue d'anthropologie des connaissances*, 3(1).

[Vanbreemersch, 2009] Vanbreemersch, N. (2009). *De la démocratie numérique*. Paris : Le Seuil.

[van Eeten, 2009] van Eeten, M. (2009). Where is the Governance in Internet Governance? *GigaNet Annual Symposium*, Sharm-el-Sheikh, Egypt, November 14<sup>th</sup>, 2009.

[van Schewick, 2010] van Schewick, B. (2010). *Internet Architecture and Innovation*. Cambridge, MA: The MIT Press.

[Vinck, 2003] Vinck, D. (Ed., 2003). *Everyday Engineering. An Ethnography of Design and Innovation*. Cambridge, MA : The MIT Press.

[von Hippel, 1986] von Hippel, E. (1986). Lead Users: A Source of Novel Product Concepts. *Management Science*, 32 (7) : 791–806.

[von Hippel, 2001] von Hippel, E. (2001). Innovation by User Communities: Learning From Open-Source Software. *Sloan Management Review*, 42 (4) : 82-86.

[von Hippel, 2005] von Hippel, E. (2005). *Democratizing Innovation*. Cambridge, MA : The MIT Press.

[Yang & Garcia-Molina, 2001] Yang, B., & Garcia-Molina, H. (2001). Comparing hybrid peer-to-peer systems. *Working Paper*. ([http://www-db.stanford.edu/~byang/pubs/hybridp2p\\_med.pdf](http://www-db.stanford.edu/~byang/pubs/hybridp2p_med.pdf))

[Wei, Xie & Zhou, 2008] Wei, G., Xie, M. et Zhou, G. (2008). HiPEC : A Hybrid P2P Model for Electronic Commerce. *IEEE Symposium on Advanced management of Information for Globalized Enterprises*, 2008. ([http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=4721502&tag=1](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4721502&tag=1)).

[Wellman, 2004] Wellman, B. (2004). The Three Ages of Internet Studies : Ten, Five and Zero Years Ago. *New Media & Society*, 6 (1) : 123-129.

[Williams *et al.*, 2007] Williams, C., Huibonhoa, P., Holliday, JA., Hospodor, A.,

Schwarz, T. (2007). Redundancy Management for P2P Storage. *Working Paper*.

[Wilson, 2006] Wilson, F. (2006). Spying on Myself. *AVC*.

[http://avc.blogs.com/a\\_vc/2006/01/spying\\_on\\_myself.html](http://avc.blogs.com/a_vc/2006/01/spying_on_myself.html)

[Winer, 2000] Winer, D. (2000). The P in P2P. *Scripting*, 9 septembre 2000.

<http://scripting.com/davenet/2000/09/09/thePInP2p.html>

[Wood, 2010] Wood, J. A. (2010). The Darknet: A Digital Copyright Revolution.

*Richmond Journal of Law & Technology*, 16(4).

<http://jolt.richmond.edu/v16i4/article14.pdf>.

[Woolgar, 1991] Woolgar, S. (1991). Configuring the User: the Case of Usability Trials. In Law, J. (ed.) *A Sociology of Monsters*, Sociological Review, Monograph 38. London: Routledge, 57-102.

[Zimmer, 2008] Zimmer, M. (2008). The Gaze of the Perfect Search Engine: Google as an Infrastructure of Dataveillance. In A. Spink and M. Zimmer (eds.), *Web Search, Springer Series in Information Science and Knowledge Management*, 14. Berlin: Springer-Verlag, 77-99.



## **TABLEAUX**

### **Chapitre 2**

Tableau 1. Modèle à trois niveaux précisant la terminologie relative au P2P. D'après [Schoder, Fischbach & Schmitt, 2005] et [Oram, 2001].

### **Chapitre 4**

Tableau 2. Chronologie du développement de Faroo, 2006-2012.

Tableau 3. Post « Scaling & Market Entry Barrier » sur le blog de Faroo, <http://blog.faroo.com/>, mai 2008.

Tableau 4. Clause 5 des termes de licence Faroo (à partir de février 2009).

### **Chapitre 5**

Tableau 5. Chronologie du développement de Wuala, 2004-2011.

Tableau 6. Extrait des conditions d'utilisation de Wuala, consultées à <http://www.wuala.com/en/about/terms>, version de mai 2009.

Tableau 7. Le type de connexion Internet comme barrière à l'entrée du mécanisme de troc [Section « Trading Storage » du site Wuala].

Tableau 8. Extrait de la conférence « Google Tech Talk » donnée au siège Google de Mountain View, Californie, par Dietrich Poell en octobre 2007.

### **Chapitre 6**

Tableau 9. Chronologie du développement de Tribler et de P2P-Next, 2005-2012.

Tableau 10. Extrait de la page du site Tribler expliquant la philosophie qui sous-tend le dispositif (<http://tribler.org/trac/wiki/AboutTribler>, page en l'état vue pour la dernière fois en mars 2011)

Tableau 11. Extrait de la foire aux questions (FAQ) du site Tribler (<http://tribler.org/trac/wiki/faq#LegalP2Pdoesthatexist>), consultée en décembre 2009.

## IMAGES

### Chapitre 4

Image 1. Capture d'écran du « compteur » des pairs contactés par le client P2P d'un utilisateur au cours d'une interrogation de Faroo.

Image 2. Capture d'écran montrant à l'utilisateur la construction de l'index distribué, et son stockage en local par chaque nœud/pair. Chaque utilisateur peut suivre la manière dont l'index se construit et quel rôle a sa machine.

Image 3. Capture d'écran du panneau de configuration, qui permet à l'utilisateur de voir et surveiller ses contributions de ressources au moteur de recherche distribué.

### Chapitre 5

Image 4. Capture d'écran de l'interface Wuala vue par l'utilisateur, stade de closed alpha (début 2008) sur système d'exploitation Linux. Image par Daniel Bartholomew sur LinuxJournal.com.

Image 5. Capture d'écran de la partie « réseau social » de l'interface Wuala vue par l'utilisateur, stade de closed alpha sur système d'exploitation Linux. Image par Daniel Bartholomew sur LinuxJournal.com

Image 6. Capture d'écran du programme de réglage des paramètres pour le stockage partagé dans Wuala, ainsi qu'il est vu par l'utilisateur. Image par Damien Oh sur maketecheasier.com.

Image 7. Capture d'écran du réglage de la mise en commun de ressources pour le stockage partagé dans Wuala, ainsi qu'il est vu par l'utilisateur. Image par Damien Oh sur maketecheasier.com.

Image 8. Extrait de la conférence « Google Tech Talk » donnée au siège Google de Mountain View, Californie, par Dietrich Poell en octobre 2007.

Image 9. Représentation graphique schématique du fonctionnement de Wuala, site Web de l'entreprise (<http://www.wuala.com/fr/learn/technology>), 2010.

### Chapitre 6

Image 10. De ABC à Tribler 3.x (première version)

Image 11. Interface usager de Tribler, série 4.x.

Image 12. Interface de Tribler 5.0 du point de vue de l'utilisateur (2009), montrant les résultats d'une recherche.

## **Nains sans géants. Architecture décentralisée et services Internet**

**RÉSUMÉ :** Si le concept de décentralisation est en quelque sorte inscrit au cœur même de l'Internet, son urbanisme actuel n'intègre ce principe que de manière limitée. La recherche d'alternatives au mode dominant d'organisation des services basés sur Internet – une concentration autour de grands ensembles de serveurs sous le contrôle des "géants" du secteur – se poursuit, en quête d'efficacité et de durabilité. En cherchant les meilleures solutions, certains développeurs se retournent vers les qualités persistantes d'une ancienne technologie, le pair-à-pair (P2P), qui replonge dans la topologie de l'Internet pré-commercial, mettant à profit les ressources des « nains » du réseau – ses marges ou sa périphérie.

Cette thèse explore l'approche distribuée et décentralisée de l'architecture technique des services Internet. Il s'agit de comprendre ce que dessine une architecture de réseau décentralisée du point de vue de l'articulation des acteurs et des contenus, de la répartition de responsabilités, de l'organisation du marché et de la capacité à exercer du contrôle, des formes d'existence et des rôles d'entités telles que les nœuds du réseau, ses usagers, ses unités centrales. Ce travail analyse sous quelles conditions un réseau qui répartit à ses marges la responsabilité de son propre fonctionnement, et qui suit un modèle non hiérarchisé ou hybride, peut se développer dans l'Internet d'aujourd'hui. La thèse suit les développeurs de trois services Internet – un moteur de recherche, un service de stockage et une application pour le streaming vidéo – construits sur un modèle de réseau décentralisé, les collectifs d'usagers pionniers qui se développent avec ces services, et ponctuellement, les arènes politiques où l'on discute de l'organisation et de la gouvernance de l'Internet à moyen et long terme.

**Mots clés :** Internet ; services ; architecture ; décentralisation ; distribution ; P2P ; pair-à-pair ; recherche ; stockage ; streaming ; usagers ; développement ; gouvernance

## **Dwarfs without giants. Decentralized architecture and Internet-based services**

**ABSTRACT:** Even if the concept of decentralization is embedded to some extent at the very core of the Internet, today's "network of networks" integrates this principle only partially. The dominant organizational model for Internet-based services involves large clusters of servers controlled by the "giants" of the IT sector. The search for alternatives is in progress, aiming at different ways to achieve effectiveness and sustainability. In this quest, a number of developers look back to the evergreen qualities of a relatively old technology, peer-to-peer (P2P), that leverages the socio-technical resources of the network's "dwarfs" - its periphery or "edge" - in a way that is, in fact, closer to the pre-commercial Internet.

This dissertation explores the distributed and decentralized approach to the technical architecture of Internet-based services. It illustrates the co-shaping of a decentralized network architecture and of several different dynamics: the articulation between actors and contents, the allocation of responsibilities and the capacity to exert control, the organization of the market, the forms of existence and role of entities such as the nodes of a network, its users, its central or coordinating units. This work analyses the conditions under which a network that structures itself according to a non-hierarchical or hybrid model, and delegates the responsibility of its functioning to its edge, can develop and thrive in today's Internet. The dissertation follows the developers of three Internet services - a search engine, a storage service and a video streaming application - built on primarily decentralized network models; it also follows the collectives of pioneer users developing with these services, and selectively, the political venues where the Internet's medium- and long-term organization and governance are discussed.

**Keywords:** Internet; services; architecture; decentralization; distribution; P2P; peer-to-peer; network edge; search; storage; streaming; users; development; governance

