Privacy as invisibility: pervasive surveillance and the privatization of peer-to-peer systems
Francesca Musiani

To cite this version:
Francesca Musiani. Privacy as invisibility: pervasive surveillance and the privatization of peer-to-peer systems. tripleC, 2011, 9 (2), pp.126-140. <halshs-00563789v2>

HAL Id: halshs-00563789
https://halshs.archives-ouvertes.fr/halshs-00563789v2
Submitted on 29 Jun 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.
Privacy as Invisibility: Pervasive Surveillance and the Privatization of Peer-to-Peer Systems

Francesca Musiani

Centre de Sociologie de l’Innovation, MINES ParisTech/CNRS;
francesca.musiani@mines-paristech.fr

Abstract: This article addresses the ongoing, increasing privatization of peer-to-peer (P2P) file sharing systems – the emergence of systems that users may only join by personal, friend-to-friend invitation. It argues that, within P2P systems, privacy is increasingly coinciding with “mere” invisibility vis-à-vis the rest of the Internet ecosystem because of a trend that has shaped the recent history of P2P technology: The alternation between forms of pervasive surveillance of such systems, and reactions by developers and users to such restrictive measures. Yet, it also suggests that the richness of today’s landscape of P2P technology development and use, mainly in the field of Internet-based services, opens up new dimensions to the conceptualization of privacy, and may give room to a more articulate definition of the concept as related to P2P technology; one that includes not only the need of protection from external attacks, and the temporary outcomes of the competition between surveillance and counter-surveillance measures, but also issues such as user empowerment through better control over personal information, reconfiguration of data management practices, and removal of intermediaries in sharing and communication activities.

Keywords: peer-to-peer, privacy, surveillance, invisibility, privatization, sharing, communication

Acknowledgement: This article owes a lot to the discussions that took place within the frame of the Surveillance and Society Track of the European Association for the Study of Science and Technology (EASST) Biannual Conference 2010, which took place September 2-4, 2010, in Trento, Italy. My thanks go to the convenors and participants of the track for the insights and comments on the paper I presented in Trento, closely related to the present article; I am also grateful to Marisol Sandoval, Peter Lauritsen and another anonymous reviewer for helpful suggestions and comments on earlier drafts.

In the last ten years, peer-to-peer (P2P) has become one of the most “hyped” and discussed words in the field of Information and Communication Technologies (Shirky et al., 2001; Schoder & Fischbach, 2003). The term refers to the notion that in a network of equals (or “peers”), by means of appropriate systems of exchange and communication, two or more nodes in the network are able to spontaneously collaborate, with no need of a central coordination or intermediation (Schollmeier, 2002; Schoder & Fischbach, 2003). Very frequently framed as a threat to the digital content industry – their most diffused use by the public at large being the unauthorized sharing of copyright-protected materials – P2P systems are well-suited on one hand to give free, and immediate, access to perfect copies, and on the other hand, to promote increased effectiveness, freedom, and stability in online content distribution, enhanced by the direct connections between the nodes-users of the system (Elkin-Koren, 2006).

This article addresses P2P’s ongoing, increasing privatization – the emergence of P2P file sharing systems that users may only join by personal, friend-to-friend invitation (Rogers & Bhatti, 2007). It argues that, within P2P systems, privacy is increasingly coinciding with “mere” invisibility vis-à-vis the rest of the Internet ecosystem, a tendency rooted in a trend that has shaped the recent history of P2P technology. This trend consists in the alternation between manifestations of what will be called “pervasive surveillance” of such systems – the ubiquitous and generalised search of potential infringers carried out by copyright owners on the Internet – and reactions by developers, and eventually users, to such restrictive measures. Yet, it also suggests that the richness of today’s landscape of P2P technology development and use, mainly in the field of Internet-based services, opens up new dimensions to the conceptualization of privacy, and may give room to a more articulate definition of it; one that includes not only the need of protection from external attacks, and the
temporary outcomes of the competition between surveillance and counter-surveillance measures, but also issues such as user empowerment through a better control over personal information, reconfiguration of data management practices, and removal of intermediaries in sharing and communication activities.

The first part of the article introduces the conceptualizations of privacy and surveillance that are relevant to the argument of this article. After touching upon David Lyon’s depiction of the “World Wide Web of surveillance” (Lyon, 1997), it outlines Sonia Katyal’s concept of “piracy surveillance” (2005) as pervasive detection of consumer infringement, and Frances Grodzinsky and Herman Tavani’s argument (2005) that placing the burden of infringers’ identification on copyright owners has opened up a new culture of surveillance, one that entitles copyright owners to pervasively search the Internet for potential infringers.

P2P technology’s history, as Niva Elkin-Koren remarks (2006), has been deeply informed by the frequent, almost overwhelming, association of such technology with one of its possible uses, (illegal) file-sharing. Thus, the second part of the article outlines the different generations of P2P file-sharing systems’ genealogy, starting from the moment in which the public at large first accessed them (Napster, 1999). The genealogy of P2P file-sharing systems is, in fact, a story of tensions between surveillance and counter-surveillance technologies. It is argued that the ways in which P2P systems have taken shape and evolved in the last decade are closely linked to the dialectic between juridico-technical measures restricting P2P-enabled file sharing activities, and socio-technical responses that have shortly followed each of them: in other words, to the constant attempts of surveillance technologies and sharing technologies to outrun each other. In this sense, the genealogy of P2P file-sharing systems is also a history of resistance towards regulation of user behaviour by means of digital surveillance, a concept that is receiving increasing attention in surveillance studies (Hollander & Einwohner, 2004), not only as an umbrella term for protest or opposition practices taking place between the surveyor and the surveilled, but as a “more complex, multi-directional and multi-actor (…) process” (Martin, van Brakel & Bernhard, 2009, p. 214).

The third part of the article introduces third-generation, “private” P2P networks and explores how developers and users of these systems seek to take their main weapon away from copyright holders, by placing a special emphasis on a friend-to-friend paradigm that allows users to join the system only by personal invitation of another user (Rogers & Bhatti, 2007; Le Fessant, 2009; Wood, 2010), shaping privacy as de facto invisibility from pervasive surveillance.

The fourth and conclusive part opens up to a conception of P2P systems as possible tools for the materialisation of a social, political and economic “opportunity” for Internet-based services. It suggests that, while paramount for putting into perspective the evolutions and developments of P2P systems over the last decade, the “surveillance-and-counter-surveillance” paradigm may entail an exclusively “defensive” conception of privacy; a conception that, while an important one, is only a part of the story. Other parts – enacted daily in a number of projects and applications for P2P Web search, social networking, data storage that are being developed since 2006 – are user empowerment through a better and more nuanced control over personal information, reconfiguration of the balance between users’ and service providers’ rights over personal data, and removal of intermediaries in sharing and communication activities – parts that if neglected, may lead to overlook the potential of P2P as an effective, scalable and stable way to distribute, exchange and communicate online, in a variety of ways.

---

1 The generations in P2P file-sharing systems are differentiated, and accounted for, according to benchmarks for significant changes in the sharing technologies underlying such systems. A number of authors in both computer science (Oram, 2001; Schoder & Fischbach, 2003; Laflaquière, 2005) and law (Litman, 2005; Elkin-Koren, 2006; Wood, 2010) find themselves in agreement that the technological changes detailed in section 2, in terms of both the sharing practices enabled and the juridical measures implemented, are indeed “turning points” that allow to account for a new “generation” of P2P systems. It is noteworthy, albeit outside the scope of this paper, that a “fourth generation” of P2P systems, besides the three described here, is increasingly identified in the tools that allow for streaming media (e.g. radio and television) to be distributed over a P2P network by means of a BitTorrent-like swarming technology. These tools will be addressed in more detail in the conclusive section of this paper.
1. Surveillance, privacy and P2P systems

Since the inception of the World Wide Web and the proliferation of Internet-based services that has characterized our lives in the last decade (Di Maggio et al., 2001), a number of voices have raised to warn Internet users about their daily life being increasingly monitored, in the form of traces whenever they ask for, or provide, goods and services, whether they seek information or move in real or virtual spaces (Rodotà, 2006). Thus, everyone’s virtual social representation is gradually shaped with reference to the information left by each interaction and transaction, scattered in a variety of databases, data collections, and networks, what interests us most here.

1.1. Information privacy, a multi-faceted concept

While the concept of privacy may be applied, depending on context, to an individual or group’s ability to protect or selectively reveal all kinds of personally sensitive information, encompassing anything from physical features to political orientations, the extensive theoretical discussion of the notion goes back to how it is defined, what exactly it means and what aspects it includes. Of particular relevance for this article is the discussion of the so-called “information privacy”, the relationship between the protection and dissemination of personally identifiable, sensitive data, mostly by means of collection and storage technologies.

Most of philosopher Luciano Floridi’s work addresses the ways in which the commodification of ICTs and the emergence of a global information society are increasingly challenging the right to informational privacy. To him, the most appropriate way of understanding the increasing impact that ICT is having on informational privacy is in terms of an ontological revolution that has changed the nature of the world of information from analogue to digital. By comparing three interpretations of informational privacy, reductionist, ownership-based and ontology-based, Floridi comes to the conclusion that informational privacy should be considered a fundamental right and hence that, as for other fundamental rights, by default the presumption should always be in favour of it (Floridi, 2005).

Addressing the crucial question of what is to be included in the definition of privacy and how, Herman Tavani points out – in relation, more specifically, to information(al) privacy – that “the debate about how privacy is best defined is closely related to the question of whether privacy should be viewed as a full-fledged right, or simply in terms of one or more interests that individuals have” (Tavani, 2008, p. 133); he notes how in the case of P2P computer networks, it may be possible to apply the same “publicity” principle used in the case of cookies technology to determine whether personal information accessible to the technology should be declared a “normatively private situation”, i.e., when all the information about the implications of the technology is disclosed, users are enabled to make an informed decision about whether to accept or reject it (Tavani, 2008, p. 160).

Interestingly, while addressing the interplay of technology and privacy in the digital age, Daniel Solove points out how a definition of privacy, and of the ways to protect it, cannot afford to neglect the examination and the building of an architecture regulating the ways in which information is collected and used: thus, he points out, “privacy must focus not merely on remedies and penalties for aggrieved individuals but on shaping an architecture to govern the ever-increasing data flows of the Information Age” (Solove, 2004, pp. 10-11). Along the same lines, Bennett and Raab argue that political scientists need to pay particular attention to the “regulation of personal information as a major ‘currency’ of both the contemporary capitalist organization and of the modern state”: rather than considering the inherent difficulty of providing a general definition of privacy a “show-stopper”, the privacy scholar should take the “questionable assumptions about individualism, and about the distinction between the realms of civil society and the state” as her starting point, and question these very assumptions so as to unearth privacy as a “political” object (Bennett & Raab, 2003, p. 6).

Looking at the “other side of the privacy equation”, scholarship has also considered if and how in specific contexts, the balance between individual privacy and societal responsibilities in the building of collective safety and security should be challenged, and debated much more than it currently is. In his controversial 1999 book The Limits of Privacy, Amitai Etzioni argues that privacy is in fact too
protected and privileged, to the detriment of common good and social sustainability. Citing among others the example of cyber terrorists exchanging much more easily with the rise of encrypting techniques, Etzioni controversially argues that public safety is nowadays "systematically neglected out of excessive deference to privacy": according to him, privacy should be defined and considered as an "individual right that is to be balanced with concerns for the common good – or as one good among others, without a priori privileging any of them" (Etzioni, 1999, p. 4).

All of these different conceptualizations and scholarly treatments of the concept of privacy show how the challenges posed by the current management of Internet-based services, and the governance of public space without borders that the Internet helps to create, determine at once new needs and new risks for its users, in their double capacity of consumers and citizens, as well as individuals. In recent times, the right to privacy has not only entered or updated national constitutions and transnational charts, but with other rights – such as the safeguard of environment, access to information, right to communicate, consumer protection and enlarged/more direct participation in administrative decisions – it is one of the so-called "third generation rights," the expression of new requirements raised by fast-paced developments in science and technology (Evans, 2007). Indeed, in the context of personal data protection in sharing and networking activities online, the concept of privacy is today acquiring yet new facets, due to the reconfigurations, prompted by cyberspace, of the relationship between property and personal data, and to the increasingly permeable boundaries of digital space that open up new possibilities for consumer surveillance (Katyal, 2005). Both these notions (the relationship between property and exposure of sensible data, and the permeability of boundaries of interactions in the digital space) are both very relevant to the main argument presented here. Thus, for the purpose of this study, I follow Katyal's conceptualization of "information privacy" as a person's ability to control what personally identifiable information she reveals about herself over the Internet, and to manage other people's access (notably "unreasonable searches or seizures") to that information; ability that is placed under considerable strain and constantly reconfigured by the rapidly evolving Internet ecosystem of encryption, anonymity and surveillance tools.

1.2. Information privacy, copyright protection and pervasive surveillance

David Lyon has repeatedly defined the Internet as a "web of surveillance" (1988, pp. 37-47; 1998), in which monitoring of the user/consumer is a growing phenomenon. The increased participation, especially of particular groups of users, to electronic commerce also raises issues of intrusion and surveillance (Castells, 2000). The ways in which personal data are treated are likely to influence, reinforce or weaken power structures, be they market- or politics-dependent, that rely on the tampering of privacy and related rights in order to thrive and be successful (Musiani, 2010); surveillance is implicated in the maintenance of inequalities and divisions, and raises challenges to identity anytime individual are not in full control anymore of the capacity to control communication about themselves (Lyon, 1998, p. 39). Along these lines, scholars interested in the concept of resistance have emphasized the "digital, algorithmic and rhizomatic aspects of new surveillance practices" (Martin, van Brakel & Bernhard, 2009, p. 215), noting how traditional ways of understanding surveillance as an exclusive relationship between the surveyor and its subjects excludes the emerging "surveillant assemblages" that are instrumental in conducting the surveillance, transforming its purposes and hierarchies, as well as the institution of privacy (Haggerty and Erickson, 2000).

Interested in addressing, from a juridical standpoint, the relationship between information privacy and copyright protection, Sonia Katyal has observed (2004; 2005; 2009) that a number of increasingly ubiquitous and invasive copyright enforcement strategies have seen the light in the last few years, that share the common trait of relying on private mechanisms of surveillance for their execution and control. According to Katyal, these techniques of surveillance demonstrate copyright’s increasingly tenuous relationship with information privacy: in the past, legislators and scholars have focused their attention on other, more visible methods of surveillance relating to employment, marketing, and national security, yet the phenomenon of “piracy surveillance”, as she describes the extrajudicial systems of monitoring and enforcement that detect, deter, and control acts of consumer infringement, is completely distinct from other consumer monitoring procedures, and is “in-
completely theorized, technologically unbounded, and, potentially, legally unrestrained” (Katyal, 2005, p. 227). As Frances Grodzinsky and Herman Tavani (2005) point out, surveillance has taken on such a prominent role in the debate at the crossroads of privacy, property and expression because when the burden was first placed on copyright owners to identify infringers on the network, this legitimized somehow the rise of “an entire new industry that has content owners searching the Internet for potential infringers” (Grodzinsky & Tavani, 2005, p. 247).

This “new surveillance”, as Katyal labels it, implies fundamental alterations of intellectual property rights, from defensive shield into offensive weapon, in order to record consumer activity and possibly enforce particular standards of use and expression, while proscribing activities deemed unacceptable; thus, the conflict between privacy and piracy is important not just as a showcase of an overlooked mode of surveillance, but also as a demonstration of the need to resolve conflicts between them in ways that reflect and protect the relationship between modern technology and personal freedoms (Katyal, 2004).

Along the same lines, Niva Elkin-Koren remarks (2006) that the interrelationship between law and technology often focuses on one single aspect, the challenges that emerging technologies pose to the existing legal regime, creating a need for legal reform; juridical measures involving technology both as a target of regulation and as a means of enforcement should, however, take into account that the law does not merely respond to new technologies, but also shapes them and may affect their design (Elkin-Koren, 2006, p. 15). Surveillance systems that fail to take this into account are likely to re-shape the technology in ways that diminish or impede their potential for socio-economic benefits (Elkin-Koren, 2006, p. 21), or sacrifice the most valuable aspects of cyberspace, extracting from it fundamental principles of informational privacy for the sake of unlimited control (Katyal, 2004; 2005).

While this conception of surveillance is not entirely attributable to the development of P2P technologies, or to the explosion of “cyberspace piracy” mostly through the use of file-sharing P2P tools, it is especially well illustrated within this domain as a representative of the “paradoxical” nature of the Internet, that “both enables and silences speech, often simultaneously” by offering “both the consumer and creator a seemingly endless capacity for human expression […] alongside an insurmountable array of capacities for panoptic surveillance” (Katyal, 2005, p. 228). Privacy-related issues concerning P2P networking illustrate this point well. Users of P2P networks share idle computing resources such as free bandwidth, storage space, and computing power, and provide the system infrastructure itself, which makes P2P systems “the economically and technologically optimal vehicles for digital content distribution” (Wood, 2010, p. 6), less vulnerable to bandwidth restrictions and more scalable. This also means, however, direct access to the data packets in a peer’s data stream by other peers, and therefore, that the data stream may be compromised by the other peers in the network that are contributing to the very process of data transmission (Musiani, 2010). So, users need to have some knowledge about the software they are using, and they need to be aware of what types of materials and information are being shared (or those they do not want to share), as it is “quite possible to share the entire hard drive, including sensitive information such as mailbox and private documents” (Suvanto, 2005). The direct connection between the peers also implies, in many cases, that in order for the connection to be established, the P2P file sharing software needs information such as their IP address. In some of the first, most popular P2P file sharing systems, this address has been used openly and directly exposed, thus raising anonymity problems (Suvanto, 2005). It is no surprise, then, that in order to counter such weaknesses, and their exploitation by whatever entity willing to tamper with the stability and integrity of the network, more recent P2P developments have been moving in two directions, content encryption and improved anonymity (Li, 2007; Musiani, 2010).

The deployment of pervasive surveillance measures with the aim to ensure liability and responsibility for a specific type of P2P traffic – measures constantly evolving and changing in parallel with P2P tools themselves – can account for many of the most important steps in the genealogy of P2P file sharing systems. How this has unfurled over the last ten years, and why it is relevant to understand the ongoing, increasing privatization of P2P networks, is the subject of the next section.
2. A genealogy of P2P file sharing networks

Far from being a recent development that would have started with Napster, P2P technology may be understood as one of the most ancient network architectures in the world of telecommunications (Oram, 2001). In this sense, Usenet, with its discussion groups, and the early Internet, in the form of ARPANET, may be considered as P2P systems. As a consequence, some scholars are arguing that P2P is taking the Internet back to its roots, to an era when each computer had equal rights on the network (Minar & Hedlund, 2001). The lowering of costs and the increasing availability of computational capacity (processor cycles), bandwidth, storage capacity – all accompanied by the far-reaching development and acceptance of the Internet – have determined new fields of application for P2P networks. In a recent past, this has entailed an impressive increase in the number of P2P applications and controversial discussions concerning their limits and performances, as well as their political, social and economic implications (Schoder, Fischbach & Teichmann, 2002; Smith, Clippinger & Konsynski, 2003).

The last decade has witnessed a series of evolutions in the field of P2P technologies. The wide success that applications of such technology have enjoyed has certainly been an important catalyst of their developers’ creativity, and as a consequence, of improvements in the effectiveness of P2P tools; however, developments in the field have shaped and have been shaped in return, to a large extent, by political and juridical constraints, notably the lawsuit threats that have been put on the table by some of the main actors in the digital content industry. Three “generations” (Laflaquière, 2005; Wood, 2010) of technologies have, therefore, seen the light.

I present here the different moments in the history and formation of a P2P “genealogy”. I interpret them as a dialectic process between the promulgation of laws or juridical measures posing restrictions to file sharing activities taking place on P2P networks, and the advancements and technical responses that have followed the application or the implementation of such measures. In this sense, the genealogy of P2P file-sharing systems is also a history of resistance towards regulation of user behaviour by means of digital surveillance: resistance that may be defined as successful, but at a cost that the following sections seek to discuss.

2.1. Centralised hybrids: the first generation

The first generation of P2P was actually a centralised system, using a networking model called one-to-many, allowing for a unique support for the diffusion and sharing of files by means of different nodes, while a central server supervised and controlled the traffic (Laflaquière, 2005).

Napster, Shawn and John Fanning’s creation, was in 1999 the first, most famous and widespread among these systems, including a central server aiming at speeding up, increasing and easing user activities within the network. However, files were stored and distributed, by means of the terminals owned by end users – not on and by the server. The function of the server was to establish connections between users, and facilitate file searches initiated by the users. To this aim, the server used (and stored) the list of available files on the network, and of users’ IP addresses. Users were then able to search the list for files available on the ensemble of users’ computers, and the P2P programme would then establish a connection between the two or more interested users, that would directly transfer the file between their own terminals (Elkin-Koren, 2006).

The centralised model has often been preferred in the first phases of use of this technology for file sharing purposes, as the central list or index facilitated the identification of files in a rapid and effective way. As the users needed to access the system by means of a central point, however, it was possible to invalidate the entire system by disconnecting the server, a feature that may potentially result in a strict control exercised upon users. Last but not least, users needed to register on the system so as to be recognized and connected; as a result, the service provider was able to know the identity of every user, and what he was downloading (Lemley and Reese, 2004).
2.1.1. The central index and Napster’s demise

At the end of the Nineties, the widespread propagation of digital technologies had begun to alert a powerful lobby, composed by the main actors of the digital content industries. This lobby heavily pressured the United States Congress (Napster had its legal residence in the country) for a stronger protection of copyright and intellectual property rights as a whole. The lobbying produced its results, and soon, the industry of digital content enacted an aggressive strategy of lawsuits and processes, destined to implement the set of regulations that had just been approved. The content industry started by aiming at commercial entities, summoning them to court for contributory and vicarious liability in copyright violation, because if just one of the ongoing trials was won, it would have been sufficient for having the main server stopped, thus compromising the diffusion mechanism in its entirety (Elkin-Koren, 2006; Wood, 2010).

The same features of the centralised model that promoted its technical efficiency made it, in this case, more fragile vis-à-vis accusations of vicarious liability in the copyright violations enacted by users of the service. In 2001, a tribunal of the United States ruled that Napster, the most widespread P2P service provider up to that point, was to be deemed guilty of contributory and vicarious liability – and elicited a complex debate among lawyers and IT practitioners alike on the dangers of a doctrine that assimilated capability to liability (Elkin-Koren, 2006; Murray, 2010). The court cited, as the main rationale underlying this deliberation, Napster’s capacity to control the behaviours of all users of the service by means of the index and the central research function (Wood, 2010). More specifically, the responsibility of Napster was assessed in the following terms by the ruling: the services provided by the application were specifically designed for allowing users to locate and distribute musical files; Napster was therefore materially contributing to its users’ violations, as evidence suggested that Napster had knowledge of the violations, but had not modified or “purified” the system. Moreover, the court ruled that the central index was giving Napster both the material capacity and the right to supervise its users, and that this list allowed the service provider to locate copyright-infringing material, thus allowing it the right to terminate the access of infringing users to the system (A&M Records, Inc. Vs. Napster, Inc., 2001). In short, according to copyright owners, Napster was failing to enact surveillance measures itself while having the capacity to do so, and was therefore liable.

2.2. Towards complete decentralisation: the second generation

The juridical decision concerning Napster marked the demise of hybrid-centralised P2P networks and triggered the development of a new generation of networks, less efficient in certain respects, but more decentralised. The “architects” of the new-generation P2P networks started to design these networks while giving priority to the objective of minimizing the risk of being denounced for violations (Elkin-Koren, 2006); it has been noted that P2P technology constitutes a clear example of how regulations that imply an attribution of responsibility influence product design (Baram, 2007). The main purpose of such P2P systems, thus, became to escape surveillance: prevent whatever entity willing to do so to retrace violations and, even more importantly, to derive the violator’s identity from the trace.

The first and most important consequence of the Napster verdict consisted in the initiative of the P2P operators that had followed it chronologically to decentralise their systems, with the aim to avoid the attribution of responsibility that user-initiated exchanges would imply (Wood, 2010). Second-generation technology, thus, directly connected users, without need for routing information to pass through a unique central server (as it was the case for hybrid models). Furthermore, there was no list, directory or central index in these decentralised models; while this made direction of requests and searches more complicated, it was also preventing service providers from storing user-related information, identifying users upon request of third parties, and directly facilitating connections (Rahman et al., 2009).

Partially distributed, hybrid systems like FastTrack (used by Grokster, a popular file-sharing application) switched from a central index to a function which attributed to some of the computers
connected to the system (without awareness or initiative by their users) the operational role of super-nodes – decentralised connection points indexing available files and managing research queries. These connections aimed at making the search function more rapid and avoid bottlenecks.

2.2.1. Gnutella and the serial routing of searches

Instead of using super-nodes, the models that immediately followed Napster, such as Gnutella, routed searches serially (Wood, 2010), across all nodes available on the network at a given moment. Thus, the two models created their communities of users by reuniting the IP addresses of users connected to the Internet, allowing for the creation of a “branch” of the network aimed at direct communication between users. After entering the network, a user could search for the files on every computer it was connected to, communicate and share files with these other users, without intermediaries. Gnutella has therefore distanced itself from a potential contributory liability – for, as it is an open source protocol, every user can write an application in the form of a Gnutella client, and as a consequence, there is no unique identified operator to whom responsibility for the infraction may be attributed (Biddle et al., 2001).

In completely decentralized systems, the branching system often makes searches and exchanges slower, but service providers for decentralized systems have less control on users of the service. Without a central server, a provider has limited or no capacity to supervise a violation, and cannot take out of the system one or more files, or users, as a response to an infringement of copyright law. Decentralized systems are also more difficult to keep under surveillance and to neutralize, as there is no central point; many of them, in addition, were built on open-source protocols, thus, the neutralization of parts of a system was of a very limited effectiveness as more experimented users were able, with little effort, to adapt copies of the programme’s code so that the system could continue to function.

2.2.2. The stabilization of contributions with BitTorrent

Decentralized systems’ fragmented design is also bound to contribute to the free-riding phenomenon, typical of users who wish to download but not to upload – a behaviour that may, in extreme cases, compromise the effectiveness of the system, as storage costs are not distributed between users in a balanced way. The main reason for the rise of protocols such as BitTorrent was the necessity to stabilize levels of contribution on P2P networks; these services make the collaboration mandatory as they establish limitations on a user’s download pattern according to how much, and what, that same user is uploading.

The first versions of BitTorrent included an intermediary tracker service, functioning as a searcher and aggregator of files, allowing for both uploading and downloading. Trackers kept a log, specifying which users were downloading which file, the material location of the file, and of its fragments. Logs were, nonetheless, crucial within the frame of processes against copyright infringers, as they permitted (as servers before them) to identify suspects of violations reliably.

Thus, following versions of BitTorrent eliminated the necessity of trackers, after the service had been stopped and its design reorganised after multiple lawsuit threats. BitTorrent developers considered that, with no centralized features, the new design would have been able to make it more difficult for copyright owners to trace and stop illegal file sharing. SuprNova, one of the most widespread BitTorrent tracker service, was forced to halt its service after the Motion Picture Association of America (MPAA)’s campaign against illegal file sharing. Many users of the service thus went back to the Gnutella protocol on Grokster, until this service was, in turn, shut down. Grokster was considered liable for users’ illegal behaviour in light of the fact that the service had distributed its product with the explicit aim of promoting violation, and taking a commercial advantage from it (Metro-Goldwyn-Mayer vs. Grokster, Ltd., 2005).

Other file sharing service providers in the United States sought juridical refuge in the Sony vs. Universal City Studios decision. The sentence established that the distribution of equipment destined to copy was not liable for contributory infringement if the programme was also capable to
promote substantial non-infringing uses (Sony vs. Universal City Studios, 1984; Wood, 2010, my italics). After this decision, companies that still explicitly sought to distribute software destined to file sharing made active steps towards avoiding the same kind of liability that had been attributed to Grokster. For example, LimeWire (a programme based upon a protocol of Gnutella type) started requiring its users not to infringe copyright as part of its Terms of Use, and included within its website an educational section on unauthorized file sharing.2

3. Third-generation P2P networks and the quest for invisibility

With the creation of a decentralised P2P technology, the intention of developers to answer lawsuit threats by means of technology has been especially relevant. Control is removed, totally or partially, from the hands of service providers – thereby making the tracing of user behaviour more difficult.

But second-generation P2P networks, such as Gnutella, BitTorrent and other programmes based on the same protocols, had one further common feature that placed them under serious threat: they were, in a number of ways, lacking in anonymity (Biddle et al., 2002; Wood, 2010). Giving room for the identification of the network’s end points, this type of peer-to-peer network revealed the IP address – and often, other information as well – of nodes in the network. Networks were decentralised, but not private, as they were conceived in such a way that peers would be able to communicate with all other peers in the network. Activities and uses on these “public” P2P networks were therefore not impossible, albeit difficult, to trace back – making it possible, once more, to detect violations and identify infringers in view of a lawsuit (Laflaquière, 2005).

The following step in P2P evolution came when, after the juridical offensive on commercial entities such as Napster, the digital content industry turned its attention to ordinary citizens. Thus, in 2006, the Recording Industry Association of America (RIAA) sued over fifteen thousand individuals, on the grounds of copyright violations (Wood, 2010). At the demand of users willing to circumvent this attribution of responsibility, P2P developers began working on a number of improvements, allegedly aimed at providing users with anonymity, privacy, and better control on personal data. This version, the most recent development as of today in the field of file-sharing distributed networks, is now in the position of effectively hiding user behaviour and raising new questions and frameworks of analysis for copyright on the Internet (Lasica, 2005; Wood, 2010).


In late 2001, among the technical and juridical concerns that followed Napster and preceded Gnutella, four engineers employed by Microsoft’s research group on network security introduced, in a subsequently very influential paper, a new term that would account for the Internet of the underground: “darknet” (Biddle et al., 2001). In this paper, the four engineers described the “darknet” as a “collection of networks and technologies used to share digital content” (Biddle et al., 2001, 1). The word soon spread to general media, and started to be used as a reference to a number of “clandestine” practices and tools on the Internet. Between 2003 and 2005, the term “darknet” became the label of all sorts of cyber-activities looking uncertain and menacing, from private virtual clubs to heavily-secured online databases, non-traceable with mainstream search engines, from cybercrime and spamming activities to other “obscure places” on the Internet, haven of illegality or at best, ambiguity (Rivlin, 2003; Wood, 2010).

2 “Using P2P Software Safely” at http://www.limewire.com/fr/legal/safety. As of October 26, 2010, the website of LimeWire is nonetheless under a court order to stop distributing its software. A sign obscuring the software’s home page reads: “LimeWire is under a court order dated October 26, 2010 to stop distributing the LimeWire software. (...) LimeWire LLC, its directors and officers, are taking all steps to comply with the injunction. We have very recently become aware of unauthorized applications on the internet purporting to use the LimeWire name. We demand that all persons using the LimeWire software, name, or trademark in order to upload or download copyrighted works in any manner cease and desist from doing so. We further remind you that the unauthorized uploading and downloading of copyrighted works is illegal.” (http://www.limewire.com/, last accessed November 26, 2010).
In parallel, “darknet” started to become the label for private and anonymous P2P networking tools, as opposed to their “public” predecessors (Wood, 2010). The questions of surveillance, privacy and security were introduced for the first time in a “darknet”-related juridical work in 2004, where such a network is defined as the collection of networks and other technologies allowing people to share digital content “with little or no fear of detection” (von Lohmann, 2004). This counter-surveillance element is reprised by describes elsewhere the “darknet” as a network of people using closed spaces – safe havens, virtual and real at once, with little or no chance to be detected – in order to share digital content with others, so as to avoid “the restrictions on digital media imposed by entertainment companies” (Lasica, 2005, p. 45). The vision of private networks coming out of such works is that of a supermarket of digital media, with something of a “wild west” mentality, fully able to rival products and services provided by the main actors of digital content industry by means of privacy – interpreted, and de facto treated by users and developers, as invisibility.

More generally, the term “darknet” refers to all private networks with file sharing purposes, that may be broadly defined as networks, or network of networks, distributed and decentralised (with no central index) including functions of privacy, security (encrypting functions), and user anonymity, with the main goal of sharing information with certified members of the network.

The main goal of a private P2P network is to create a closed system allowing to communicate and to exchange securely, so that detection or penetration by external entities, such as governments or companies, is avoided. Users may download or upload, and inject content in an anonymous fashion, so that outsiders or strangers to the network may not hold enough information to identify its users. Improvements in privacy and security, on which developers have been focusing most recently (Suvanto, 2005), give room to improved anonymity, and the lack of a public entry point into the network makes it difficult, and more often impossible, for strangers to find out what is happening, more specifically what is being shared, on this type of network.

3.2. Among friends, like in social networks

Beyond the connotation of illegality it may have assumed, the principle underlying file-sharing “darknets” is what users commonly refer to as “friend-to-peer” (F2F) networks, meaning that direct connections are established between recognized friends only.3

This enables a response to surveillance activities, such as blocking and filtering, potentially a lot more effective than approaches relying only on encryption (aimed at preventing the service provider from filtering exchanges) and indirectness (preventing one user of the network from knowing the user he is communicating with, thereby enhancing anonymity; Le Fessant, 2009). The limits of both these responses – still subject to “man-in-the-middle” or “Sybil” attacks, respectively4 – can be overcome with the introduction of a friend-to-friend networking paradigm within P2P networks: indeed, what social networking (popularized by Facebook) would look like, if social links were also network links (Figueiredo et al., 2008; Musiani, 2010).

In friend-to-friend connections, every user hosts his personal page on his computer, with, in particular, all information or data he considers personal. He authorizes his friends, one by one, to access his computer. For this, he sends to each friend a secret key, which will be used by the friend

---

3 This model shares some similarities with the consumer-to-consumer (or C2C) communities (described e.g. by Murray, 2010, 41), inasmuch as the design of the networks underlying such communities allows for the removal of intermediaries and enables direct interactions between the owner and the potential recipient of a resource; moreover, it emphasizes the need of trust between the two or more actors in the exchange, which is a crucial feature in enabling friend-to-friend transactions/connections in F2F networks, as well. However, the C2C model is generally meant to describe trading communities, and more specifically, trading communities that share material, not digital goods, albeit by means of electronic transactions; therefore, most of the issues raised by copyright, surveillance and anonymity, that are at the center of the present paper, are less important or outright negligible in the context of C2C.

4 A “man-in-the-middle” attack is a form of attack on mutual authentication between correspondents in a network, in which the attacker makes independent connections with the victims and relays messages between them, making them believe that they are talking directly to each other over a private connection, when in fact the entire conversation is controlled by the attacker; in a Sybil attack, an attacker subverts the reputation system of a network by creating a large number of pseudonymous entities, using them to gain a disproportionately large influence in the network.
for the first connection to his computer by means of the P2P application. At the moment of this first connection, a new secret is shared, which will be used for all the following connections: the first secret is not subsequently reusable, avoiding any intercepting by a third person (Le Fessant, 2009).

This mechanism of identification and content distribution is currently used in a number of applications for sharing of personal data and information, such as pictures and amateur videos (Peerpyle, 2007; Move&Play, 2009; GigaTribe, 2005, most recent version 2010). It enables quicker access to content (as it is directly available on the user’s computer, without need of copying it on a site), and makes it accessible only to friends using the same programme and subject to identification. GigaTribe summarizes the principle of private file sharing this way: “Today due to the wide range of digital entertainment, everyone has a continuously growing virtual library of photos and videos. All of these files are directly created or saved on your hard drive. An easier and more secure solution for sharing this library with your friends is to allow them to access your hard drive. Your files remain up-to-date on your personal hard drive in security and eliminating an extra step in transferring these files to an outside server.”

Figure 1: GigaTribe screen capture (http://www.gigatribe.com/en/screenshots, 2010)

The mechanism of adding friends, based upon the exchange, by a private channel, of a secret key that can be used just one time, has two implications. On one hand, it enables automatic identification, from that moment on, of all people connecting to a particular user’s computer – and cuts off any non-identified entity. The user knows every other user connected to the computer, and can choose the extent of trust he gives to each of them – as a consequence, choose the personal data they can access: Le Fessant describes it (2009) as a shift from an effort to maintain user anony-
ity, towards reciprocal knowledge of identity between private peers so as to reinforce user security. And while some applications restrain the use of private file sharing to personal data and information, placing them in the realm of "legality", there is no reason why other types of information, such as copyright-protected music, movies or programmes, should not be shared as well. Indeed, they increasingly are – and with "little or no fear of detection".

3.3. Friend-to-friend networks: the key to invisibility

Private file sharing by means of friend-to-friend P2P networks, by placing a special emphasis on a paradigm that allows users to join the system only by personal invitation of another user, is, chronologically, the last so far in a series of attempts by surveillance technologies and sharing technologies to outrun each other. By enforcing in an increasingly strict manner the juridical and technical measures against previous generations of "public" or semi-private P2P networks, the actors of digital content industry, and occasionally States themselves, have led developers and users of these systems to seek further ways to neutralize the main weapon of copyright holders, identification devices – and, in the process, all involved actors have contributed to the shaping of a definition of privacy as de facto invisibility from pervasive surveillance.

<table>
<thead>
<tr>
<th>P2P Generations</th>
<th>Sharing</th>
<th>Privacy</th>
<th>Surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>First generation, Centralised (one-to-many) hybrids</td>
<td>Central server to facilitate connections/ direct, P2P-type transfers Ex.: Napster</td>
<td>Index and central research function make it transparent, for whoever undertakes a search, which user has what content. Users register on the system.</td>
<td>While it is easy to track down users, surveillance is not aimed at them, but selectively at company servers themselves, as stopping the main server results in the entire system’s demise.</td>
</tr>
<tr>
<td>Second generation, Decentralised networks</td>
<td>Supernodes, serial routing of searches, stabilization of contributions. Ex.: Grokster, Gnutella, BitTorrent</td>
<td>No list, directory or central index. This complicates direction of requests and searches, but prevents service providers from storing user-related information and identifying users upon request of third parties. However, lack of anonymity at end points – networks are decentralised, but not private.</td>
<td>Surveillance becomes more pervasive inasmuch as it concentrates at end points – users/ordinary citizens. The “publicity” of these networks gives room for copyright owners to retrieve IP addresses and other information “at the edges”.</td>
</tr>
<tr>
<td>Third generation, Darknets/Private P2P/Friend-to-Friend</td>
<td>Direct connections, encryption, anonymity, secret keys. Ex.: GigaTribe, Groove, WASTE (and other applications for sharing of personal data and information, such as pictures and amateur videos)</td>
<td>Direct connections are established between recognized friends only. Every user hosts his personal page on his computer and authorizes friends, one by one, to access his computer, by means of secret keys.</td>
<td>Blocking and filtering surveillance activities are powerless. Provided that only trusted users join the network, the system is closed and invisible.</td>
</tr>
</tbody>
</table>

Table 1: Co-evolution of privacy, surveillance and sharing technologies across P2P generations

4. Conclusions: Beyond the “surveillance-and-counter-surveillance” paradigm?

This article has addressed the ongoing “privatization” of P2P systems. It has argued that this process may be intended as the last step of P2P technology’s recent history that has seen the alternation between forms of pervasive surveillance and attack of such systems, and reactions by developers and users to such restrictive measures. Moreover, it has argued that as a result of this dynamic, privacy in P2P networks is increasingly coinciding with “mere” invisibility vis-à-vis the rest of the Internet ecosystem. As a conclusion, I open up to some dimensions in the conceptualization of privacy in P2P that a special emphasis on surveillance and counter-surveillance measures may be leading to neglect.
Indeed, P2P technology is also increasingly deployed as a tool for the materialisation of a social, political and economic “opportunity” for Internet-based services. A variety of P2P-based projects and applications are currently seeing the light, proposing to explore the potential of P2P technology fully by serving diverse necessities of use (P2P Web search, Faroo, beta 2006; distributed data storage, Wuala, closed alpha 2007; distributed social networking, Diaspora, alpha 2010).

In doing so, they are elaborating in practice a more articulate concept of privacy, one that includes – beyond the need of user protection from external attacks and the temporary outcomes of the competition between surveillance and counter-surveillance measures – issues such as user empowerment through an improved and more nuanced control over his personal information; the inherently social functions of P2P technology; the removal of intermediaries (in the words of Elkin-Koren, 2006) in sharing and communication activities; the plurality of possibilities offered by the “legal” uses of decentralised network architectures.

![Diaspora Alpha](https://joindiaspora.com/)

This is, perhaps, the main lesson that surveillance studies as a field can learn from the history of P2P systems: indeed, as pointed out by several scholars of resistance to digital surveillance (e.g. Martin et al., 2009), technological surveillance is better understood as a complex multi-actor, multi-level framework, able to look beyond the binary opposition between surveyor and surveilled. The ways of evading, countering, circumventing pervasive surveillance have deeply informed the socio-technical evolutions and developments of P2P systems, an essential part of the history of this technology; but a thorough understanding of P2P as both target and source of law, and co-producer of user rights, cannot do without a careful observation of how this technology prompts reconfigurations of where personal information is stored and data is exchanged; of the frontiers between nodes and networks; of how available tools are performed by developers and users. In a nutshell, reconfigurations in the attribution, recognition and modification of the rights of users and service providers.

The reading of P2P history proposed here is also an invitation to renew the political debate on decentralised distribution systems. It aims at underlining the importance to retrace how these reconfigurations happen within the context of a variety of uses beyond file sharing – thus, in “legal” territory with respect to the “copyright wars” we have provided accounts of here. Not only a richer definition of privacy, beyond its framing as “escape from surveillance”, but also a more articulate
conception of “legality” of P2P technology than the avoidance of its use for piracy purposes, are likely to emerge from a careful observation of these practices and devices. In order to start building on it, though, actors involved in “copyright wars” need to acknowledge first and foremost that juridical and technical measures aiming at the elimination of a specific type of P2P traffic – as well as the responses they may elicit – could entail the loss, or the serious damaging, of potential economic and political benefits of the P2P model.

Without this acknowledgement, and despite the rich ongoing debate on the definition and conceptualization of informational privacy in the digital age, privacy on P2P networks may soon be reduced to a simple “invisibility from the guardians”, as a disguise for a reckless Far West of exchanges. It may be treated merely as such, when discussed in the political – and controversial – venues that shape the future of the Internet.

References


About the Author

Francesca Musiani

PhD candidate and researcher at the Centre de Sociologie de l’Innovation (CSI), a mixed research unit (unité mixte de recherche) of Mines ParisTech and CNRS located in Paris, France. Instructor at Université Pierre et Marie Curie in Paris, France. Her dissertation research focuses on features and implications of alternative peer-to-peer technologies; her research interests include ICTs for development, the evolution of legal systems and rights in the digital age, and online participatory practices of fan communities. She maintains collaborations with the University of Padova, Italy, where she obtained her first Master’s degree in Organizational Communication; she also holds a Master of Arts degree in International Law and the Settlement of Disputes from the United Nations-mandated University for Peace in San José, Costa Rica, and has worked as a journalist at the United Nations Headquarters in New York. She is the author of Cyberhandshakes: How the Internet Challenges Dispute Resolution (...And Simplifies It), published with EuroEditions in 2009 thanks to a publication scholarship (beca de publicación) from the European Foundation for the Information Society, and the co-recipient of the 2010 Second Charlemagne Youth Prize of the European Parliament for her contribution to the volume You Are Here (Broken Dimanche Press, 2009).