Cloud Computing : Centralization and Data Sovereignty

Primavera de Filippi, Smari Mccarthy

Cloud Computing: Centralization and Data Sovereignty

Primavera De Filippi, Smari McCarthy

Abstract: Cloud computing can be defined as the provision of computing resources on-demand over the Internet. Although this might bring a number of advantages to end-users in terms of accessibility and elasticity of costs, problems arise concerning the collection of personal information in the Cloud and the legitimate exploitation thereof. To the extent that most of the content and software application are only accessible online, users have no longer control over the manner in which they can access their data and the extent to which third parties can exploit it.

1. Introduction

"Cloud computing" has become a popular, yet poorly defined term in online service provision. By aggregating a large number of computing resources together into a few clusters of very large dimensions, Cloud computing has created an imbalance in authority structures that is very similar to the structural changes witnessed during the Industrial revolution. Just as the industrial revolution has progressively alienated workers from the means of production, today, most of the means of online production (in terms of hardware, software, content or data) are concentrated within the hands of large Internet service providers.

Although Cloud Computing constitutes a great opportunity for small start-ups to compete in the market for online services without the need to make massive initial investments, exporting all their infrastructure and data into the Cloud is decreasing the capacity of users to control the manner in which their resources are being held. Given that everything can be stored, processed, or executed on any computer system regardless of its whereabouts, most of the means of production are increasingly owned or at least *de facto* controlled by large companies.[1]

The trend is clear. Resources are moving away from end-users, towards centralized systems that possess huge processing power and storage capacities. Users' devices are devolving from personal computers to laptops, smart phones or integrated devices whose main function is to access particular sections of the Cloud through browsers or mostly dumb applications. While front-end processing is perhaps becoming slightly more common in the form of in-browser application, data storage is heavily biased towards centralized back-ends.

The implications are many: users are giving away their content under a false ideal of community; they are giving away their privacy for the sake of a more personalized service; they are giving away their rights in the name of comfort and accessibility; but, most

[1] Cloud computing is based on the centralization of resources. To the extent that content is centralized, control is also centralized. Regardless of who is the actual owner the data stored in the Cloud, the manner in which resources can be accessed, used, or even just transferred from one place to another is ultimately controlled by the Cloud provided. For a general overview of the specificities of Cloud Computing, see e.g. Michael Miller (2009), "Cloud Computing: Web-based applications that change the way you work and collaborate online", Que Publishing, Indianopolis, Indiana.

importantly, they are giving away their freedoms and, very frequently, they do not even realize it.

The paper will analyze the impact of Cloud Computing on society. By analyzing the way the Internet has developed over time, it will draw attention to the fact that the Internet has been and is evolving into an increasingly centralized architecture that might strongly impair the rights of end-users and endanger the privacy and confidentiality of information stored into the Cloud. These problems are exacerbated by the international character of the Cloud, which extends over multiple jurisdictions but does not account for national boundaries.

Regulating the Cloud has turned out to be an extremely challenging task, which has not yet been properly addressed by the law. With this paper, we do not purport to come up with a solution, but merely to propose a series of recommendations on how to address these challenges by public and private means.

2. The Emergence of Cloud Computing

1. – *Definition of Cloud Computing*

Given its recent and very fast adoption in everyday language, the actual definition and scope of Cloud Computing are still under debate. In part, this stems from the fact that Cloud Computing does not actually provide much in terms of new technology, but rather an alteration of the use of older technology to serve new types of business structures.

The underlying idea of Cloud Computing dates back to the 60's with the concept of "utility computing" - the dynamic provision of computing resources according to the client's needs.[2] As for the term "Cloud Computing", telecommunication operators already employed term "cloud" in the early 90's as a means to demarcate the boundaries of responsibilities between users and service providers. However, it is not until 2006 – when Amazon launched its new Amazon Web Service (AWS) – that the term "Cloud Computing" eventually became mainstream[3] and rapidly evolved into a popular business model, which, in spite of its popularity, is still difficult to define.

NIST's definition of Cloud Computing[4] is perhaps one of the most comprehensive, but is not universally accepted any more than any other definition. For the purposes of this paper, we consider Cloud Computing to represent the sharing or storage by users of their infrastructure or content on remote servers that are accessible online. This can be achieved at the level of the infrastructure (IaaS), platform (PaaS), or software (SaaS), each with their share of structural nuances and potential threats. This paper will focus on the concept of public Clouds, intended as a variety of applications that users can access and use through web browsers as if they were installed on their own computers or devices.[5] Although not all public clouds are browser-based (for example Dropbox's public shares), this focus does not

---

[2] See Douglas Parkhill (1966), *The Challenge of the Computer Utility*, Addison-Wesley; exploring the the similarities between the on-demand supply of electricity and the elastic provision of hardware and software resources.

[3] See e.g. Rachael King (2008), "Cloud Computing: Small Companies Take Flight". Businessweek.

[4] See the NIST Definition of Cloud Computing; Peter Mell and Timothy Grance, NIST; available at http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

[5] Cloud Computing can be implemented at various levels of abstractions and deployed either internally of externally. In the common sense of the term, Cloud Computing refers to the concept of a "public Cloud" as a service offered by a third-party that dynamically provides a series of resources accessible on-demand through the Internet, often via web applications. This can be contrasted to the concept of a "private Cloud" as a service for private networks allowing a company to host applications or virtual machines on its own premises.

come out of thin air, as the browser is increasingly used as a catch-all approach for user applications and is increasingly being developed with this specific intent.

Although such cloud services are generally seen as advantageous to end-users, in terms of flexibility of access and scalability of costs, these benefits come at a price. While the Internet was regarded by some early in its existence as a possible implementation of a decentralized market economy,[6] we see it moving towards a thoroughly centralized market where the power of the service providers increases as the power of end-user terminals decreases, as is apparent with netbooks and low-end laptops, mobile phones, e-book readers, embedded networked computing appliances in cars and other consumer devices. Although their relative computational capacity has increased substantially over time, heavy processing is increasingly performed in the Cloud and only the results are displayed to the users, so neither high processing power, large amounts of RAM, nor even permanent storage are nowadays required on the user-side to perform most everyday operations. A smart phone connected to the Internet can be just as powerful as any computer because it borrows storage capacity and computational resources from the thousands of machines that constitute the Cloud; any complex processing is done remotely while the front end simply deals with presentation. The technical characteristics of the terminal are no longer relevant as (a) software is for the most part executed through online servers, and (b) data no longer resides on end-user devices, but is instead stored in the Cloud.

The current trend suggests that most of the computing activity that is today performed locally on end-user computers will eventually shift into the Cloud; moving from a peer-to-peer decentralized computing environment to a centralized client-server environment. Whether or not this is desirable, from the perspective of end-users, depends on various philosophical aspects, but also technical details regarding the way the Cloud is implemented and on the policy of the Cloud provider, in particular, in terms of privacy and data protection. The problem is, however, that policy is inherently malleable. In practice, there is no privacy policy, uptime assurance or data protection mechanism that can eliminate the added operational risk created by shifting to a third party infrastructure. At best, the risk can be minimized by not storing sensitive data[7] and mitigated by not relying on one single cloud platform.

2. – *The changing face of Networked Services*

a. Trends towards centralization

The Internet was designed as a decentralized system to maximize resilience and eliminate the possibility of a single point of failure. Due to this design parameter centralized services were uncommon on the early Internet. As it became increasingly commercialized, service providers were mostly small scale companies, schools and cooperatives that utilized the distributed nature of the network. Most early websites were informational resources for

[6] During its early phases, the Internet was often regarded by many pioneers and visionaries as a potential implementation of a pure market economy characterized by free exchange of information, low transaction costs and very few barriers to entry. See, e.g. Eric Schlachter (1994), Cyberspace, the Free Market and the Free Marketplace of Ideas, in Hastings Communications and Entertainment Law Journal (Comm/Ent) [16 Hastings Comm/Ent L.J. 87]; Yannis Bakos (1998), The emerging role of electronic marketplaces on the Internet, in Communications of the ACM, Volume 41 Issue 8; James C. Bennet (2001), The End of Capitalism and the Triumph of the Market Economy, in Network Commonwealth: The Future of Nations in the Internet Era.

[7] In the context of data protection, sensitive personal data is defined to include religious beliefs, political opinions, health, sexual orientation, race, membership of past organisations, etc (see e.g. 8 of the Data Protection Directive). Here, the term is used to refer to any information that is considered (subjectively) valuable and whose dissemination might (potentially) have a negative effect on the data subject.

local communities, competing with older peer-to-peer (P2P) systems and protocols, such as e-mail and Usenet, working from a very limited set of use-cases and metaphors.

There was a strong momentum towards community-based websites and user-driven journalism in the late 1990's, with articles and feedback emerging on online web based forums, which were slowly replacing Usenet.[8] Before the advent of blogging platforms such as Wordpress.com, Livejournal.com or Blogger.com, it was not uncommon for small groups of people to set up a web server to host personal home pages, frequently running custom made software managed by somebody in the group.

Likewise, instant messaging and interactive discussions were generally done through direct communication between peers and on decentralized platforms, such as Internet Relay Chat (IRC), as opposed to centralized systems which have since emerged, such as ICQ, Microsoft Messenger or Skype.

As time has passed, small, local services have been replaced with larger, more central ones. Dmytri Kleiner notes that the dot com boom "was characterized by a rush to own infrastructure, to consolidate independent internet service providers and take control of the network."[9] He describes the situation as a kind of land grab where investors tried to replace the smaller service vendors with larger ones on every scale, from low level telecommunications infrastructure to high level services such as news aggregation, e-mail and video.

This centralization trend has further continued with increasing market consolidation, currently yielding an ecosystem comprising of services like YouTube for video, GMail for e-mail, Google News for aggregated news, Flickr for photo sharing, and MSN and Skype for instant messaging and voice/video conferencing. Many alternatives exist catering to more specific needs, scattered along the long tail of a Pareto distribution, but with a seemingly increasing scale parameter.

Network effects are such that the more users are on a platform, the more valuable the platform is to each user. In spite of their significance in the context of social networks, network effects are not, as such, a sufficient justification for there to be only one centralized social networking platform.[10] The network is fully capable of allowing for decentralized systems, as various peer-to-peer protocols have demonstrated.[11] It is possible to devise a peer-to-peer infrastructure based on an open protocol, which would allow users to keep control over their own data, and even to use network in a limited way locally on their computer, without the need for any Internet connection.[12]

---

[8] Online news started with Bruce Parello's "News Report" on the University of Illinois' PLATO system in 1974, but by the late 1990's most large newspapers had at least some online presence.

[9] Dmyti Kleiner, The Telekommunist Manifesto, Institute of Network Cultures, Amsterdam; http://www.networkcultures.org/_uploads/%233notebook_telekommunist.pdf

[10] Natural monopolies are justified by large economies of scale: a producer's cost curves decrease when the scale of production increases. Network effects describe the increase in value of a good or service derived from the standardization of that good or service. While natural monopolies often comes together with network effects, like in the case of the telephone network, network effects do not necessarily lead to natural monopolies, like in the case of the Internet network.

[11] Decentralized protocols are ubiquitous on the Internet. Giving an exhaustive list would be unpractical, but common examples include the Domain Naming System protocol (DNS), the SMTP protocol for e-mails, Bittorrent and Gnutella for file-sharing, Skype (which uses centralized coordination servers but attempts to make calls directly between peers), and, finally, FreeNET, i2p and TOR for anonymous navigation.

[12] Batch processing of downloaded e-mail is a well known example of this. This practice dates back to the the time when Internet connectivity was temporary, for example with costly dial-up connections. People would download all mails, compose replies and other mail as necessary, and then reconnect to send.

Interoperability between systems operated by different vendors is at the heart of this, but individual vendors are not legally required or financially motivated to support interoperability, and increasing concentration of the market and the consequent concentration of power in the hands of a few enterprises is preventing this from happening.

The concept of Cloud Computing then becomes not an issue of mere convenience for users, but a primary objective of vendors who wish to increase their market share. Cloud services, whether they're infrastructural, platform based, or software as a service, present a fiction of decentralization to the user in the form of network effects, while the service is increasingly operated by large companies that leverage their position to limit interoperability.

Because of their dominant position, large service providers can exert a degree of subjugation never conceived of by smaller and more local services, and a degree of control that would be impossible in a peer-to-peer network. This creates a series of legal issues in terms of control, privacy, and confidentiality of information that will be specifically addressed in the following sections.

b. Case study: two social networking sites

The case of social networks is particularly interesting given their manifest evolution from a local and community-centric to a global and extremely centralized architecture. Prior to the globalization of social networking sites such as MySpace, Facebook, and Google+, smaller scale social networking sites were common within local communities, such as hugi.is, an interest-based social network in Iceland, irc-galleria.net, a Finnish website providing social networking and photo gallery services to IRC users, and cu2.nl, a Dutch social network offering forums and photo galleries, amongst other things. Most early social networks did not manage pair-wise relationships between users. User relations were typically flat and unrestricted, with all users of the system seeing each other's profiles and general information, but they were commonly pseudonymous and contained very limited private information. Initially introduced in such systems as MySpace, Orkut and Bebo, pair-wise relationships have since then become part and parcel of any system intending to provide social networking, although symmetric relationships are not always necessarily the desired format. Twitter was the first major social network to demonstrate the value of asymmetric relations. Today, most social networking websites provide similar features and characteristics. All provide public and private messaging systems, albeit with variable levels of service and emphasis.[13] Some systems allow photographs or other media to be added, such as Facebook and MySpace in particular, which allow photo albums, videos and other rich media, sometimes including third party applications.

Accepting these variations on the theme and acknowledging the untold other differences, we will focus the remainder of this case study on two social networking sites; one local, the Icelandic site Hugi,[14] and one global, Facebook.

To begin with, it has to be noted that Hugi cannot be understood as a decentralized service. Rather, it is an early example of a centralized social networking service.

[13] While they all provide users with a way to communicate with each other, different platforms provide different means of communication. Some allow threaded messaging while others only allow linear messaging. Some restrict the number of characters allowed in messages, for example 140 on Twitter, 450 in Facebook public status updates and 10000 in OkCupid private messages, while others do not impose any such practical restrictions.

[14] Hugi was originally operated by Sí minn, the former state telecoms company which was privatized in 2005 with the sale of 98.8% of its shares to Skipti. It is now operated by Skjá miðlar ehf. For more information, see www.hugi.is

Technologically, Hugi is very similar to the early Facebook.[15] Even today, apart from the improved friendship management, the technology behind Facebook is not far removed from that of Hugi. Facebook has a more developed user interface and gives different weight to different features such as internal chat, external chat through XMPP, status messages and other aspects of messaging, but most features are primarily user experience tweaks which have come along over various iterations of the Facebook user interface.[16]

Until 2003, a large portion of Icelandic people aged from 16 to 24 were actively contributing on Hugi in polls, forums, articles and other interactive communications. Today, however, most of the user-base has shifted to Facebook. As of 2011, it is estimated that over 65% of people in Iceland have accounts on Facebook.[17]

While there are certainly many elements of user interface that influence people towards using Facebook, as the various interface changes to Facebook have shown, it is hard to believe that the trigger is merely a technical one. Rather, we claim that the key factor for the shift from Hugi to Facebook was essentially due to the more integrated and international nature of the latter, as opposed to the local character of the former.

In order to back up this claim, an online questionnaire was sent to some former users of Hugi and current users of Facebook. The results reveal that the scope of the service (i.e. its extension in the Internet landscape) weights very strongly in the mind of end-users. Despite a general inclination towards the private management of personal data, all users have declared to value the size of the community and the worldwide scope of the platform above other factors.[18]

As a result of their difference in scope, the two services are not even considered to serve the same function by many users.[19] Hugi is little more than a communal sounding board that maintains a local culture fitted to meet the needs of its original operator, Síminn, a telecommunications company. Facebook, on the other hand, is both an agora and a marketplace. Like Hugi, it is controlled by a single company, but, unlike Hugi, it has reached global significance. As a commercial start-up, the goal of Facebook is to increase the number of users on the network, as well as their dependency upon it, so as to lock a maximum number of users into the system.[20]

---

[15] Developed in the PHP programming language with MySQL as a database, Hugi does not provide much in the way of Web 2.0-style services beyond the level of user interaction presumed in such a setting; e.g. there is no post-loading processing which accesses server data, as through AJAX or other asynchronous HTTP requests.

[16] It can be expected that if Hugi had not been "neglected" similar updates would have followed there, although perhaps not with as great rapidity. In conversation with Hugis webmaster, in May 2011, it was said that, although Hugi had seen better times, a large cause of its decline was the neglect of the site's original owner.

[17] As of 2011, Iceland ranks first in terms Facebook penetration, with over 65.76% of the population on Facebook or 203,140 in total. Web developer Brian Suda noted on Twitter (http://j.mp/pVFK2N) that Facebook's internal advertising service estimated reach for advertisements targeted at the Icelandic market to be greater than the population of Iceland. For more updated statistics, see http://www.socialbakers.com/facebook-statistics/iceland .

[18] In a small and informal questionnaire (n=30) amongst former users of Hugi, when asked whether, all other things being equal, they would prefer a service such as Facebook, but with their personal data hosted within Iceland, exactly half said they would; when asked if they would prefer a service where their data was hosted on their own private computer, 64% said they would. Younger people, in particular, seem less concerned with sovereignty over their own data, while older users appear more concerned about the locality of their data. Yet, all of those questioned said that the size and international aspect of Facebook mattered either much or very much.

[19] In the same questionnaire amongst former users of Hugi who also use Facebook, 82.15% claimed that Facebook and Hugi serve different roles, with the rest claiming that they only partially serve the same role.

[20] As for 2011, Facebook is valued at roughly 80 billion dollars (according to a recent private-market transaction on SharePost, an online marketplace for private investments) and has over 500 million users;

According to current estimates, roughly 10% of the world's population has Facebook accounts,[21] giving this centralized platform a higher penetration than any system seen before. This case study shows the trend clearly in terms of social networking, but we believe the conclusions of this analysis to apply, by and large, to the majority of applications provided by large centralized companies over the Internet.

3. Legal Issues of Cloud Computing

It takes only very basic examples to show the danger of over-centralization in the sphere of the Internet. In addition to the most common examples, such as Google and Facebook, there are a very large number of actors whose operations are crucial in the everyday life of many Internet users. The more the level of dependency increases, the more the effects of not having control over the content or infrastructure become apparent, although some of the implications might remains very subtile. In this section, we will illustrate the manner in which the Cloud distinguishes itself from standard client-server architecture by virtue of its centralized character, and how this might endanger both the privacy of end-users and the confidentiality of information. Finally, we will address the issue of transnationality and data sovereignty in order to understand whether it can actually be resolved in the context of Cloud Computing.

1. – *Centralized Control*

Today, no matter how much one tries to keep it secret, there exist many mechanisms or devices that collect personal data and communicate it to third parties without the consent of the data subject.[22] Most often, however, it is actually the user who willingly communicates information to a variety of interested parties. On the Internet, this is done on a daily basis through blogs, forums, newsgroups, mailing lists, search engines, etc. It has been argued that there is no reasonable expectation of privacy on public fora, but there exists an often unacknowledged distinction between data explicitly published and metadata created as part of the publishing activity, not to mention data provided to a service for private reasons, such as search queries to a search engine, or draft blog entries. In other settings, a user may wish to grant some people access to data, such as reviewers for a draft or a comment

---

meaning that each user's contribution, if we ignore the network effect, is about $160. Of course, given the nature of network effects, the most recent user added is always the most valuable. With 7% of humanity registered on the world's largest social network, the only way for Facebook to increase shareholder value is to aggressively reach out to an ever-growing group of users, while minimizing the risk that current users leave.

[21] As of July 2011, Facebook claimed to have 750 million active users. https://www.facebook.com/press/info.php?statistics

[22] Spyware programs (which are a form of malware) are malicious software that collects personal data about users without their consent. As users perform tasks such as browsing the Internet, spyware programs collect information about users and their behavior. Although commonly acknowledged in the digital world, similar devices are commonly deployed in the physical world, in the form of eavesdropping, interception of written communications, video surveillance through CCTV, and, most recently, identification via biometric data and geo-localization by means of GPS tracking and networking technologies. For a more detailed overview of the mechanisms and the consequences of pervasive surveillance in modern societies, see, e.g. David Murakami Wood (2008), Towards Spatial Protocol: The Topologies of the Persavise Surveillance Society, in Alessandro AUrigi and Fiorella De Cindio (Eds), Augmented Urban Spaces: Articulating the Physical and Electronic City; Ashgate Publishing.

intended for family members on social media. The existence of access control lists in software creates an expectation of privacy.[23]

While this is not a problem in itself, outsourcing data, software and hardware resources to a third party's architecture necessarily requires some consideration. Security risks, privacy concerns, lack of interoperability and user's lock-in are only few of the problems that might derive from the fact that users do no longer have control over their own resources. Indeed, as many user no longer control nor understand their infrastructure, they are increasingly controlled by those who do know how to control the infrastructure - and by those who own it.

Cloud Computing introduces an additional layer of concern. Although apparently analogous to traditional server-client architectures, there persists an significant difference between the Cloud scenario and other existing outsourcing scenarios. The reason is that, in the case of Cloud Computing, huge amounts of data can be gathered together into large data-centers often interconnected to each other.

The problem arises when the information given to separate (and apparently independent) services is actually aggregated together by one single entity (either because it is the common provider of said services, or because it has acquired the data from third parties). Even though information had been voluntarily provided by users, aggregated data might provide further information about users, which they did not necessarily want to disclose. This can be critical because, if one single entity were to provide a large variety of services and the data collected through all of these services were to be processed into an integrated framework of analysis, that entity would fundamentally be able to know much more about its user-base than what has been voluntarily disclosed by each individual user.

Technically, this is already a possibility, and, as a matter of fact, this is already part of reality. Let's take a look at Google. With a mission to "organize the world's information and make it universally accessible and useful", Google offers a large variety of services (mostly for free), whose ultimate purpose is not only that of presenting information in a more organized way, but also that of gathering as much information as possible. Services such as Google Mail, Google Documents, Google Calendar, Google Maps, Google News, Google Reader, Orkut, Youtube, Picasa - and many more - are all intended to collect information about the users of that service. Even a service apparently as harmless as the Google search engine is in fact able to collect very important pieces of information. A cookie (whose expiration date is irrelevant for any practical matter) is stored into every computer so that it can be identified at every subsequent connection.[24] While the Citizen's Rights Directive of 2009 (which amended the E-Privacy Directive of 2002) now requires a system of "opt-in" for the use of cookies, explicit consent is however not necessary when the cookie is 'strictly necessary' to deliver a service which has been explicitly requested by

[23] See Matwyshyn, 2009, "Harboring Data: information security, law and the corporation", Stanford Law Books - in particular the chapter dealing with social networking.

[24] Every time a user connects to Google's search engine, a cookie is stored on the user's device, with an expiration date of two years. The expiration date is pushed ahead of two years whenever that cookie is accessed by any of Google's sites and it is detected that the cookie is about to expire. By virtue of this cookie, Google is able to store an almost permanent and unique ID on every user's device, as Google will either keep the same unique ID in the cookie, or at least be able to associate the old ID with any new ID that is issued. Although Google claims that the purpose of the cookie is to remember user preferences, the cookie is also be used for the purposes of profiling. See http://www.google.com/privacy/privacy-policy.html - "When you visit Google, we send one or more cookies to your computer or other device. We use cookies to improve the quality of our service, including for storing user preferences, improving search results and ad selection, and tracking user trends, such as how people search. Google also uses cookies in its advertising services to help advertisers and publishers serve and manage ads across the web and on Google services."

the user.[25] In the case of Google search engine's cookie, although it does potentially enable Google to collect all manner of information about users, this cookie is presented as a valuable service to the users, who would otherwise be unable to enjoy the benefits of personalized search results and customized advertisements. Most users are either unaware of the privacy implications of such services, or value the service beyond their perceived personal risk. Younger users, in particular, are less disturbed by the panopticon-like sharing of information, both on social networks and to companies providing cloud-based services, but are yet not entirely flippant about their approach to privacy - indeed, it appears that their approach may be more nuanced than that of older users.[26] Increased demand for clear privacy settings in software and understandable privacy policies appears to be slowly improving this gap in awareness.

Since most of these services are either available online or automatically synchronized whenever a user connects to the Internet, Google can keep track of every user activity performed on its system. This data can be very valuable for the purposes of mass profiling (i.e. understanding the preferences of the user-base as reflected by the behavior of each individual user) and user profiling (i.e. understanding the preferences of each individual user through the analysis of its specific interests, activities, and social surroundings).[27]

However, Google, being a corporation, is ultimately not interested in monitoring the activities of its users, nor in gathering information about the socio-demographics of its user-base, but rather in the maximization of profits. Profiling is necessary for Google to know what users want, so as to eventually offer them the most personalized results and the best kind of advertisements. The greater the user-base, the most accurate the profiling can be, and the higher the profits that can be extracted from a system of customized advertisement dependent upon the interests of each individual user. In this case, the fact that the end-users do not pay for the service means that they themselves are the product being sold, or rather, statistics about them are. There is no reason to assume malice here, but there is reason to draw attention to privacy concerns.

Various companies have built successful business models around the realization that, instead of getting money in exchange of a service, it is often more valuable to provide services for free in order attract a maximum number of users. By accepting the terms of services, users agree to share most of their data and information with Google, regardless of

---

[25] See Article 5 of Directive 2009/136/EC (the Citizen's Rights Directive): "Article 5(3) shall be replaced by the following: "3. Member States shall ensure that the storing of information, or the gaining of access to information already stored, in the terminal equipment of a subscriber or user is only allowed on condition that the subscriber or user concerned has given his or her consent, having been provided with clear and comprehensive information, in accordance with Directive 95/46/EC, inter alia, about the purposes of the processing. This shall not prevent any technical storage or access for the sole purpose of carrying out the transmission of a communication over an electronic communications network, or as strictly necessary in order for the provider of an information society service explicitly requested by the subscriber or user to provide the service.";

[26] See Danah Boyd and Eszter Hargittai, "Facebook Privacy Settings: Who Cares?", First Monday http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3086/2589

[27] Mass profiling is more concerned with the general trends and navigation patterns of the user-base than with the actual preferences and activities of each individual user. User profiling focuses instead on the personal and distinctive characteristics of users and is therefore more likely to infringe upon their right to privacy. For an overview of the various techniques used for the profiling of users in a Cloud environment, see, e.g. Olfa Nasraoui and Carlos Rojas (2003), From Static to Dynamic Web Usage Mining: Towards Scalable Profiling and Personalization with Evolutionary Computation, in Workshop on Information Technology, Rabat, Marocco, and, in particular, Gang Ren; Tune, E.; Moseley, T.; Yixin Shi; Rus, S.; Hundt, R. (2010), Google-Wide Profiling: A Continuous Profiling Infrastructure for Data Centers, in Micro, IEEE, volume 30, issue 4.

the privacy or the confidentiality thereof.[28] Hence, although the majority of Google's services are offered for free, users pay - willingly or not - with their own data, which is only later turned into profit by Google AdSense or other forms of advertisement.

In this context, the scope of the Cloud is extremely important. By offering such a wide variety of services, Google is able to obtain different pieces of information which pertain to different fields of endeavor. When users search for something on the web, Google can learn about their interests; when users read their emails on Gmail, Google can learn more about their personal or professional life; when users check out a location on Google Maps, Google can learn where each user has been or wants to go. The greater the scope of the Cloud, the greater is the amount of data that can be gathered together and the more valuable is the information that can be obtained with the processing and correlation of such data.[29]

While this is likely to help Google increase its profit, the collection and processing of user data into a common integrated framework can also benefit the users when it comes to increasing the quality of the service. Many users are therefore not merely agreeing, but even eager to share their personal data and information with Google in order to obtain a more customized and integrated service. Google Calendar is more valuable because it can be integrated with Gmail for e-mail reminders and notifications and with Orkut and Google+ for discovering new events and remembering the birthdays of some friends. As the value of a service increases not only with the number of users connected to that service but also with its degree of integration with other services, the wider is the portfolio of services offered by Google, the most users will be attracted to these services.

2. – *Privacy & Confidentiality*

There is an inherent security risk in the use of the Internet to transfer sensible information and personal data. As a general rule, information wants to be shared, and most of the value that can be extracted from it emerges from the usage and communication thereof. However, whenever it is published on the Internet, the privacy and confidentiality of information is necessarily put at risk.[30] Given the global scope and international character of the Cloud, these risks have considerably increased with the deployment of Cloud Computing. Every bit of information that has been published into the Cloud becomes accessible from anywhere and at anytime, yet, once it has been exported into the Cloud, users lose the possibility to control their data, which can no longer be accessed, edited or retrieved without the consent of the Cloud provider.

---

[28] Google privacy policy states that Google may collect all kind of personal information provided by users themselves, log in information gathered whenever users access one of the various Google's services, user communications, information gathered by cookies stored in users' devices or collected by third party applications, and location data in the case of location-enabled services such as Google Maps or Latitude. For more details on Google privacy policy, see http://www.google.com/privacy/privacy-policy.html

[29] Google's privacy policy clearly states that Google will be pooling all the information they collect from all of their services. Google reserves the right to "combine the information you submit under your account with information from other Google services or third parties in order to provide you with a better experience and to improve the quality of our services." See http://www.google.com/privacy/privacy-policy.html

[30] The advent of Internet and digital technologies introduced a series of concerns that might significantly affect users' willingness to communicate personal data and confidential information over the Internet. Given that there can be no perfectly secure mechanism to transfer information, publishing information on the web necessarily involves the risk of data loss or spill over. See e.g. Bob Blakley, Ellen McDermott, Dan Geer (2001), Information security is information risk management, in Proceedings of the 2001 workshop on New security paradigms, New York; and Eric C. Turner; Subhasish Dasgupta (2003), Privacy on the Web: an Examination of User Concerns, Technology, and Implications for Business Organizations and Individuals, in Information Systems Management, Volume 20, Issue 1.

The advent of Cloud Computing has introduced a series of new challenges concerning the way in which information can be transferred or processed,[31] most of which have yet to be resolved. This requires more careful attention to be paid to the actual or potential consequences of Cloud Computing on the privacy and confidentiality of personal and governmental information. What kind of information can be shared into the Cloud? Can anything be kept private in the Cloud? How to make sure that data protection regulations are actually being respected by every player involved in the provision of a Cloud-based service?[32]

On the one hand, as an attempt to reduce the risks of abuse by third parties, the law may restrict the ability of certain institutions to rely upon the services of a Cloud provider by introducing a series of procedural and/or substantive barriers. The reason is that information stored in the infrastructure of a third party may have weaker protection than information that remains in possession of users. The chances for inadvertent exposure increase substantially with every new intermediary and with every new layer of abstraction. While securing the infrastructure is obviously very important, it is not sufficient if the interface or application running on that infrastructure has not been properly secured as well. Although users need a way to log into the system in order to transfer data from or into the Cloud, this could constitute a significant security risk unless proper access control and secure transfer protocols have been adopted. Likewise, even though users are made to access the services by password, unless there is filesystem level encryption of the data with a key held only by the user - which is impractical in most cases - the operator of the service or anybody else who gains physical access to the servers can peer into the stored data. In more extreme cases, attacks on the hardware can be used to extract information that is resident in runtime memory.[33] In most cases, security issues are due to lack of or poor application of cryptography and a general lack of tradition for security. Various campaigns have tried to remedy this, such as the Tactical Technology Collective's ONO Robot campaign, Survival in the Digital Age,[34] and the Electronic Frontier Foundation's HTTPS-everywhere campaign.[35] Yet, regardless of the degree of protection promised by the cloud provider, the security and confidentiality of information is ultimately determined by the weakest link in the chain. Insofar as data is transferred through several intermediaries, only one of them needs to be violated for any malicious user to obtain the relevant information.

---

[31] See Gutwirth (ed.) (et al) Computers, privacy and data protection: an element of choice, Springer, 2011

[32] A lot of discussion centres on the privacy and confidentiality issues surrounding the cloud, for an overview of the current debate, see e.g. Hon, W. Kuan, Millard, Christopher and Walden, Ian, The Problem of 'Personal Data' in Cloud Computing - What Information is Regulated? The Cloud of Unknowing, Part 1 (March 10, 2011). Queen Mary School of Law Legal Studies Research Paper No. 75/2011. Hon, W. Kuan, Millard, Christopher and Walden, Ian, Who is Responsible for 'Personal Data' in Cloud Computing? The Cloud of Unknowing, Part 2 (March 21, 2011). Queen Mary School of Law Legal Studies Research Paper No. 77/2011.

[33] An interesting example is the Cold boot attack, allowing anyone with physical access to a computer to retrieve encryption keys from the operating system after restarting the machine. The attack relies on the "data remanence" of DRAM and SRAM memory in order to retrieve memory contents that remain readable for a short period after power has been removed. For more information, see J.Alex Halderman, Seth D. Schoen, Nadia Heninger, William Clarkson, William Paul, Joseph A. Calandrino, Ariel J. Feldman, Jacob Appelbaum, Edward W. Felten (2008): Lest we remember: Cold Boot Attacks on Encryption Keys, in Proceedings 2008 USENIX Security Symposium.

34 The Tactical Technology Collective and ONO Robot produced a series of animated films to raise awareness about the digital traces users leave behind. Its main aim is to engage people in better understanding the information and communications technologies they are using, so that they can decide when and if they want to take risks. For more details, see www.onorobot.org

[35] HTTPS Everywhere is a Firefox extension produced as a collaboration between The Tor Project and the Electronic Frontier Foundation. It encrypts communications with a number of major websites using Transport Layer Security. For more details, see http://www.eff.org/https-everywhere

Accordingly, a distinction must be made between the use of Cloud Computing for storing or processing data on the parts of consumers, end-users or small companies, as opposed to government and multinationals. While individuals are generally free to share information in a decentralized global environment (even though they are often not fully aware of the terms set out by the service providers and of the consequences of storing information in the Cloud),[36] in the case of an institution - such as a business, corporate, or governmental institution - privacy laws often prohibit or limit the disclosure of personal information to third parties.

The disclosure of information by government agencies is restricted both by internal rules and public regulations on data protection, whereas a series of standards established by different bodies of law regulates the possibility for a business or corporation to export information into the Cloud. For instance, in the USA, the Health Insurance Portability and Accountability Act (HIPAA) establishes a series of rules regulating the use and disclosure of identifiable health information, which can only be transferred to a service provider that promises to comply with the same set of standards (often incompatible with the terms of services established by a cloud provider). Similarly, the Violence Against Women Act precludes domestic violence service providers from disclosing information without the consent of the data subject, unless compelled by statute or a court (Public Law 109-162 as amended by Public Law 109-271); tax preparation laws provide statutory and regulatory protection that limits the disclosure of tax return information without the taxpayer's consent (Internal Revenue Service rules - 26 U.S.C. § 6713 and § 7216; 26 C.F.R. §301.7216); whereas the disclosure of personal information concerning the financial situation of a consumers by a financial institution is precluded under the Gramm-Leach-Bliley Act (15 U.S.C. § 6802); and the disclosure of video rental and cable television subscribed records is protected under the Video Privacy Protection Act (18 U.S.C. § 2710) and the Cable Communications Policy Act (47 U.S.C. § 551). Similar rules apply in Europe and in a variety of other jurisdictions. Although the actual content of the law varies according to the jurisdiction, a certain degree of harmonization has nonetheless been achieved in various parts of the world. In Europe, for instance, the European Data Protection Directive heavily regulates the processing, transfer and disclosure of personal information.37 The Directive concerns the processing of "personal data" - broadly defined as "any information relating to an identified or identifiable natural person".[38] It establishes a series of basic conditions that must be fulfilled with regard to personally identifiable information, which can only be collected and processed to the extent necessary as to fulfill a particular purpose, as well as

[36] While many users do not even bother to familiarize themselves with the terms of services of the cloud computing platform they wish to use, doing so is often not an easy undertaking even for those who try to understand the consequences of entering into such agreement. Besides, it is fairly common that the provider reserves the right to vary the terms and conditions on which the service is provided without notifying the users. For more details, see Dan Svantesson, Roger Clark (2010), Privacy and consumer risks in cloud computing, in Computer Law & Security Review, 26 (4), 391-397.

[37] For more information, see http://ec.europa.eu/justice/policies/privacy/index_en.htm. Any entity holding or processing personal data must comply with a set principles of good practice, according to which data must be fairly and lawfully processed, for limited purposes and in an adequate, relevant and not excessive manner. It must remain accurate, be securely kept no longer than necessary and it must be processed in accordance with the data subject's rights.

[38] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, Article 2(a).

an additional set of restrictions on the collection and use of sensitive data (i.e. personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership and data concerning health or sex life),[39] which can only be processed with the explicit consent of the data subject. The Directive also introduces an obligation for anyone processing personal data to notify the data protection supervisory authority of the member State in which they operate, and to provide proper information to the individuals whose personal data is being processed. Data subjects always have the right to refuse that their personal data be used for advertising or marketing purposes. Finally, the Directive provides that the transfer of personal information outside of the EU can only be done if the laws of that country provide an adequate level of protection[40] (unless the company to which the data is transferred actually guarantee to comply with European data protection laws).[41] Those provisions have been discussed in the ECJ's case of *Lindqvist*,[42] which clarified the application the Directive to the uploading of personal data on Internet websites. Although it was held that posting personal data (e.g. individual names, telephone numbers, hobbies, etc) on the Internet qualifies as the processing of personal data for the purposes of the Data Protection Directive, the court held that the mere posting of such data on an Internet website could not be regarded as a transfer to third countries, provided that the server infrastructure is actually located within the EU. While this is likely to exempt most European web operators from the legal regime regulating the transfer of personal data, European laws limiting cross-border data transfers might however have a considerable impact on Cloud Computing, whose scope is likely to extend beyond national boundaries. Indeed, the law effectively prohibits exporting personal data to any cloud provider whose servers are located in countries with weak data protection laws.

On the other hand, certain jurisdictions have actually introduced legislation that might ultimately hinder the privacy and confidentiality of information for the sake of protecting national security and public order. This is the case of certain countries whose laws can oblige Cloud providers to communicate to the authorities any information that constitutes evidence of criminal activities. This means that government agencies can, under certain circumstances, require the disclosure of personal or confidential information by third parties. For instance, in the USA, although the Electronic Communications Privacy Act (ECPA) provides a series of protections against the access by governmental agencies to

---

[39] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, Article 8(1).

[40] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, Article 25. Specific agreements have been made with the U.S. in order to semplify the procedure for any US company that certifies to comply with the Safe Harbor Privacy Principles - a set of 7 principles that establish the minimum standards to be respected in terms of access, security, data integrity, notice, and opt-in or opt-out choices.

[41] The U.S. Department of Commerce in consultation with the European Commission developed a "Safe Harbor" framework (whose principles can be seen at http://export.gov/safeharbor/) in order to bridge the different privacy approaches adopted by Europe and the U.S. and provide a streamlined means for U.S. organizations to comply with the European Directive on Data Protection.

[42] Criminal Proceedings against Bodil Lindvist ("ECJ Case C-101/01"): a Swedish woman posted information concerning her volunteer work at a church on her website, including identifiable information concerning her colleagues. The Swedish data protection authorities commenced proceedings against her for having posted such information without obtaining permission from the Swedish data protection authorities and the individuals concerned. The Gota Court of Appeal referred a number of questions to the ECJ to clarify the interpretation of the Data Protection Directive.

personal information held by third parties (18 U.S.C. § 2510-2522 and § 2701-2712), these protections have been subsequently weakened by the USA PATRIOT Act, which entitles the FBI to compel - following a court order - the disclosure by U.S. Internet service providers of any record stored on their servers (50 U.S.C. § 1862). The consequence is that, as opposed to personal information that remains in possession of the data subject, data published in the Cloud is more likely to be handed out to a governmental body, because, in the absence of proper notice, the data subject does not even have the opportunity to object.

While most individuals, businesses, corporate and governmental institutions do store their data online on databases and remote file systems operated by third parties, many are nevertheless reticent to export personal data and confidential information into the Cloud, because they are concerned that it might end up in wrong hands (e.g. advertisers or malicious users) or that it might be seized by a foreign governmental body.

3. – *Transnationality*

The international character of the Cloud introduces an additional layer of complexity to an already complex problem. Information stored in the Cloud can be subject to a variety of different laws according to the location where it is stored, processed or transmitted. In order to provide a service to end-users, Cloud providers might avail themselves of the services of different Cloud providers located in different jurisdictions. In addition, regardless of whether or not the service is being partially outsourced, data is frequently transferred from one data center to another in order to be processed across multiple jurisdictions. This is generally done on the basis of technical constraints and on the grounds of network efficiency, but also depending on legal or economic factors (e.g. taxation, hardware cost or price of electricity). As a result, it is often difficult to determine in advance and with certainty the actual location of information stored in the Cloud: a file being served from Luxembourg at one moment could be served from the Philippines at the next. Each jurisdiction may have pros and cons in terms of legal environment, such as different approaches to intermediary liability limitations, in the US provided under §230 of the Communications Decency Act, and in the European Union under the e-Commerce Directive (2000/31/EC), but equivalent legislation does not exist in the majority of the developing world. The varying jurisdictions may also raise questions of consumer protection, for instance in terms of warranty and merchantability. It is unclear whether a user whose data is stored in another continent, however temporarily, can expect protection from system faults, loss of data, or leaking of private data, although in most cases legal claims would be made to the hosting provider, who most likely is operating out of a country with similar jurisdictional constraints as the user.

The huge amount of data stored outside of national boundaries has become a critical issue that is directly related to the problem of effective jurisdiction - i.e. the question of government control over domestic data. While government control can be exerted over information stored within the national jurisdiction of a country, it can be extremely difficult to practically enforce after the data has been exported into the Cloud.

The reason is that it is almost impossible to provide a definition of what constitute “domestic data”. Data, as such, does not have any nationality but merely inherits the law of the territory in which it is located. It has become increasingly common in recent years that actors intentionally push data through multiple jurisdictions in the hopes of accumulating different types of protections. In this regard, a crucial problem that emerges from the international character of the Cloud is the issue of forum-shopping. Different servers and data-centers located around the world can be used to take advantage of certain laws and/or to circumvent others. Unless it has been contractually precluded to do so, a Cloud provider with data-centers in more than one jurisdiction could theoretically move information from

one jurisdiction to another in order to benefit from the most favorable laws. This can be used, for instance, as a means for any service provider to bypass domestic regulations on data protection.

On the flip-side of this, as data passes through jurisdictions it can also accumulate the weaknesses of those jurisdictions. If it is, for instance, much easier or unnecessary for police in a particular country to obtain a court order to inspect private data from Internet hosting providers than is otherwise practiced, or if there are no data protection laws in place with appropriate penalties for exposure of private data, that weakness could lead to a weakening of the overall privacy of the user's data anywhere. While this does mean that, in theory, police may be able to enforce law over national boundaries, in practice, actual police collaboration mechanisms are usually not sufficient. Europol and Interpol, for instance, can only operate within certain boundaries and have a number of resource constraints, while national police generally does not have sufficient leeway, resources or contacts to coordinate with law enforcement in other countries, especially if those countries are very far away.

In a context designed not to take into account national boundaries and where everything can travel from one place to the other in a completely transparent manner, the real challenge is to determine who can exert control over what. Given the international scope of the Cloud, identifying the applicable law for every piece of information stored in the Cloud is a challenging task because it is difficult to establish the jurisdiction that every bit of information actually belongs to. Moreover, given that data can transfer from one Cloud to another and from one jurisdiction to the other, different laws might apply to the same bits of information at different moments in time. Some services, such as Twitter,[43] explicitly state in their terms of service that activity on their service falls under a particular jurisdiction, so as to reduce their own risk exposure. The US ninth circuit is particularly common in this sense, due to the dominance of Silicon Valley over cloud services.

Regardless of its origin or destination, the same data will not be subject to the same legal regime according to the country where it is located and the nationality of the Cloud provider. For instance, according to the USA PATRIOT Act, the government could potentially seize any piece of information stored in a US data-center or by a US company, without the data subject even being aware of it (50 U.S.C. § 1862). Other countries, however, do not necessarily share the same rule. By giving out information to the US government, a foreign company could therefore potentially violate the laws of its own country to the extent that it discloses personal or confidential information to third parties without the consent of the data subject.

An increasing number of companies and governmental agencies located outside the U.S. are becoming reluctant to release their data into the Cloud - as they are concerned about their data falling within the hands of US providers, or even just entering into US territory, where it would become subject to laws allowing for the US government to access that data.[44]

The European Union addressed this issue in 1995 with the Data Protection Directive, which stipulates that personal data cannot be transferred to countries outside the

[43] http://twitter.com/tos ; "All claims, legal proceedings or litigation arising in connection with the Services will be brought solely in San Francisco County, California, and you consent to the jurisdiction of and venue in such courts and waive any objection as to inconvenient forum."

[44] For instance, even though the BlackBerry system has been accredited by security agencies in the United States, Australia, New Zealand, Austria and Canada, the French General Secretariat for National Defense has released a circular stating that BlackBerry handhelds should not be used by ministries and State officials because they constitute a threat to France state secrets. The reason is that all e-mails sent from a BlackBerry handheld are transferred through servers in Canada, Britain, and in the United States - which makes them vulnerable of being seized by U.S. authorities.

EU that do not provide an "adequate level of protection".[45] This was enacted not in response to Cloud Computing, but rather due to a general concern that data should not be transferred to non-EU countries without some adequate controls (e.g. Binding Corporate Rules; Commission's finding of adequacy, etc). In implementing this Directive, certain countries, such as Germany, introduced even stricter requirements that must be satisfied in order to comply with German data privacy law.[46] Although the Directive was passed before the widespread deployment of the Internet (and is therefore slightly obsolete nowadays), it nonetheless has strong implications for Internet service providers and Cloud Computing. Indeed, according to the Directive, national data protection laws apply to all information located in the territory of a Member State, regardless of its origin or destination.[47] The result is that, while this is likely to reduce the risk of personal data being illegitimately exploited without the consent of the data subject, this is also likely to reduce the possibilities for Cloud providers to outsource their services in the EU - because, even if data is merely being processed in a Member State, it might be difficult to export it after it has entered the EU.

4. Data sovereignty

Finally, data sovereignty is an important problem, which is often not sufficiently taken into account. In view of the advantages that can be derived from Cloud Computing in terms of costs and flexibility, many private and public institutions are tempted to export both their data and IT systems into the Cloud. Yet, many of them might be discouraged to do so to the extent that they cannot ensure a minimum standard of sovereignty over their own data. The difficulty to know with certainty which law applies to information stored into the Cloud creates strong legal uncertainty and raises a number of challenges that still have to be addressed by the law.

---

[45] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, Article 25 introduces the principles that Member States shall provide that the transfer to a third country of personal data which are undergoing processing or are intended for processing after transfer may take place only if, without prejudice to compliance with the national provisions adopted pursuant to the other provisions of this Directive, the third country in question ensures an adequate level of protection; where the adequacy of the level of protection afforded by a third country shall be assessed in the light of all the circumstances surrounding a data transfer operation or set of data transfer operations; particular consideration shall be given to the nature of the data, the purpose and duration of the proposed processing operation or operations, the country of origin and country of final destination, the rules of law, both general and sectoral, in force in the third country in question and the professional rules and security measures which are complied with in that country.

[46] See section 11 of the German Data Protection Act (Bundesdatenschutzgesetz - BDSG) that specifically addresses the requirements that German data controllers must comply with when transferring data to a third party abroad.

[47] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, Article 4 (National law applicable) specifically states that each Member State shall apply the national provisions it adopts pursuant to this Directive to the processing of personal data where: (a) the processing is carried out in the context of the activities of an establishment of the controller on the territory of the Member State;; (b) the controller is not established on the Member State's territory, but in a place where its national law applies by virtue of international public law; (c) the controller is not established on Community territory and, for purposes of processing personal data makes use of equipment, automated or otherwise, situated on the territory of the said Member State, unless such equipment is used only for purposes of transit through the territory of the Community.

Transnationality is an important aspect of this, but even within a jurisdiction the ability of the owner of data to exert authority is limited when it is held by a third party. Secret subpoenas can lead to governments gaining access to hosted data from cloud services without the owner of the data being notified, as is thought to have happened in the case of Internet activists Birgitta Jónsdóttir, Jacob Appelbaum and Rop Gonggrijp, after Twitter successfully petitioned to unseal a court order demanding their data.[48] It has not become clear whether other cloud services were served similar orders in that case.

Similarly, a third party could potentially gain access to a cloud user's data without the owner having any ability to detect such activity. While such breaches could equally come from outside or from inside the cloud, cloud providers have an ethical, but not a legal obligation to inform users of such breaches.

The new security breach notification requirements under the amended E-Privacy Directive introduce an obligation to notify concerned individuals of any security breaches involving personal data (Article 4). Those requirements are meant to increase the accountability of data holders, encourage more investments in data security, and provide an opportunity for all affected individuals to mitigate their damages. However, while this is likely to reduce the risks associated with security breaches and tampering with personal data, the Directive only applies to public communication service providers (e.g. telecom operators, mobile phone communication service providers, Internet access providers), whereas private and corporate networks have been explicitly excluded from the scope of the Directive.[49]

In the case of most Cloud services, there is thus no obligation for the Cloud provider to report any security breaches. Issues concerning data sovereignty can ultimately only be resolved by storing personal data only on private devices, and using public clouds only for public data. This includes distinguishing between sensitive and non-sensitive database entries and files, but further implies a transition back to a peer-to-peer Internet topology in terms of service rendition. Overall, the added risks, both legal and practical, suggest that users need to actively seek ways to protect their own interests.

4. Recommendations

a. Private measures and legislative limitations

In spite of its dangers and drawbacks, Cloud Computing is being adopted by an increasing large number of institutions, businesses and individuals. As the number of users increase, the infrastructure of the Cloud needs to be continuously expanding. Data centers are rapidly evolving to meet an exponential growth in the number of users and the increasing amount of data they produce. As new needs arise, the underlying technologies making up the Cloud also need to evolve in order to satisfy specific users' needs, criteria or expectations. In view of its complex and integrated nature, the Cloud relies on a variety of different technologies

[48] A large number of articles have covered this case, examples include, amongst others: https://www.computerworld.com/s/article/9204138/U.S._subpoenas_Twitter_for_Wikileaks_info and http://www.theregister.co.uk/2011/01/08/feds_subpoena_twitter/. Wired in particular published an article calling for Twitter's response to the subpoena to be adopted as an "industry standard": http://www.wired.com/threatlevel/2011/01/twitter/

[49] Directive 2009/136/EC amending Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector (the E-Privacy Directive) - recital 55: "In line with the objectives of the regulatory framework for electronic communications networks and services and with the principles of proportionality and subsidiarity, and for the purposes of legal certainty and efficiency for European businesses and national regulatory authorities alike, Direc-tive 2002/58/EC (Directive on privacy and electronic com-munications) focuses on public electronic communications networks and services, and does not apply to closed user groups and corporate networks."

designed - designed for different purposes and fields of applications - whose core functionalities are constantly expanding.

The high speed at which the technologies underlying the Cloud are evolving is such that Cloud providers are devising new mechanisms to regulate the way in which these technologies can interact with the rights and expectations of their clients, usually by means of specific Service Level Agreements. However, since most commercial Cloud providers are more interested in making profits than in protecting the interests of their user-base, users should be wary of their privacy online and understand the risks involved with losing control over the data stored in the Cloud.

In particular, given the degree of legal uncertainty that is emerging in the Internet landscape, there is a real need for the law to be reformed in order to better accommodate current and future users concerns in terms of data security and privacy. Yet, the law does not seem able to follow the pace at which Cloud Computing is evolving. Eben Moglen points out that Cloud Computing can never truly be regulated, as any regulation of the Cloud will be preempted by a change in the way the Cloud is defined, or in which jurisdiction it operates. "The cloud means that we can't even point in the direction of the server anymore" he states, adding that "You can make a rule about logs or data flow or preservation or control or access or disclosure but your laws are human laws and they occupy particular territory and the server is in the cloud and that means the server is always one step ahead of any rule you make."[50]

The legal framework is unable to deal with the flexible and dynamic character of the Cloud. The length of the legislative process cannot compete with the speed at which private actors can identify and rapidly implement technical or contractual mechanisms to avoid the constraints formerly introduced by the law.

b. Intermediary liability and responsibilities

SLAs traditionally contain wide disclaimers of liability that serve to protect the service vendor. The dynamic character of the Cloud is such that any service provider could decide at any given time to out-source part of its infrastructure and operations to third-party providers, without ultimately informing the other parties to the contract. Although the operation is generally not visible to end-users, it might nonetheless affect the quality and reliability of the service as a whole. In order to preclude any responsibility in the eventuality of failure, most of the services provided to end-users are offered under specific Service Level Agreements that stipulate that the service provider cannot be held responsible or liable for the activities performed by third-party contractors. While these can be justified for business reasons, they should stand out as a warning for end users to avoid these services even though they do not currently realize the dangers they entail.

In addition, service providers should avoid selling black-box services or using them as part of their infrastructure, since this limit the degree to which they can make guarantees to their clients. Guarantees effectively enforced by an upstream infrastructure provider necessarily rely on the ability or willingness of that provider to comply with the provisions of the SLA. This creates unavoidable problems. For example, after the service Reddit moved its operations entirely to Amazon cloud services, infrastructural problems with the Amazon cloud has, on numerous occasions, caused service outages for Reddit. Further, since these services are hosted on Amazon's infrastructure, Reddit's privacy policy can only be enforced up to the level where Amazon steps in. Outsourcing entails the transfer of responsibility from an organization to another. This creates a potential risk for end-users

[50] Eben Moglen, Freedom in the Cloud; speech given at New York Internet Society. Transcript: https://www.softwarefreedom.org/events/2010/isoc-ny/FreedomInTheCloud-transcript.html

who have entered in a contractual relationship only with the last actor in the supply chain (the Cloud vendor). Users are thus left without direct recourse against the other actors involved in the actual provision of the service, which are not necessarily informed of the terms and conditions of the end-user agreement.

c. Privacy enhancing technologies and data protection

As more and more services carry heavy privacy and confidentiality burdens, the potential threats to privacy increase. To begin with, Service Level Agreements could be developed to better reflect the privacy and confidentiality concerns of users and smaller vendors.

Yet, service Level Agreements and privacy policies are useless in the face of events which are irrevocable, such as the exposure of private data. Users of Sony's PlayStation network know all too well that this danger is not a hypothetical one.[51] Most service vendors have done little or nothing to protect the security of their users. Firesheep (a tool which enabled users to easily hijack sessions from other users on the same wireless network) showed that Facebook's unwillingness to provide HTTPS was providing a privacy risk as well as a risk of identity theft.[52] Facebook's response was to add an optional HTTPS browsing feature, which most users have never taken notice of.

A strong step towards data protection and user security could be made if service vendors were to start offering privacy-by-design by default. This would include HTTPS-only browsing, communication with clients offered over PGP or other e-mail encryption, and by promoting client awareness about data protection and privacy issues. It has been frequently pointed out that the general public is not highly concerned with the technical complexities of privacy and security[53], but this lack of awareness can be addressed on many angles. First, the development of intuitive user interface motifs for data security. Current design motifs are targeted at technical audiences - indeed, the development web browser security features over the last decade has run across multiple failed motifs, and still many users are unsure of appropriate security methods. Peter Eckersley of the Electronic Frontier Foundation has proposed an alternative scheme of sovereign keys[54], but an acceptable solution to both user interface issues and appropriately intuitive security technologies is probably still far away. In the meantime, user education and public awareness projects could go a long way towards increasing security on the user end.

The problem is that the risk of private data being illegitimately accessed or stolen cannot be resolved exclusively at the service end. Asides from the implementation of stronger security mechanisms, it would be ineffective to protect users' data by providing encryption at level of the service, since the key would ultimately be stored in the same place as the lock.

The risks derived from losing control over the infrastructure can be mitigated in different ways. One way consists of using Cloud-level server virtualization but insisting on the use of on-disk encryption with remote key management, or other privacy enhancing

[51] In April 2011, Sony suffered a breach in the Playstation online video game network. As one of the largest Internet security break-ins, this breach led to the theft of personal data, such as names, addresses, birth dates, passwords and possibly credit card numbers belonging to 77 million user accounts. This required Sony to shut down the network, and although Sony given notice of the breach to its customers, no information has been provided as to how the data might have been compromised.

[52] HTTP session hijacking (sometimes called "sidejacking") is when an attacker gets a hold of a user's cookie, allowing them to do anything the user can do on a particular website. On an open wireless network, cookies are basically shouted through the air, making these attacks extremely easy. For more details, see http://codebutler.com/firesheep.

[53] See, for instance, Whitten & Tygar, *Why Johnny Can't Encrypt: A Usability Study of PGP 5.0*, available at http://gaudior.net/alma/johnny.pdf

[54] See https://www.eff.org/deeplinks/2011/10/how-secure-https-today

methods. Another way to mitigate those risks is to abstract storage and computational capacity in such a way that data can be hosted securely on a remote host with specific access keys that are only available to one user (so that processing can be done arbitrarily at any given time by only that user). Essentially, this amounts to formalizing the Cloud not as a service to dumb client devices but as extensions of smart client devices. These clients can in turn become dynamic servers, controllers of their own data.

Various arguments have been made about the complexity of strong encryption and privacy technologies and how average users have little interest or ability to apply them. However, this claim has been taken at face value with remarkably little scrutiny. Conversely, smaller networks catering to more local communities could distribute the risk and limit the scope of potential damage.

d. Peer-to-peer alternatives, interoperability and network neutrality

The original design of the Internet was optimized for efficiency, flexibility and autonomy, as opposed to hierarchy and authority. Although large vendors can strongly benefit from Cloud-based services, smaller vendors would appear to benefit more from open protocols, federated or peer-to-peer (P2P) services, and a higher degree of interoperability.

Network effects require a service to reach a critical mass of users before the public can perceive it as valuable. Slow adoption of a new service can cause it to fail, even if technically superior to existing alternatives. In the case of many Cloud services, network effects can constitute an important barrier to entry. However, this is only true when services are not interoperable with each other - which further encourages Cloud providers to use closed and proprietary systems in order to reduce the risk of new entrants invading the market.

Because of the barriers to entry introduced by the dominant service providers, the only way for a new service to enter the market is to be far superior than whatever is currently available in the market - in terms of service, speed, and reliability - so as to provide sufficient incentives for users to move from one system to the other in spite of the shifting costs. In a poorly competitive environment composed of a few large commercial organizations using closed or proprietary formats, it is virtually impossible for any new entrant to compete without huge investments in technical infrastructure, application software and advertisement.

One of the only ways to compete with the dominant players in the market is for a very large number of (very) small players to gather their efforts together into the creation of one large integrated infrastructure.

While this can theoretically be achieved in many different ways, P2P technology is definitely the most appealing alternative for end-users. Often unable to fend for themselves due to lack of resources or lack of technical expertise, users have sought out service providers to get important Internet services. Yet, end users are the ones who can benefit the most from P2P services since they can acquire greater control over their personal data, in addition to obtaining greater vendor mobility (i.e. the ability to choose which vendor they wish to deal with). This could reduce costs for the users and generate more competition in the market. Effective use of P2P services could also guarantee that end users maintain the "right to oblivion" by making it possible for them to remove their personal data from the Cloud at any time (even though anything that has been intentionally copied by a third party could still be made available to the public).

The emergence of P2P alternatives to centralized services has encouraged some of the dominant players to introduce new barriers to entry. If consumer lock-in is no longer sufficient to eliminate competition, the solution is to attack the infrastructure of the Internet, by acquiring priority access to the network. That way, it becomes impossible for others to

compete on equal grounds, because regardless of the quality of the service, it will always be slower, and therefore less valuable.

In order to preserve competition in the market, net neutrality should therefore be respected. This can be achieved either by regulating the extent to which private parties can operate ex-ante (e.g. by introducing an obligation of non-discrimination), or by regulating the market ex-post with the tools that are already available under competition law.

5. Conclusion

Cloud computing is a new model of computing fueled by the shift of control from end-users towards increasingly centralized services providers. There are many consequences to the deployment of cloud computing: some intended, others unintentional; some good, and others bad. Many are already noticeable and measurable, while others can only be foreseen by analyzing the trends that have been set.

The advantages offered by Cloud computing are clear: infrastructure providers can benefit from strong economies of scale, whereas Internet service providers can benefit from enhanced flexibility and scalability of costs. From the perspective of end-users, the main advantages are the possibility to access data from anywhere and at any time - regardless of the device they are connected from - and the ability of avail themselves of the computing power and storage capacity of the cloud. Further, it allows clients to outsource the obligation of maintaining complicated infrastructure and having to maintain up-to-date technical knowledge, while externalizing the cost of purchasing and running the infrastructure.

This does not, however, come without costs. Exporting data to the cloud means that users can no longer exercise any kind of control over the use and the exploitation of data. Data stored in various data centers can be processed without the knowledge of users, to be further redistributed to third parties without their consent. If everything has been stored in the cloud, cloud providers can ultimately determine everything that users can or cannot do. As most Internet users are no longer in charge of their own data and are no longer capable of managing their own infrastructures of production, storage, and distribution, the control is all in the hand of few corporate entrepreneurs.

After the industrial revolution governments were urged to exercise their authority for the creation of labor and consumer protection laws, and are today faced with a similar situation as regards to the digital revolution. The claim that governmental intervention has become necessary in order to promote civil liberties and to protect fundamental rights on the Internet is not unfounded. At this point in time, however, the power dynamic is not yet so set in stone that structural changes cannot remedy the problems providers and users are faced with. P2P technologies and protocols, open standards with good interoperability mechanisms, strong encryption made widely available to users, better service level agreements and policies amongst cloud providers, greater awareness of privacy and data protection issues amongst users are amongst the methods which can be employed to reduce the risks inherent in Cloud Computing, and return the Internet back to its distributed origins, lest it rain.