Neutrality in internet regulation: three regulatory principles

Pascal Francq

24nd October 2014

Abstract

The concept of network neutrality, although disputed, is generally conceived as the need to ensure a fair use of the internet to all stakeholders. Its most widely defended interpretation postulates that putting hard constraints on how network operators control the data transferred is sufficient for an effective implementation of the concept. Except for security and technical management issues, network operators cannot discriminate against data based on their transmitter, receiver, content or a combination of these criteria. This essay argues that such a regulatory framework is not sufficient for a true democratic development of the internet. Through different examples, it will show that multiple biases exist in today's internet. It therefore defends an extension of the regulatory scope to other internet stakeholders and the adoption of new regulation mechanisms and new regulators. It proposes three regulation principles and discusses how regulation mechanisms in keeping with these principles can increase online neutrality and meet some of the current challenges of the internet.
This essay does not claim to answer all the ongoing questions, nor does it claim that the different regulation mechanisms proposed are easy to implement. It must be seen as a contribution to the continuous debate regarding the balance which must be established between the freedom of all internet stakeholders and the regulations required to ensure a fair use of the internet for all.

Pascal Francq is with the Paul Otlet Institute (Belgium) and with the ICTEAM department at the Université Catholique de Louvain (Belgium). Contact him at pascal.francq@otlet-institute.org.

Table of Contents

1 Introduction

The internet plays a crucial role in today's global economy as it constitutes a major component in the daily lives of a growing number of people. There is no doubt that the internet extends considerably the way people may communicate with each other, and that it has revolutionised the production of and the access to knowledge. The internet is also a growing factor in our modern economies and generates a flow of technological innovations.

How to ensure equal access conditions for all stakeholders (internet users and economic stakeholders) is one of the main questions raised in this context. The concept of network neutrality has been presented as an answer. For the *Oxford English Dictionary*, neutrality is a "*state or condition of not being on any side*". This implies not supporting a particular viewpoint rather than another. Applied to the online context, it suggests that the internet should act as a neutral moderator for online contents and services, and that it does not provide an environment that favours some stakeholders over others. However, in practice, the concept covers different regulatory scopes.

The most accepted interpretation is limited to the stakeholders responsible for the data transmission by forcing them not to discriminate against data based on their transmitter, receiver, content or a combination of these criteria [110]. They are only allowed to make technical decisions related to the management of the internet. A few exceptions are accepted, such as protecting network integrity or blocking unsolicited contents (mostly spam and malware). This approach, which emerges from the opposition between network operators and providers of online contents and services, defends a purely liberal view of regulation: make online contents and services compete, and let the "invisible" hand of the market decide which ones survive. However, several scholars have pointed to the fact that this traditional interpretation does not cover all the problems related to online neutrality [77].

This essay argues that the current regulatory framework does not meet the challenges of today's internet, and that an extended one must be developed. First, section 2 presents a global view of the internet and its stakeholders. Section 3 will then show — through an analysis of some technological, social and political aspects of the internet — that true network neutrality is impossible to ensure within the current regulatory settings. In fact, considering that most internet users can independently choose the online contents and services that are best for them, this is a theoretical perspective. Section 4 therefore proposes a broader scope for internet governance based on three generic principles. Each of these principles will be illustrated through current topics regarding internet regulation. It also discusses the form that internet regulation authorities responsible for the implementation of this new regulatory framework and its verification should take. In particular, it defends the establishment of independent regulatory agencies that represent all internet stakeholders in a more democratic fashion. Section 5 then analyses several current situations that can be regulated with mechanisms based on these principles to ensure a safer online environment respectful of enhanced internet user rights. These principles could also help technology developers design and operate online services that are more respectful of true online neutrality. Finally, section 6 concludes the essay.

2 The internet object

Before discussing its regulation, it is necessary first to describe "the internet". The term is mostly used as a singular reality that can be defined unequivocally. In reality, the internet is formed by a number of disparate elements grouped into several layers of "networks" entangled with each other. Let us present these elements shown in Figure 1.

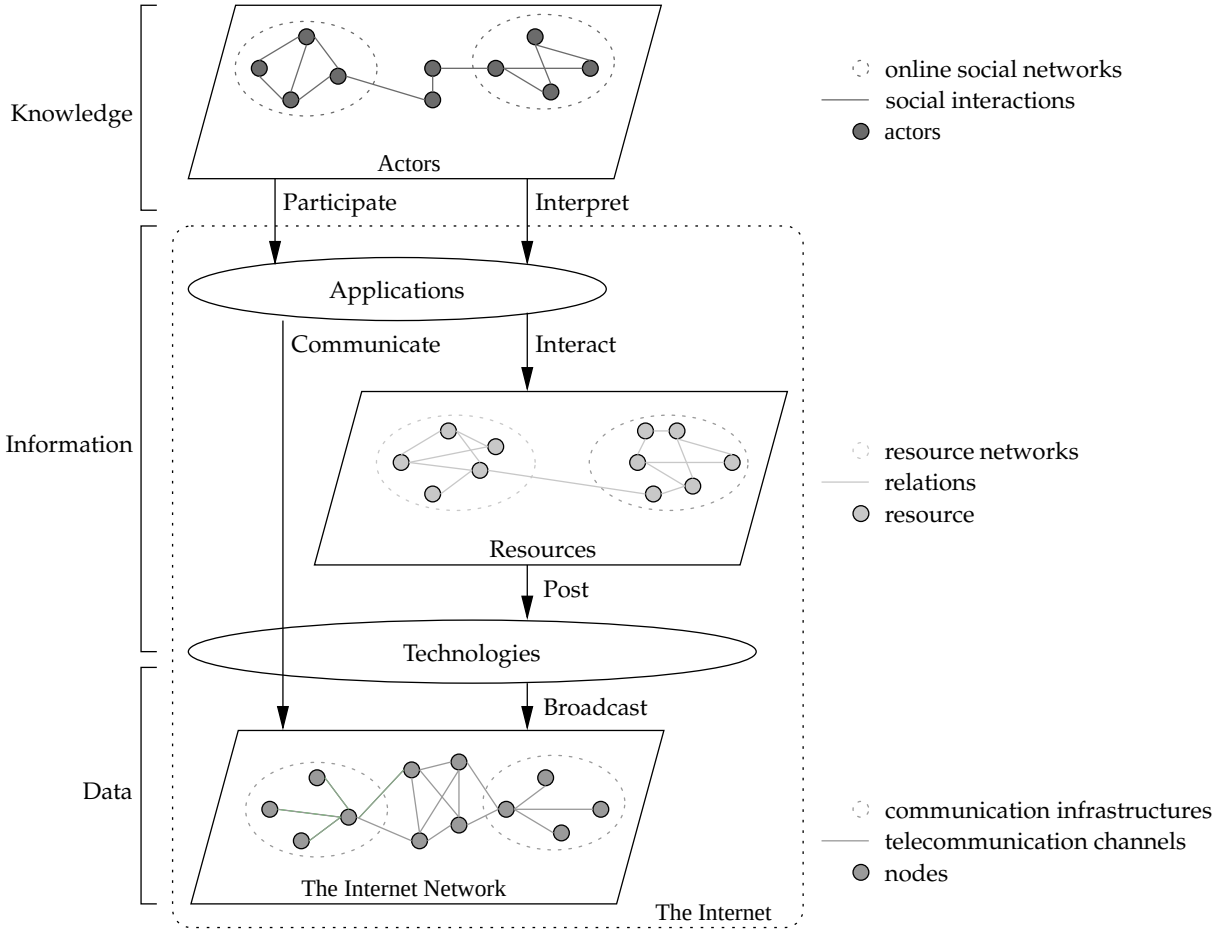

Figure 1: The internet.

2.1 The internet

One must distinguish between *the internet network* and *the internet*. The first characterises the *communication infrastructures* that allow a set of connected devices (called *nodes*) to exchange *data* (series of 0 and 1). A communication network needs at least two nodes (a *transmitter* and a *receiver*) and a protocol, i.e. a method used by the devices to contact themselves and a set of communication rules. Most communications through the internet network are done thanks to the Internet Protocol (IP) and the Transmission Control Protocol (TCP). Known as TCP/IP, they implement *packet switching*. They are based on a basic transfer unit called a *datagram*. When data must be sent over the internet network, they are divided and encapsulated into datagrams. Each datagram takes its own path through the network, and TCP/IP ensure that the lost datagrams are re-emitted and that datagrams are correctly reordered at the destination. Each node on the internet network is associated with an IP address[1], and the only information used by TCP/IP to route a datagram is its origin and destination IP addresses. Moreover, the internet

[1] In practice, some mechanisms exist that allow several devices of a same sub-network (such as all home personal devices) to share a same IP address vis-à-vis the internet network.

network adopts a decentralised approach: the path followed by a given datagram is not computed globally once and for all, but is determined by a series of local decisions at each intermediate node (mostly routers). On the TCP/IP, there is a stack of application protocols (data encryption, email, Web, etc.). All these protocols form a layered architecture called the *TCP/IP model*. The implementation of the internet network follows the *end-to-end principle*: intermediate nodes should be limited to basic functions only (mostly conveying datagrams), while the end hosts manage the application-specific functions (such as a website or a browser).

Datagrams are transmitted through *telecommunication channels* (optical fibres, electromagnetic radiation, etc.). The internet network is the interconnection of communication infrastructures based on different telecommunication channels (home Wi-Fi networks, internet provider networks, internet backbones, etc.). In practice, the communication infrastructures are hierarchically organised, and one can differentiate three different levels. First, the *internet backbones* ensure a global interconnection between cities, states and continents. Most datagrams go through them. Second, the metropolitan area networks (MAN), sometimes called the *last mile*, are used by the internet users to access the internet network. MAN connect buildings to the internet network, and mobiles to cell towers. Third, local area networks (LAN) are intermediates between a device (computers, smartphones, etc.) and a MAN. In office buildings, computers are connected to the internet network through a LAN managed by organisations occupying these buildings. Today, in most homes, a LAN (using Wi-Fi) allows multiple devices and inhabitants to access the internet network.

Originally, the internet provided only remote access to a distant computer[2]. The first users soon imagined other features, such as the exchange and broadcast of *resources* (for example interpersonal messages), and developed technologies that implement them. The architecture consisting of the communication infrastructures and these technologies characterises the internet, with the internet network being only one of its components. One of the characteristics of many online resources is that they are organised in *resource networks*. The best known of these resource networks is the World Wide Web (or simply the Web), where the resources are webpages linked by hyperlinks. Many other such networks exist: each online forum, for example, is a resource network consisting of messages grouped into discussion threads. Resource networks may be interconnected through relations: a forum post can, for instance, contain a link to a website.

Applications[3] compose the highest layer of the internet. They are used to produce, diffuse (for example when an email is sent) or access resources (such as a list of relevant documents related to a given query). An application (such as a Web browser) proposes one or more features (for example rendering a webpage) and implements one or more technologies (HTTP and HTML for simple webpages). Traditionally, applications are separated between *servers* and *clients*. Servers propose resources[4] while clients access these resources. On the Web, for example, Web browsers (client applications) interact with Web servers (server applications) to exchange documents (mostly HTML). From an economic point of view, servers act as producers of online resources and clients as the consumers of these resources. It should be noted that some applications are both clients and servers. So-called Peer-to-Peer (P2P) applications, such as *BitTorrent* or *Gnutella*, access resources (mostly files) proposed by others (client role) while providing their own resources (server role).

The internet network is responsible for the transmission of data. Internet applications provide information though the access to online resources. It is only when this information is interpreted by humans that individuals give it sense and generate *knowledge*.

2.2 The internet stakeholders

Computers and technologies never act independently (although they can act automatically by delegation[5]): they are used as mediators between stakeholders, who may be individuals or organisations. Through certain applications, stakeholders engage in social interactions and participate in *online social*

[2] In fact, the internet is rooted in a project of interconnecting distant computers known as the ARPANET.

[3] An application is a particular socio-technical mechanism, generally proposed through software, with which an internet user, or a group of internet users, interact with the internet.

[4] In practice, a resource (such as a webpage) is built by several applications which can be hosted on multiple computers.

[5] A search engine, for example, responds automatically to requests formulated by internet users. But a search engine is a collection of applications developed by engineers that provides an online service financed by a commercial company.

networks[6], which can be interpersonal, inter-organisational, or a mix of individuals and organisations. *Facebook* and *Twitter* are popular applications hosting online social networks. The emergence of these applications is rooted in the early years of the internet. Newsgroups, for example, have existed since the end of the seventies. The majority of the stakeholders active online interact with the internet to search for information, send emails, post a picture on their *Facebook* page, listen to music, etc. I call them *internet users*.

Some stakeholders are not shown in Figure 1: the *internet operators* that provide the communication infrastructures, applications and online resources to the internet users. While it is well known that the internet has its roots in a research project funded by an American defence governmental agency, since April 1995, its development has mainly been carried out by commercial companies [33]. In this discussion, it is necessary to distinguish between the interventions of internet operators in communication infrastructures, and those of others.

Network operators run the communication infrastructures forming the internet network that allows devices to exchange data. *Backbone operators* provide the internet broadband infrastructures while *access operators* (sometimes called internet service providers) manage the last mile. A same company can be active at both levels. Their business model implies fees paid by those who want to access the internet through their communication infrastructures. There are a relatively limited number of network operators — mostly telecommunication companies.

On the contrary, there are a huge number of stakeholders that develop applications and provide online resources. In particular, many of them propose *online immaterial commodities* in the form of online contents (such as online articles) or online services[7] (for example a search engine, an email service or a cloud computing service[8]). I call those proposing digital goods online immaterial commodity providers or simply *providers*. From an economic point of view, providers are producers of online immaterial commodities and internet users are the consumers of these commodities[9]. While some providers are non-profit organisations (such as the Wikimedia Foundation that hosts *Wikipedia*), others are commercial companies with multiple business models. Besides traditional fee models whereby internet users have to pay them directly (such as some newspaper websites), some providers propose their immaterial commodities "freely" to internet users (search engines and social networking applications are the best known examples) and generate revenue indirectly (the most widely adopted model being online and mobile advertising).

In fact, internet users can also be seen as providers, for example when they write a blog or post a message on *Facebook*. But I make a distinction between a "normal" provider that wants to produce mass online immaterial commodities and a "normal" internet user who does not. I therefore consider Wikimedia Foundation and Facebook as providers while those who write notices in *Wikipedia* and use *Facebook* as internet users. This distinction will become clear while reading this essay.

To put it simply, in the terminology used in the rest of this essay, network operators are active in the communication infrastructures, while providers propose online immaterial commodities. Some companies are both network operators and providers. Several access operators also offer some online services to their customers, such as an email account. But since most of these services depend on a link between an access subscription and these services (i.e. if internet users change their access operator, they lose access to these services), most internet users prefer pure providers. Similarly, the big providers (such as Google) have their own communication infrastructures for performance and security reasons. But, so far, they do not offer internet access to internet users[10] and are not subject to the current regulation mechanisms. So, currently, we may consider that the major global internet players are mainly pure

[6]These online social networks were popularised by HORWARD RHEINGOLD as "virtual communities" [87]. Since several sociologists criticise RHEINGOLD's expression [62], I prefer to use "online social networks".

[7]From a purely technical point of view, online services are also based on content exchange. For example, when a search engine receives a query (content), it builds a webpage (content) that contains the results (including hyperlinks to these results). But I prefer to differentiate online immaterial commodities for which the content is the core (such as an online article) from those for which the content is only the support (the core of a search engine is to propose a list of relevant documents for a given query, with the webpage presenting the results being only a means to interact with the users).

[8]Cloud computing is the generalisation of the use of online services for tasks that until recently were executed locally on computers, such as document storage (Dropbox, Google Drive, etc.) or software use (Google Docs, ERP software, etc.). In fact, it can be seen as a return to the early days of computing when users executed tasks on mainframe computers (online cloud services) through specialised terminals (desktop computers, smartphones or tablets).

[9]Unsurprisingly, providers mainly run server applications while internet users run client applications.

[10]It should be noted that some providers (such as Google and Facebook) have plans to propose direct internet access in future.

network operators or pure providers[11].

In this essay, I shall discuss online immaterial commodities and providers, the latter often having the same name as the online immaterial commodity they provide. I choose to use italics when I refer to an online immaterial commodity. For example, the company Facebook provides the social networking application *Facebook*.

[11]The biggest attempt to create a major player in the infrastructures and resource networks was the merge of AOL and Time Warner, but in the end it was a failure.

3 The network neutrality utopia

The previous section showed the three constituent components of the internet: the communication infrastructures, the online contents and the online services to manage these contents (production, diffusion and mediation). To ensure the neutrality of the internet, the widely defended approach refers mostly to the communication infrastructures. It postulates that the network operators are the only internet stakeholders to be in a position that allows them to discriminate against all the other internet stakeholders. In other words, if the network operators are limited to technical choices only, a non-biased internet provides a "neutral" environment where "autonomous" stakeholders can act "freely". But, as I will explain, this is clearly not the case in practice [34]. In fact, several biases exist on the internet, both regarding its constituent components and internet users. My point here is not to systematically criticise internet operators, in particular commercial companies, and I admit readily that these operators play a crucial role in the development of the internet and, therefore, in its social, political and economic benefits. My aim is to show that the current regulatory framework is unable to ensure network neutrality.

3.1 Online bias framework

First, I shall present a framework for the analysis of biases proposed by BATYA FRIEDMAN and HELEN NISSENBAUM in their article *Bias in Computer Systems* [35]. Since the internet is a network of networks of computer systems, their framework is relevant in studying different online neutrality problems. FRIEDMAN and NISSENBAUM define a bias as something that "*systematically and unfairly discriminates against certain individuals or groups of individuals in favour of others*". They also notice that neither an unfair discrimination that it is not systematic, nor a systematic discrimination that is not unfair are biases.

FRIEDMAN and NISSENBAUM introduce a bias topology based on three categories that can be applied to the internet. First, there are *preexisting biases*, which may be individual or collective, that depend on existing institutions and social practices. They are incorporated in communication infrastructures and in online contents and services without being questioned. For example, Google's search engine implements a well-known principle inherited from bibliometrics: the quality of a webpage (a scientific article) can be measured through the hyperlinks that reference it (the papers that have cited it) [14]. Second, technical constraints or technical considerations introduce *technical biases*. An example is the way some nodes manage network congestion by not conveying all the datagrams they receive (I discuss this below). Finally, when the internet is used by internet users, their practices, which usually appear gradually without being imagined by developers, imply some *emergent biases*. Collaborative tagging applications (*Delicious*, *CiteULike*, etc.) represent a typical case. Here, internet users can freely assign tags (keywords) to resources (webpages, scientific articles, etc.), with the idea being that the "internet community" manually indexes online resources (which cannot be done by a small group of experts). But, while the developers imagined that internet users would use tags that describe only the resources themselves, studies show that some of them also describe their relationships with these resources (by using tags such as "mystuff", "toread", "jobsearch", etc.) [27].

I propose adding a fourth category, i.e. that of *financial biases*, that regroups all the discriminations related to the financial weakness of some stakeholders. In fact, financial biases are, strictly speaking, preexisting biases rooted in the capitalistic development of the internet. However, as I will show later on, the introduction of a specific category for financial biases is motivated by the fact that specific public policies can reduce them.

3.2 Biases in communication infrastructures

Communication infrastructures transmit datagrams between two end points of the internet (for example a server hosting a webpage and a personal computer with a browser), by routing each datagram through a set of intermediate nodes managed by the network operators. These stakeholders take two types of decision: one related to the physical construction of their infrastructures, and the other to the rules applied during the routing of the datagrams across the different networks that form the internet.

Remember, in the current regulatory framework, the role of network operators is supposed to be limited to technical choices only [110]. But deciding where to install which communication channels (optic fibres, electromagnetic radiations, etc.) is not a technical choice only: it is a business decision too. As network operators are mainly commercial companies, their business decisions depend only

on their incentives to invest. Here, the situation depends on the kind of network operator (subsection 2.2). For those who deploy private networks for their own use, their motivation is the quality of their online services (in particular the redundancy of the data and their proximity with internet users to ensure short response times). For the others, they certainly will not invest in a high-quality network if they have no guarantee that a corresponding market exists. This is especially true for the "last mile" networks that can count less on economies of scale than broadband networks. All locations are thus not equal regarding internet access, and a financial bias exists. While highly populated areas provide multiple access services, ensure competition between different access operators and allow internet users to choose the best suited access service, some areas have restricted access to the internet or no access at all. It is probably difficult to force commercial stakeholders to invest to no avail, but this shows that the physical location is a discriminant factor regarding internet access. While industrially advanced nations have very good network coverage overall, this is not the case in every region of the world. Even in countries like the United States, high connection qualities[12] are not always available [16]. Moreover, we know that many communication infrastructures (in particular mobile) must probably be updated to manage future data flows. Existing discriminations will therefore continue.

The way TCP/IP protocols route the datagrams also implies several technical biases. The size of the internet network and the amount of datagrams transmitted make it impossible to compute an optimised routing for all datagrams. The intermediate nodes transmitting a given datagram are therefore chosen progressively by local heuristics[13]. These heuristics can favour one datagram over another due to a lack of a global vision of the internet. When an intermediate node assumes that some network congestion exists, it decides to delay the retransmission of a datagram or to simply drop it (mostly when its buffer is full). The order in which incoming datagrams are forwarded by the nodes follows the "first-in-first-out" principle. Regarding each datagram separately, this seems to be a neutral choice. But a datagram is always part of a larger flow at the application level (having voice communication, exchanging emails, etc.), and a delay in the datagram transmission can be discriminating for one application (such as voice communication) and not for another (email exchange for example). The identical treatment of each datagram flow as is the case with the "first-in-first-out" principle, discriminates indirectly against some applications (and their users) in favour of others. These are not "intentional" biases on behalf of network operators to discriminate against specific datagrams, yet a *de facto* discrimination exists. In other words, even purely technical choices lack neutrality in managing datagrams.

In any case, it is impracticable to limit network operators to technical decisions only [99]. In particular, if illegal and harmful content[14] is to be stopped, one must decide which content is illegal or harmful, and which is not. Such decisions involve choosing which code signatures define malware and which word combinations characterise spam. These decisions go far beyond "simple" technical choices. Moreover, many network operators, for fear of legal procedures, apply rules that lead to *overblocking* [69], i.e. legal and safe data become unavailable because blocking techniques discriminate too much. For example, in December 2008, after an issue with the cover of the 1976 Scorpions album *Virgin Killer* shown on the corresponding *Wikipedia* notice, some British access operators blocked the whole *Wikipedia* website [75].

3.3 Biases in online services

As just shown, even with the current regulatory framework, network operators introduce several forms of discrimination. A closer look at other internet operators teaches us that discrimination exists at every layer of the internet. In fact, each online service used to produce, diffuse or mediate online contents influences — not necessarily always intentionally — the way people interact with these contents. Let us illustrate this with three cases: search engines, recommendation systems and social networking applic-

[12] The connection quality is a combination of criteria that can include upload and download bandwidths, connection time and quantity of data transferred.

[13] For the *Oxford English Dictionary*, a heuristic "*solves problems or makes decisions by trial and error or through empirically-derived rules (often used to obtain approximations when more formal or exact methods are too slow or complex)*".

[14] Harmful content implies that its production, distribution or consumption harm people. This is the case with child pornography, malware or spam. Illegal content only refers to a definition given by the law, for example the distribution of copyrighted works. Harmful contents are mostly illegal but this is not always the case. First, one may defend the idea that considering some contents as illegal is more harmful than letting them be diffused (this is all the debate about the prohibition of racist words). Second, the law may lag behind the practices (spam is not considered as illegal everywhere). At any rate, the illegal character of harmful content is a more general question and must be discussed prior to internet regulation itself.

ations.

First, let us mention search engines. Search engines are today's most used tools to explore the Web, and the only ones for many internet users [83]. Each search engine implements three processes: (i) the indexation of online documents, (ii) the selection of documents to retrieve for a given query and (iii) the ranking of these documents (the top ones are supposed to be the most relevant). Each process discriminates against another. First, due to an exponential amount of online documents, search engines introduce a technical bias by indexing only a very small part of it (called the surface Web). Moreover, the criteria used to choose which documents to index are never explained[15]. Secondly, despite recent developments in pattern recognition and information retrieval, the selection process is still a very complicated task (technical bias). As noted daily by readers, search engines often retrieve unrelated documents. Finally, the main principle behind their ranking mechanisms raises some questions. It assumes that the structure of hyperlinks between online documents determines the most "serious" ones [14]. The first result of the *Google Search* for the query "truth about Holocaust" is well known: when entered, it proposes a denialist site[16]. Of course, Google does not endorse this website, but its official position is not to interfere *a posteriori* on its ranking algorithm and to let internet users alone decide which pages are highly ranked through the hyperlinks they create. If many websites (created by internet users) integrate hyperlinks to a given document, it will have a good ranking. However laudable this position may sound, it is far from being neutral. Search engines clearly choose a Darwinian process to decide which documents are ranked first: those which gather the most hyperlinks are the most fit to survive anonymity on the Web (preexisting bias). These biases do not mean that search engines deliberately manipulate which documents are proposed first to internet users, but that they make hypotheses and take decisions that are not neutral when developing their services. One may argue that internet users accept some discrimination from search engines in order to retrieve only "relevant" documents. And American search engines frequently invoke the first amendment (free speech) to protect the results of their ranking algorithms [39]. But, since commercial search engines (*Google*, *Bing*, *Yahoo!*, etc.) do not document clearly how they work, nobody knows which documents are indexed and which criteria are used to select and rank them. And search engines can decide to censure certain websites. A known example is the systematic discrimination between 2006 and 2009 by *Google Search* against the website of Foundem, a potential competitor [39].

But search engines are not the only online services to distort contents; recommendation systems do it too. Take the product suggestions by online shops (such as *Amazon*). Here, recommendation systems propose products (such as books) to consumers based on previous purchases, stating that "those who have bought this book have also bought these". With time, some algorithms become very complex by including several criteria (such as taking different assessment behaviours into account) [54]. In practice, some discrimination arises because these algorithms mostly suggest popular products (i.e. those which were sold most and had many positive reviews). They are therefore rarely able to make alternative suggestions as a librarian or a bookseller could do (such as "this book has the same atmosphere as this one"). This discrimination has multiple origins. First, a preexisting bias exists with the hypothesis which the algorithms are based on: the quality of a product is defined by its popularity among consumers (through its sales and positive reviews). In other words, the recommendation strategy is based on the law of the market only. This may sound reasonable, but it is not a neutral choice. Second, there is also a technical bias. Since these recommendation systems combine information on products with respect to the preferences of internet users (purchases, reviews, etc.) in order to make suggestions, the products for which less information is available are under-represented. In other words, the most popular products tend to become even more popular. Third, recommendation systems may also introduce an emergent bias when they are based on ratings by internet users (such as *Digg* or *reddit*). A recent study showed that people are influenced by prior ratings when they rate [73]. The researchers developed a website that published news that internet users could rate positively or negatively. These Internet users were divided into three groups. In addition to a reference group, there were two other groups whose votes were manipulated. For the first one, a positive vote was added each time by the experimenters, and for the second one, a negative vote was added each time. The experiment revealed that internet users tend to compensate negative ratings with better ones, and that they give even better ratings when positive ratings exist. Again, I do not claim that online sales websites try to promote some products rather than

[15] It is probably a mix of the popularity of the corresponding website, the number of hyperlinks pointing to the document and the depth of the document in the structure of the corresponding website.

[16] This result was reproduced on January 3, 2014.

others[17] or that rating systems manipulate internet users by displaying existing ratings, but only that they are not neutral.

Social networking applications are not neutral either (*Facebook*, *LinkedIn*, etc.). Everybody knows how they work. Internet users can freely pick up "friends", follow their online activities and endorse them ("like"). On a regular basis, users are prompted by the application to add "friends" (mostly through a "friend of a friend" mechanism). First, a preexisting bias is rooted in the choice to encourage internet users to have as many "friends" as possible. Second, social networking applications assume that each "friendship" is equally "important". Yet people think differently: they have several circles of friends with different priorities. But implementing multiple levels of "friendship" is difficult to manage for internet users and applications. Social networking applications therefore do not integrate such mechanisms and consequently introduce a technical bias. Furthermore, it may become addictive for some internet users (we have to go to *Facebook* very often in order not to miss any information concerning a close friend), and it forces many of us to rethink our conception of friendship (emerging bias). One may argue that it is a virtue of the internet to propose new ways of interacting with people. This may be true, but forcing us to change the way we interact is not something neutral. Moreover, social networking applications can also discriminate against some internet users. Examples include LinkedIn, which temporarily banned Syrian internet users in 2009 [77], and Facebook, which removed a group of 1000 internet users that criticised the Moroccan monarchy [72].

Actually, there are no explicit regulation mechanisms concerning online services (such as forcing some service providers to document their services). But this does not mean that nothing governs how online services work. As LESSIG put it in his well-known expression "*Code is Law*" [65], online services implicitly regulate social practices and social norms. Synchronous communication applications are one example used by LESSIG to illustrate this. Initially, only text was available. Many physical attributes of the real world (disability, lack of beauty, etc.) were hidden. But today, images and videos are also available, and physical differentiations become visible online. This changes the nature of relationships in cyberspace. Online services also reinforce the role of popularity as the structuring factor of online interactions: search engines retrieve the most popular documents, online sellers recommend the most popular products, *Twitter* suggests following the most popular accounts, etc. Finally, online services tend to specialise our interests and tastes: online sellers recommend products similar to those previously bought, *Facebook* proposes friends of our friends as friends, etc.

This discussion may sound far-fetched. But on September 8, 2010, a French court used similar reasoning to sentence Google and its chief executive for public defamation[18]. The court held that the "autocomplete" feature of its search engine is discriminatory. In fact, an individual prosecuted Google because, when his name was entered in the search field, *Google Search* suggested the additional keywords "rape", "condemned", "Satanist", "jail" and "rapist". This and other more recent cases[19] show that online services are not neutral and that their providers can be considered as responsible for the corresponding discrimination.

3.4 Biases in online contents

Communication infrastructures and online services thus influence the way online contents are produced, distributed and mediated, and impact how internet users interact with these contents through the internet. Beyond the biases introduced by these services, the online practices of internet users also create emergent biases regarding contents. In particular, two phenomena deserve to be discussed. First, let us mention online manipulations whereby malicious or dishonest stakeholders exploit communication infrastructures and online services to intensify their biased behaviour, and to distort the way online contents are presented to the internet users. Second, let us mention what I call the *fetishism of the internet*

[17] The behaviour of these algorithms, enhancing the sales of a "limited number" of products, has competitive advantages for the sellers: they can negotiate interesting prices when they buy the most sold products and their stock does not need to contain a lot of each product.

[18] `http://tinyurl.com/bwfmxxd`, decision of the Tribunal de Grande Instance de Paris, 17ème chambre (consulted on January 30, 2014).

[19] `http://tinyurl.com/dxsrhxh`, the French Cour d'appel de Paris Pôle 2, chambre 7 sentenced Google on December 14, 2011 because its "autocomplete" feature suggested the keyword "rogue" when the name of a company was entered (consulted on January 30, 2014). `http://tinyurl.com/qc8cgdh`; a similar court ruling was also issued in France by the Cour de cassation, Première chambre civile on June 19, 2013 (consulted on January 30, 2014).

[34]. I define it as a misunderstanding of the functioning of the internet by most internet users who see it as a black box and are unaware of its biases.

Let us go back to search engines. One might think that *search neutrality* requires that search engines retrieve a balanced presentation of different points of view regarding a given question formulated through a query. But this objective is impossible to reach [39]. As explained above, their ranking algorithms only express the position of documents in the hyperlinks graph representing the Web and, for technical reasons, only a small part of the internet is indexed by search engines. These technical biases could be described as a form of censorship (although not necessarily intentional) which internet users are unaware of. There are also several emergent biases. First, some stakeholders artificially manipulate the ranking of a given document, for example by creating a set of hyperlinks pointing to it, a technique sometimes called "Google bombing" [2]. Second, because most internet users do not know the technical limits of search engines, they consult only the first documents retrieved [47]. Therefore, they almost always miss several of the results and have a truncated view of the object of their search.

The list of the most visited websites places *Wikipedia* at the top[20]. The "*multilingual, web-based, free-content encyclopedia project*" gives access to millions of articles written collaboratively by (sometimes autonomous) contributors. An often cited study published in 2005 shows that the quality of 42 scientific articles from *Wikipedia* and *Encyclopædia Britannica* is similar [36]. Researchers have shown that the quantity of contributors and readers of an article explains its quality: the more people consult a given article, the faster errors are discovered and corrected [102]. In practice however, they also pointed out that the articles which are revised often are not necessary those which discuss controversial subjects. In other words, many articles do not have many contributors and readers, and some of them are written by a few contributors only who reflect their very personal position. It is always possible to consult the "discussion tab" to evaluate the number of contributors to an article (and verify its quality). But most internet users do not know this or do not take the time to make this verification. Since all articles are presented the same way *prima facie*, there is therefore a risk that the same level of confidence is given to all of them. Yet, with its *WikiScanner* that links article revisions and the IP addresses used to edit them, VIRGIL GRIFFITH detected that commercial companies and governmental organisations purge *Wikipedia* of information considered as disturbing [38]. Furthermore, while *Wikipedia* defends a strict "neutral point of view" in its articles, each linguistic version comes with its own preexisting biases (mostly cultural). For example, a comparison of articles about famous personalities in the English and Polish versions of *Wikipedia* shows that perspectives vary across cultures [18]. Another study reveals a gender bias: women contribute to only 9% of the revisions and represent only 16.1% of the editors [59]. *Wikipedia* is, of course, a wonderful and useful project, but it cannot ensure true neutrality to its readers.

What is more, as described previously, recommendation systems of online sellers tend to present the most popular products. This principle is also used to rank products when an internet user performs a search on the corresponding website. As for search engines, this ranking can be manipulated. Suppose I wrote a book on "the internet" and that, in order to favour sales, I wished to see it ranked well on a given online sales website. To achieve this, I could create a dozen dummy accounts on that website. With each account created, I buy a copy of my book and write a very good review of it. Because most books do not have many reviews, I would push mine to the top artificially.

Since its foundation in 1996, *Internet Archive* proposes a digital archive of the Web. It tries to prevent valuable knowledge being lost due to the non-persistent character of the Web[21]. In practice, it tries to store a copy of many websites[22]. For internet users, it should be a guarantee that relevant contents cannot disappear. But in 2002, the archives of the website *Xenu.net* — well known to be very critical to Scientology — were removed under pressure from the lawyers of the Church of Scientology [13].

We can multiply such examples to show that no provider can guarantee neutrality of the contents it produces and/or mediates. I am not saying that online contents are useless or always dangerous. My point is that they tend to reinforce the thoughts of the most active internet users, and that they can be influenced by some unfair practices which are difficult to detect.

[20] `http://www.alexa.com/topsites`, Alexa Internet is a company which provides Web traffic data (consulted on January 30, 2014).

[21] There is no guarantee that a document placed online remains accessible in time (for example because the website hosting it closes).

[22] Not all pages of a given website are stored, mostly for technical reasons (for example some dynamic Web pages).

3.5 Biases against internet users

So the internet is not perfect and, sadly, many online services may have a dual use. Despite its weaknesses, one may assume that the internet provides an environment where all internet users are equal with respect to its biases. But there is an internet asymmetry between people that materialises in the digital divides (geographical, social, educational, etc.). These digital divides increase online biases for some internet users while they leave others unaffected.

A geographical digital divide is clear. Depending on where one lives, the connection quality of internet access (if any) varies. In addition to people who are completely disconnected from the internet, those who are confronted with low bandwidths or irregular connection sessions also face increasing discrimination. Even the availability of high connection qualities everywhere cannot prevent a geographical digital divide, since internet users face different access environments depending on where they live (constitutional rights, internet access pricing, consumer rights, privacy, etc.). Free speech is certainly the best example of this disparity. In 2012, the OpenNet Initiative reported that 47% of internet users face censorship of online content [80]. But differences also exist between countries with comparable political systems. For example, since June 2006, Belgium has extended the right to the protection of sources to bloggers who diffuse information to the public [51], while many democracies still limit it to "traditional journalists".

Within the same country, there is a social digital divide between citizens, even in "advanced" democratic countries. With an internet penetration rate of 77.3% estimated in June 2010[23], the United States (US) appears to be a highly connected place for citizens with very broad rights. But a closer look tells a different story: while New Hampshire peaks at 90.1%, more than 40% of US citizens living in Mississippi are unconnected[24]. Such a significant difference (34.2%) is also observed with respect to their per capita personal income in 2012 (23.7% with $44,129 for New Hampshire and $33,657 for Mississippi[25]). Certain categories of citizens are thus more likely to be excluded from the Net (the unemployed, the poor and the elderly). Meanwhile, more and more information is available only on the Web (such as what to do to obtain scholarships), many administrative tasks become difficult to do offline[26], and many companies mainly use the Net to recruit (which implies employment discrimination for unconnected people). In a context where the internet is seen as a powerful tool to enhance the political participation of citizens, those without "good" connection qualities are second-class citizens. This discrimination becomes worse with the current economic situation.

While the abovementioned points are well known, there is also a little-known digital divide in terms of education. Many of the internet's weaknesses may be offset with good use practices, but not all internet users know these practices. While most internet users can *use* online services, only a few *understand* them. Everybody can, for example, enter a few keywords in a search engine. But few of them master query techniques sufficiently with logical operators to target their research needs precisely [37], or do not know the ranking principles well enough to interpret the results provided by search engines. The advent of the "Web 2.0" (blogs, social networking applications, collaborative tagging applications, etc.) is supposed to allow anyone to contribute to online content. But a study carried out between 2000 and 2008 shows that Americans with secondary school education contribute most to "Web 2.0" generated contents [91].

All of these elements demonstrate that the internet cannot be seen as an extension of the public sphere as idealised by HABERMAS [40]. In practice, it is impossible for billions of people to express their viewpoints or defend their ideas. In parallel, those who defend extreme ideas (terrorists, racists, etc.) and use online services skilfully can occupy a larger part of cyberspace than the popularity of their minority ideas suggests [9]. Several scholars go even further and speak of cyberbalkanisation to describe the fact that most internet users reinforce their own ideas through online interactions rather than being confronted with a "neutral" presentation of all opinions on a given subject [96].

The internet can even create emergent biases in the offline world. One scary example comes from companies that evaluate people's solvency based on their online social networks (such as *Twitter*, *Facebook* or *LinkedIn*) [29]. They assume that highly connected people with many "friends" and "followers"

[23] `http://www.internetworldstats.com/stats26.htm#america`, *internet World Stats*, run by Miniwatts Marketing Group, evaluates world internet usage and population statistics (consulted on January 30, 2014).

[24] *Ibid*.

[25] `www.statsamerica.org`, STATS America is a service of Indiana University (consulted on January 30, 2014).

[26] In Belgium, banks close down many branches and favour their online banking system.

have a greater chance to find a new job quickly if necessary (and are thus in a better position to continue to reimburse their debts). In other words, they discriminate against unpopular people online. In today's big data context where an increasing number of decisions are made based on digital information (partially gathered from the internet), citizens without an active online life risk being systematically disfavoured.

3.6 Current regulation scope

Unlike the collective imagination, the internet has always been regulated since its creation. Its core, the TCP/IP protocols, presupposes that IP addresses are centrally managed. Initially administered by JON POSTEL, the management gradually became the responsibility of different organisations. Today, it is the Internet Assigned Numbers Authority (IANA), a department of the Internet Corporation for Assigned Names and Numbers (ICANN), that manages the IP addresses and the Domain Name System (DNS)[27]. In practice, ICANN determines general rules and allocates blocks of IP addresses, and delegates some executive tasks to third-parties (for example the commercial company Verisign manages the ".com" domain name, and blocks of IP addresses are allocated into smaller ones by regional internet registries).

Several non-profit organisations play an important role in developing technological norms for the internet. The three best-known are probably the Internet Architecture Board (IAB), which proposes new technologies for the internet network (in particular protocols), the World Wide Web Consortium (W3C), which works on standards for the Web (such as HTML), and the Institute of Electrical and Electronics Engineers (IEEE), which produces several standards (for example the IEEE 802 LAN/MAN group of standards).

Since it became public, the internet has also been regulated by states. Here, is its necessary to make a distinction between democratic regions and authoritarian regimes. Most people live in (more or less) authoritarian regimes, where internet stakeholders, especially internet users, lack freedom in how they interact with the internet (censorship, systematic control, arrests, etc.). In democratic regions, internet regulation was first seen as a telecommunication problem. This explains why network neutrality is mostly limited to regulating how network operators transmit data over the internet network. With time, some democratic regions have enhanced the regulation of the internet including free speech (the European Union (EU) prohibits racist words published online), privacy (some regions provide large online privacy rights), or illegal and harmful contents (spamming is considering an illegal activity in several countries).

Some of these regulation rules will be discussed below. Suffice it to say that, in practice, most of these regulation rules impact network operators and almost never impact providers. Moreover, the providers of the most popular services used by internet users living in democracies are based in the US where internet state regulation mechanisms are less implemented than in some other regions (such as privacy in the EU).

The previous subsections show that network operators and providers also regulate the internet directly and indirectly. First, as argued by LESSIG [65], the engineers who develop the technologies deployed on the internet influence how internet users interact with it. By choosing the rules that specify the content to be blocked or by defining the algorithms that rank documents and products, engineers favour, often unintentionally, some stakeholders over others. Second, internet operators, which are mainly commercial companies, take many decisions with their business interests in mind. When network operators choose to deploy high broadband communication channels in some regions or when providers defend some loose privacy rules in order to better profile their internet users (and thus better target online advertising), they do not necessarily do so for the common good.

So there is a double myth concerning internet regulation: that its current scope is limited (whereas in fact it is just state regulation which is limited in democracies), and that it is sufficient to ensure fair and neutral internet use for every stakeholder.

[27] The Domain Name System (DNS) is a mechanism that associates domain names (such as `www.google.com` or `www.otlet-institute.org`) with IP addresses.

4 An extended internet regulatory framework

I have tried to show that the current regulatory framework cannot guarantee any form of network neutrality. I shall now put forth the argument that an extended regulatory framework could decrease internet asymmetries, and that, if some general principles are respected, new regulation mechanisms would do more good than possible harm. Or, in other words, I believe that by restricting some players a little more, we could protect the majority more effectively.

4.1 Origins of the network neutrality principle

Historically, original internet users have never welcomed any form of state control even though the internet was initially funded by United States (US) governmental agencies, mainly the Defense Advanced Research Projects Agency (DARPA). Most of them were rooted in the hacker culture. Emerging in the early sixties, the hackers formed a developers' collective interested in producing the best possible software [66]. Progressively, they become digital activists who defended the free diffusion of knowledge and the free access to computer resources, and who denounced the increasing control of the computer world by commercial companies [98]. The early internet was developed by hackers for hackers: they were both the producers (mostly voluntary) of online contents and services, and their consumers. The only component that they did not control was that of the communication infrastructures.

In 1996, their vision of the internet crystallised in BARLOW's *A Declaration of the Independence of Cyberspace* [4]. Since then, many earlier cyberactivists have advocated the self-management of the internet by its users. They perceive public regulation mainly as attempts by states to limit individual freedom and promote the commercialisation of the internet. Scholars have also argued that the success of the internet lies in its adoption as a communication platform without state intervention [41]. The idea of a neutral internet whereby everyone can find information in order to act as an independent and autonomous citizen is certainly shared by every proponent of democracy. In fact, there is some similarity between packet switching and the way US democracy was conceived: a few commonly admitted rules (the TCP/IP protocols) at federal level (the internet) and great freedom of action left to citizens (the nodes and the packets), without a single stakeholder (a hypothetical central node) dictating the overall management of the state (the internet) [33].

There are thus multiple origins of the network neutrality principle inherited from hackers. Another one is a preexisting bias I call the *Silicon Valley habitus*. The expression, derived from PIERRE BOURDIEU's definition of a habitus [12], refers to the social norms and social practices shared by most engineers who develop the technological artefacts of the internet which structure these developers and their artefacts. It has two key characteristics: the belief that technological innovations would be considerably curbed if the private sector is regulated, and a Saint-Simonian faith that these technologies (in particular computers) always have a long-term positive influence on humanity. Google's mission statement "*to organise the world's information*" and its unofficial slogan "*Don't be evil*" best illustrate this.

Added to this is the US phobia of everything that resembles state intervention (as the Obamacare programme has recently demonstrated again). Many internet researchers and engineers do not even seem to conceive that a form of internet regulation organised by the state can be useful (even if the state does not regulate itself).

Finally, some economic stakeholders defend the network neutrality principle because its current interpretation, i.e. constrain only network operators, is in their interest [99]. In particular, providers oppose any form of differential pricing for users of communication infrastructures based on the quantity of traffic they generate.

4.2 A case for an extended internet regulation

Of the multiple motivations of those defending a limited regulatory framework, the one to certainly save is the democratic vision of the internet shared by most of the defenders of network neutrality. But, as I argue in section 3, in today's online context, it does not correspond to the current reality. A relatively small number of people (engineers, highly educated internet users, company executives and technology evangelists) take technical, business and ideological decisions that structure how millions of internet users interact online and interpret the contents they consult. Therefore, I think that new internet regulation mechanisms should be adopted, and that they should not be limited to network operators.

One recurring argument against any extension of online regulation is the fear that it will encourage governments to control the internet. I do not subscribe to this point of view. First, no matter what democracies do, authoritarian states control "their part" of the internet anyway. Second, as the media disclosures on the surveillance programme PRISM have shown, the current online regulatory framework in democratic states does not prevent them from organising a massive (online) surveillance system [7]. I intend to show that an extended regulatory framework can provide new protection for citizens in democracies. Third, although the US government has gradually withdrawn from the management of the internet network, it maintains potential control through the Internet Corporation for Assigned Names and Numbers (ICANN) [74]. In March 2014, the US government announced that an international group to be created will receive the remaining control over the internet network management when the contract with ICANN expires on September 30, 2015[28]. But no details exist today on the nature and organisation of this international group. Furthermore, making an international independent group responsible for some technical decisions (in particular those related to network protocols, IP addresses or domain names) will not solve all the problems described above. So, while the US government's possible renunciation of its power over the internet network is welcome news, it does not eliminate the need for extended regulation rules. Fourth, it is possible, particularly in democracies, to introduce regulation mechanisms that exclude governmental influences (as is the case for the regulation of food or pharmaceutical industries). I therefore think that this argument should not avoid proponents of democracy to defend new regulation mechanisms.

A parallel can be drawn with the food industry. Due to a lack of competence, most people cannot evaluate which product is healthy, and malicious food producers can use dangerous additives to make their products look better than they are. To promote public health, states introduce a regulatory framework to supervise food safety. The online context is similar: most internet users cannot evaluate all the implications of consuming online immaterial commodities, while pernicious internet stakeholders can use technologies to track people's actions or control the broadcast of contents. It therefore seems natural that democratic states can provide protective internet regulation mechanisms. Considering the important social and economic role of the internet in many countries, it can even be considered as a public good, something "*which all enjoy in common in the sense that each individual's consumption of such a good leads to no subtraction from any other individual's consumption of that good*" [90]. If such a position prevails, states must participate directly or indirectly in certain aspects of the development of the internet. In fact, several intergovernmental organisations (UN, OECD, French Constitutional Court, etc.)[29] have already recognised that internet access constitutes a sort of fundamental right.

Ideally, since the internet is global, new regulation mechanisms should be introduced at international level. But this is impractical because of the legal heterogeneity between different geographical areas, even if they share similar values. Take free speech, for example. The First Amendment to the US Constitution provides very broad rights regarding free speech, the European Union (EU) believes that a minimum level of censorship is acceptable (for example, racist words are prohibited), and too many countries simply ban free speech altogether. On the other hand, it is certainly possible to find international agreements on specific subjects, such as a global fight against child pornography [48] or spam. Some scholars even call for an international treaty on cyberwar, in particular to try to differentiate military infrastructures from public ones as targets of cyber attacks [84]. But the adoption of international internet regulation mechanisms is a very long and erratic process. Several initiatives have been launched in the past, but we have not witnessed any real progress [75]. It is therefore necessary to adopt a multi-level approach. On the one hand, countries should start to negotiate some minimal international treaties within existing intergovernmental organisations. On the other hand, national and regional regulation mechanisms should be adopted quickly. A comparison can be made with the fight against tax evasion: states take (sometimes severe) local measures although they cannot dismantle all tax havens globally. While less effective than global internet regulation mechanisms[30], local ones are much easier to implement (it is "only" necessary to convince the political majority). The best strategy

[28] `http://tinyurl.com/pyeq4mw`, official Website of the US National Telecommunications and Information Administration (NTIA) (consulted on July 25, 2014).

[29] See, for example, the May 2011 report of the UN Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (`http://tinyurl.com/3czyazo`, consulted on July 25, 2014), or the OECD statement in June 2011 (`http://tinyurl.com/ofczcz7`, consulted on July 25, 2014).

[30] Let us use spam as an example. It can be blocked with specific software installed on email servers, and local regulation mechanisms can force email providers to install such software. But, to stop the cyber criminals that send it, international collaboration is needed.

seems to be a bottom-up one: democratic states set a good example by themselves and, if possible, work in cooperation with others (for example through the EU) or within intergovernmental organisations to adopt global regulation mechanisms. This prevents some countries from boycotting any regulation mechanism from the start.

Concerning the regulation itself, LESSIG rightly identifies four distinct but interdependent constraints that regulate stakeholders: legal constraints, social norms[31], the market and the architecture (the "code" in the case of the internet) [65]. Legal constraints can be achieved by public authorities through laws (for example, the EU directive on privacy in electronic communications) and the delegation of regulatory powers to regulators (such as those whose mission is to ensure that data privacy law is applied). Regulators have several advantages: they can build and maintain the necessary technical skills to manage the complexity of certain regulatory and supervisory tasks; they can rapidly implement new mechanisms (either by creating additional constraints to react to new threats or by relaxing existing ones to avoid adverse unintended consequences); they can take the context into account more easily; and they can perform investigations or audits. Contrary to previous technological (r)evolutions (telegraph, nuclear power, etc.), today's technologies (not only computer-based ones) evolve much more quickly in comparison to the "legislative time". While in June 2012, the EU was still in the process of specifying how cookies[32] (a technology developed during the first half of the nineties) should be implemented by websites, the first prototypes of Google Glass were built. It therefore seems unrealistic to bet on laws only to regulate the internet. In fact, laws should define the scope of what technologies cannot be used for and which social norms are unacceptable, while regulators should specify the implementation constraints on technologies and verify their online uses. Let us take spam to illustrate this. Since everybody agrees that it pollutes the internet and causes harm[33], it is necessary to qualify it as a criminal activity and to sanction those who send it (role of the law). As cyber criminals rarely reside in the same jurisdiction as their victims and international collaborations are currently uncommon, actions must be taken for the "local protection" of internet users from spam. To fight against it, network operators and email service providers should be forced to deploy tools that identify spam in order to stop it (regulators act on the architecture). In a democratic context, it is also necessary to verify that anti-spam techniques do not discriminate against "normal" content or spy on the activities of internet user (regulators act on social norms). Subsection 4.4 will briefly discuss the organisation and attributions of an ideal regulator.

4.3 Three regulation principles

In addition to regulating network operators, current regulation mechanisms restrict online contents (sometimes harshly), privacy (in some regions) and occasionally competition (for example by forcing Microsoft to allow its users to choose a browser other than *Windows Explorer*). But a regulation framework for the internet must not only extend the current regulatory framework (to provide better protection), but also limit it (such as limiting the state's cyber surveillance in democracies). Moreover, whenever possible, existing "offline" regulation rules should be extended to the online context rather than inventing new ones.

The proposition presented here is first to define some general principles that internet regulation mechanisms must respect, and then to specify *ex post* constraints and specifications when necessary. I believe that many questions raised by internet regulation may be solved with three hierarchical ordered principles: a *constitutional principle*, a *safety principle* and a *competition principle*. I will show next that each principle presupposes that existing offline regulation rules can be "naturally" extended to the online context. It should be noted that internet stakeholders always reside physically in a given jurisdiction (also if it may be sometimes nearly impossible to determine in which one). Some of the mechanisms presented below may imply the establishment of new laws, while others presuppose that new attributions are conferred to regulators.

[31] The importance of social norms shows that internet users can, through their behaviour, directly regulate the internet. For instance, if they were more aware of privacy issues, they could force providers to better protect it (by boycotting those online services they feel are too intrusive). This underlines the necessity of digital education for citizens (especially for digital natives).

[32] A cookie is a piece of information stored by a website on the computer of an internet user that it can retrieve during future visits. It is used, for example, to remember the language of a website: the first time an internet user browses a multilingual website, he or she chooses the language; the next time, this language will be automatically used. Since cookies can store any kind of data, it is a real concern regarding privacy.

[33] Not only does spam fill our mailboxes, but its circulation is paid for by everyone since every internet user contributes to the funding of the communication infrastructures. Spam therefore differs from junk mail which is paid for by the senders only.

4.3.1 The constitutional principle

The constitutional principle seems evident at first glance. While some laws guarantee a right offline (such as the privacy of personal mails), a similar provision must exist online (personal emails remain private too). Constitutional rights, such as freedom of speech, privacy or freedom of association must be guaranteed by states on "their part" of the internet. Let us illustrate this with a hypothetical case. On two social networking applications, *Sina Weibo* and *Facebook*, an internet user posts a message criticising an authoritarian regime and calling to collective action. This message is then removed by the corresponding provider. On the one hand, concerning *Sina Weibo*, nothing can be done since its provider is located in China where the government has an arbitrary vision of freedom of speech and fights any attempt at collective action [52]. On the other hand, a fight in the US courts should be launched (for example by an organisation such as the Electronic Frontier Foundation) in order to condemn Facebook.

But the constitutional principle also has some surprising consequences. Take private property for example. While everybody knows that some private property problems occur with the online sharing of copyrighted materials, most people ignore that some of them can also happen in syntactic worlds (*World of Warcraft*, *Second Life*, etc.). In 2009, the Finnish courts sentenced a *World of Warcraft* player for the unauthorised access to another player's account to steal his possessions in the game [103]. Another example is the freedom of information legislations: many constitutions guarantee a right to access public information. With public information increasingly stored in digital documents and databases, the constitutional principle can have technical implications such as using open standards only. In fact, the freedom to access public information cannot be conditioned by the use of a particular product (such as a specific spreadsheet). If some countries decide to consider the internet as a public good (just like water), internet access with minimum connection quality should be ensured for every citizen. Finland made such a choice in October 2009: a connection with a minimum speed of 1 Mb/s should be available to everyone as of July 2010[34].

Many countries suppose that citizens can make "fair uses" of "unfair methods". The best example are sit-ins that involve the occupation of given area (sometimes to block access to another area) as a form of protest. Some countries also seem to recognise this sort of action as legitimate in the online context. For example, in 2001, a German court acknowledged the legitimacy of a distributed denial of service against a flight company to protest against its participation in the expulsion of asylum seekers [72].

An important constitutional issue related to information and communications technology is cyber surveillance. The internet and connected devices (video camera, radio-frequency identification tags, sensors, etc.) provide unlimited capability to collect and process data describing citizens and their actions in detail. Theoretically, democracies govern how states can monitor their citizens, and, sometimes more severely, what private stakeholders (companies, private detectives, etc.) can do. In practice, however, there is a lack of control. In particular, from the state's security perspective, we are witnessing an ontological modification of our juridical status. All democratic constitutions used to guarantee that we were innocent until the contrary had been proved. Today, we are all considered guilty in power since without any *a priori* reason or prior mandate, the criminal investigative and intelligence-gathering stakeholders of many states collect our personal data. It seems that, under the fear of terrorism, the democratic balance between security and privacy is lost. More than ever, a renewed debate regarding this balance is urgently necessary because many questions still remain, and many of them cannot be solved with new technologies only [67]. Existing regulation rules concerning cyber surveillance in democracies, such as the Patriot Act in the US, should probably be revisited to limit what can be done (not only the scope of cyber surveillance but also the data conservation period) and to ensure a better protection for citizens.

4.3.2 The safety principle

Malware, online scams, cyber attacks and online harassment demonstrate that the internet is not always a safe place to be. The safety principle presupposes that public authorities ensure *digital safety*, which can be defined as an activity of regulating the handling, broadcasting, and storage of digital data in ways that prevent damages. This definition encompasses the protection of internet users against aggressive

[34] `http://tinyurl.com/pgsgk7a`, website of the Finnish national broadcasting company YLE (consulted on September 18, 2014).

actions (from criminals, companies or governments) and their right to access the necessary information to make rational use of the internet.

Aggressive actions comprise two well-known categories: personal data protection and internet security. The first category refers to the ownership of the data uploaded by internet users. It includes the conditions under which personal data are stored (mostly the security level implemented by operators to protect them), transferred to third parties (which should be prohibited without the explicit consent of users) or used (some countries impose operators to declare each particular use of personal data). Several practical questions are raised with respect to data ownership. Can internet users easily retrieve and/or remove the data they uploaded to a given online service (such as a social networking application)[35]? Can internet users prevent a provider from using the data they uploaded for purposes not documented at the time of the upload (for example if the service was acquired by another company)? The difficulty is that privacy is treated differently from one region to another. In the EU, for example, a directive provides a large set of online privacy rights to the citizens and requires the states to ensure that they are respected [17]. In fact, as members are forced to transpose EU directives in their national laws, privacy can be considered as a constitutional right (and thus covered by the constitutional principle). But in the US, another democratic region, an opposite view prevails: most Americans are more open to the use of personal data by online services if their uses are documented [85], and personal data uses are based on sectoral auto-regulation mechanisms [46]. The problem of privacy will be discussed briefly in subsection 5.4.2. The second category of aggressive actions is related to data downloaded by internet users. The goal is to block as much as possible upstream malicious code (such as virus) or malicious contents (for example spam). Since the constitutional principle has priority, technical mechanisms deployed to block malicious data (in particular filtering techniques) cannot violate it (freedom of speech must be guaranteed, for example). In fact, in democracies, overblocking is probably the biggest problem since most operators do not document their filtering techniques and tend to block rapidly and widely [69]. Regulators should be the principal stakeholders to limit as much as possible these aggressive actions. They should be responsible for the verification of compliance with applicable laws (when they exist) and for setting the rules that network operators and providers must follow. For instance, regulators should be able to inspect network operators to check that they do not keep communication data over an existing legal period or that their filtering techniques do not overblock.

The right to information has two dimensions. First, it requires that producers must provide detailed information about their products (as with food products that must list all their ingredients). With respect to the online context, this requires internet operators (networks, online contents or services) to document how their technologies work. As explained in section 3, there are several technical biases that are linked to technological limitations (indexation of the Web, limited information on many products, etc.). While detailed information cannot suppress these biases, it can help internet users to take them into account when they interact online. Many regions, such as the EU, require that access operators provide transparent details about the information needed to assess the connection quality of their access services [44]. But, currently, this information does not concern the filtering techniques they use. So the disclosure obligations should be extended. Some researchers have asked that search engines provide more transparency about retrieval changes and publish clear guidelines about their ranking algorithms [104]. Search engines are against this, arguing that it would reveal crucial information to competitors. But other economic stakeholders have similar constraints (such as pharmaceutical companies that must reveal the active components of their drugs), and revealing the criteria and their influences does not imply that the exact formula must be disclosed. I think that regulators should require network operators and providers to document their practices (use of personal data, recommendation algorithms used, filtering techniques, etc.) once their online immaterial commodities are consumed by a "certain number" of internet users[36]. This is especially true, as regulation mechanisms can be slow in adapting to the evolution of the internet, so that information is the main tool for internet users to protect themselves while interacting online [65]. Other general information could also be available such as the energy footprint of the different services. An experiment has shown that it is possible to estimate the total carbon footprint of a news website by evaluating the energy necessary to generate the content, transport it through the

[35]Several providers preserve data that internet users have uploaded (posts, pictures, etc.) even after they "delete" their account from an online service. In practice, "deleting an account" is interpreted by too many providers as "make an account invisible online" and not as "deleting all the data related to that account" what is the intent of internet users.

[36]The threshold can be a fixed number (for example 10,000 internet users) or depend on the level of privacy of data managed by the online service.

internet and display it on end-user devices [82]. One approach to protect partly what companies claim as their intellectual property could be to only give regulators access to this information. The latter then have the obligation to inform the public about the main aspects and the principal biases of these online services. Second, as explained above, some countries provide a freedom of information that allows citizens to access general interest data held by governments. In the online context, regulators should verify that these data are accessible through the internet with well-defined application programming interfaces (API) and open standards only. It would allow citizens and civil organisations to access public data from several government sources and analyse them with computer programs.

As for food safety, most guarantees provided by the safety principle can be ensured by regulators. Some regulation mechanisms may need specific laws. For instance, a "right to forget" as proposed by MAYER-SCHÖNBERGER [70], and currently discussed in the EU, will probably require setting a data conservation period beyond which stored data must be deleted without the explicit confirmation of an internet user. But, since it is crucial to adapt regulation mechanisms rapidly to the current reality (not only to protect internet users but also to promote innovation by companies), a purely law-based approach is unrealistic.

4.3.3 The competition principle

Everybody agrees that the development of the internet is characterised by a continuous flow of fast innovations. That would have been impossible if a single (governmental) entity had controlled every aspect of the internet and prevented private stakeholders (internet users, non-profit organisations and commercial companies) from organising a large part of it. In fact, since the nineties, the development of the internet has been left almost exclusively to private stakeholders, in particular commercial companies. Without them, many useful online immaterial commodities would not exist. But it is also clear that this evolution was possible because the internet remained "*neutral to ensure that competition remains meritocratic*" [110]. In order to ensure a similar evolution in future, a competition principle must prevail such that no internet stakeholder is favoured. In other words, as the evolution of the internet is mostly guided by the market, it is necessary to ensure that the market works efficiently. Economic history has shown that this is impossible without certain forms of regulation. In particular, we must prevent stakeholders, in particular commercial companies, from using their dominant position on the market to kill competition. This is particularly important since several internet companies hold a *de facto* monopoly (such as Google, Amazon and Facebook). The competition principle raises at least four issues: regulation of monopolies, open access to infrastructures, company taxation and provider switching.

But we must first define what the online market is. The current regulation framework, which presupposes that only network operators should be controlled, implies that the internet is formed by two markets: the communication market (infrastructures and transport services corresponding to the communication protocols of the TCP/IP model) and the online immaterial commodities market (online contents and services corresponding to the application protocols of the TCP/IP model). Since the proposed regulatory framework is also aimed at regulating the providers, considering all the online immaterial commodities as one unique market can hide some problems. In fact, while it can be justified by the horizontal integration strategy of some providers (such as Google), most of them do not have this strategy. I propose therefore to define three categories of interdependent markets that will, with the communication market, serve the following discussion: one market per online service type (search engine, email, etc.), one market per platform (*Windows*, *Android*, *iOS*, etc.) and one market per source of revenue (for example online advertising).

Monopolistic positions can emerge when there are some horizontal integrations (a stakeholder uses the revenue generated from its dominant position on a market to invade other markets) or vertical integrations (a stakeholder controls the whole supply chain in a given set of markets and blocks newcomers). Since the competitors on these markets do not have similar resources, they have no chance to compete. Google is an example of horizontal integration: building upon its dominant position on the search engine market, it captures the most of the revenue of online advertising and develops new services for other markets (emails, online office suites, etc.). Google also has a vertical integration strategy with *Android* and the default Google applications installed on every device using it. Another example of vertical integration is Cisco Systems: its develops protocols, routers and maintenance services for the communication market where it is a dominant stakeholder. In the past, similar situations have led to

governmental actions such as the breakup of AT&T in 1982 [24][37], or the obligation for Microsoft to offer *Windows* users the choice to use a browser other than *Windows Explorer* in 2009 [28]. In February 2014, the European Commission (EC) took a first step in this direction when it obtained that Google must guarantee "*that whenever it promotes its own specialised search services on its web page (e.g. for products, hotels, restaurants, etc.), the services of three rivals, selected through an objective method, will also be displayed in a way that is clearly visible to users and comparable to the way in which Google displays its own service*"[38]. Further similar actions must be taken. Regarding the position currently held by Google (and its similarity with that of AT&T before its breakup), one might wonder if separating Google's different activities (the search engine, *Android* and the other online services) could enhance competition. From the antitrust perspective, the main question is if Google has become an indispensable gateway to cyberspace [86]. The point here is not to punish Google for its success. In fact, nobody can deny the major innovations that Google has made available to the public (for example *Google Earth*) or its volunteer participation in many scientific projects (such as sharing its computing infrastructure with researchers). And other internet stakeholders benefit from some sort of power position on their market to limit competition (such as Cisco Systems cited above, or Apple, which excludes *Adobe Flash* — a technology used by many websites — from its devices or *AppGratis* from its mobile application store). I think that it is legitimate to study if some additional hard constraints on particular stakeholders (such as Google) can ensure broader competition and benefit the majority of internet users (also if they hinder these stakeholders).

A second issue is related to the organisation of online services. An online service is mostly viewed as a black box while it is technically divided into different components. For instance, a search engine may be seen as a vertical integration of three components: an indexing component that constitutes an index characterising online documents, a selection component that picks documents based on a given query, and a ranking component that orders the selected documents. In the offline world, to boost competition, regulators sometimes impose open access mechanisms. For example, rail transport is broken down into infrastructure and transport. In Europe, for a long time, one single (national) company managed both. During the 2000s, the EC forced the separation of these activities to allow different carriers to compete fairly for the transport of passengers and freight in a given infrastructure. Similar regulation mechanisms could be implemented online. POLLOCK proposed, for instance, to divide a search engine into two components (the software and the service) [81]. An extension of this proposal based on three component categories can describe each online service: the data (indexed Web pages, past purchases, etc.), the software working on that data (ranking algorithms, product suggestions, etc.) and the infrastructures that run that software (in particular the big data centres needed). To increase competition and promote innovation (for example, better ranking algorithms), some of these components could be available to different competitors, at least as long as legal constraints are respected (such as protecting privacy or classified material) since the constitutional and safety principles have priority. If such a proposal is adopted, a newcomer on the search engine market can, for instance, propose an innovative ranking algorithm (which is impossible today because of the investments needed to build an infrastructure and the time needed to gather enough past experiences to identify relevant documents). More generally, open access implies that the internet must remain an open and neutral platform. Experiences in internet mobile access indicate that it promotes innovations [5]. Currently, common protocols allow online service providers to propose integrated services to their costumers (such as email exchange between internet users with different providers). But sometimes, stakeholders do not respect this technological neutrality. In 2007, members of the UK Parliament complained because proprietary technologies used by Microsoft to design a governmental information portal made it difficult to access for citizens using browsers other than *Windows Explorer* [3]. Regulators should avoid this kind of situation.

Economic globalisation has always questioned company taxation, in particular through benefit transfer mechanisms. If states agree to avoid double taxation, many multinational companies seek to pay taxes which are too low or not to pay tax at all for some of their activities [78]. The internet reinforces this tension[39] and several problems specific to the online context have been pointed out: the possibility for a company to impact a national economy (through large volumes of online transactions with residents) without being physically established in the country that collects the taxes on profits; the at-

[37] It should be noted that AT&T has reconstituted much of its former position.

[38] `http://tinyurl.com/m6crmf4`, European Commission website (consulted on February 12, 2014).

[39] Printed books are good examples, since the value-added tax (VAT) varies from one country to another. It is 6% in Belgium, 5,5% in France and 3% in Luxembourg. Since Amazon sells its books for Europe from Luxembourg, it benefits from a reduced VAT compared to Belgian and French booksellers, which weakens the position of booksellers.

tribution of the value generated by geolocalised data through online services; and the qualification of benefits generated from new economic models [79]. A serious discussion of this problem falls outside the scope of this essay. However, some general rules can be established. First, the states of all the parties involved in an online transaction should tax part of the benefits generated by it. This is legitimate since the existence of an online transaction can generate public spending in every state (for example, if legal action is launched in its wake). Of course, to avoid a bureaucratic burden, only companies with a minimum volume of online transactions should be targeted. Second, the partition of the benefits generated by an online transaction must take into account the asymmetry of the costs for the producers (which may be null). For example, if an American buys a CD from the online website of a British rock band, all the costs of the operator may be localised in the UK (hosting the website, production, shipment, etc.). In this case, most benefits must be "assigned" to the UK where it will be taxed for the most part. These rules are not specific to the online context: companies should never be allowed to take advantage of fiscal mechanisms to avoid legitimate taxation. But the immaterial character of most commodities generating online revenue makes financial regulation more complicated.

While the first three issues may sound controversial and require a political and economic debate, the fourth one is not. It concerns the ease with which an internet user can switch from one service provider to another. The mobile segment of the telecommunications market shows that it is a crucial element in order to ensure competition. Currently, nothing exists that facilitates such a change. For example, if an internet user wishes to change his or her email provider, it will be difficult for him or her to transfer past stored emails in ideal conditions[40]. This becomes even more difficult for non standardised services (such as switching from a blog hosting provider) where internet users face real "lock-ins". While this issue can be seen as a particular case of data ownership as covered by the safety principle (what can internet users do with "their" data), an economic perspective may be useful to treat it in a less controversial context. As for other sectors (such as people who change their mobile operator while keeping the same phone number), regulators must force providers not to prevent their customers from switching (for example by proposing an export feature based on standardised data formats).

4.4 Internet regulation authorities

I have argued that the proposed extended regulatory framework presupposes new attributions for internet regulation authorities (already existing or new ones to be created). This raises new issues such as the jurisdiction of these regulation authorities, their attributions, their status and their organisation. All these questions should be part of a public discussion to ensure that new regulation mechanisms and the corresponding regulation authorities protect citizens rather than threaten them. I shall discuss some elements of this debate here.

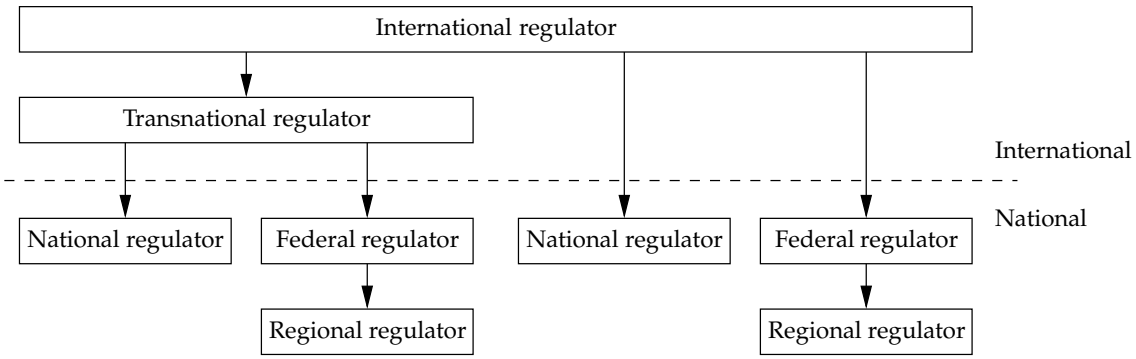

Figure 2: Regulators.

The question of jurisdiction, already mentioned in subsection 4, is the first one that comes to mind. The differences between political regimes (democracies, authoritarian regimes, etc.), legal frameworks (for example, in some federal states, each subdivision has its own constitution) and cultural conceptions (such as free speech) require a layered approach to be adopted (Figure 2). So, in practice, there must be

[40]Internet users can send all the past emails to the new address, but they may lose important management information such as folders, original sending and receiving dates, filters, etc.

a mix of international and national regulators. International regulators should be the place where states try to negotiate intergovernmental treaties on globally accepted online threats (child pornography, cyber criminality, etc.). Proposed regulation mechanisms should first be discussed within national regulators and a consensus must be reached internationally. Of course, such an adoption process rarely leads to concrete decisions. But since international cooperation is needed to fight global threats, it is more efficient to adopt minimalist but effective international agreements rather than running behind a hypothetical global treaty. Could the International Telecommunication Union (ITU), an agency of the UN, be such a global regulator? Many researchers disagree because the ITU does not include all stakeholders of the internet, and because governments have too much power. In fact, membership of the global regulator should be limited to national regulators (rather than national governments as is the case with the ITU). If these national regulators, at least some of them (in particular those in democratic countries), are organised to take into account all national stakeholders (as discussed below), the global regulator cannot be exploited by those who want to restrict online freedoms everywhere and always. One touchy question is to decide who will be responsible for managing the IP addresses and the domain names. Currently, it is a US non-profit private organisation, the ICANN, of which the US Department of Commerce (DoC) has a final and unilateral oversight over some of its operations [6]. Some argue that the ICANN functions must be taken over by a global regulator, but the US governments defend the current situation so far [75]. As already mentioned, on March 14, 2014, the DoC "*announced its intent to transition key internet domain name functions to the global multistakeholder community*"[41]. But no information was given as to how this would be achieved (nature of the regulatory entity, number of regulatory entities, etc.). Transnational regulators can be another layer at international level for countries participating in supranational institutions. The EU is a good example, since it provides several directives related to the internet (privacy, cyber criminality, etc.) that member states must respect[42]. In such cases, it makes sense that a common transnational regulator exists. But transnational regulation mechanisms will probably also remain minimalist due to the lack of agreements. Even in the EU, for example, countries may have different views of a same legal framework. The less binding Irish interpretation of the EU directive on privacy is certainly one of the elements that motivate US providers to settle in Ireland (or is it the opposite?). National regulators will therefore retain a significant role. Depending on the state organisation, a national regulator could be a combination of a federal regulator and regional regulators. Let us take Germany as an example. Privacy is a competence shared by the Länder (each one has its own constitution) and the federal state, so regulators have to be federal and regional.

Bilateral treaties should be another attempt to build global regulation mechanisms. One example could be the protocol *International Safe Harbour Privacy Principles* signed by the EU and the US in 2000 to ensure that EU privacy rules are taken into account by US stakeholders when they handle the personal data of EU citizens. It specifies that "*self-certifying to the U.S.-EU Safe Harbor Framework will ensure that EU organisations know that your organisation provides 'adequate' privacy protection, as defined by the Directive*"[43]. But one problem with international treaties is that verifying that all signatories are willing to respect them scrupulously can sometimes be difficult. In fact, there is lack of US companies that follow the *International Safe Harbour Privacy Principles,* making it a poor regulation mechanism for EU citizens to protect their personal data overseas [89].

A wide spectrum exists for the status of local regulators that depends on the degree of independence vis-à-vis political authorities. In authoritarian regimes, regulators depend directly and uniquely on the political authorities. Mixing sharp control and online propaganda, this sort of "ministry of the internet" scares every proponent of democracy and explains largely why many people defend minimalist regulation mechanisms. But regulators can also be independent of political authorities. Well-known examples are public independent regulatory agencies whose role is defined by law but which act without political interferences (for example those which regulate the food and pharmaceutical industries). Normally, transparency (in particular in their composition), publicity of their decisions (including the motivations behind them) and consultation of stakeholders are some of their characteristics. The Commission nationale de l'informatique et des libertés (CNIL), which regulates data privacy in France, illustrates such a public independent regulatory agency. But independent regulators can also emerge from private initiatives. The Internet Watch Foundation (IWF), a British charity, exemplifies such a case. Its mission "*is*

[41] `http://tinyurl.com/pyeq4mw`, official website of the US National Telecommunications and Information Administration (NTIA) (consulted on July 25, 2014).

[42] In fact, member states have an obligation to transpose EU directives into national laws.

[43] `http://tinyurl.com/5t7y7d3`, website of the protocol (consulted on February 13, 2014).

the elimination of child sexual abuse images online"[44], i.e. it acts as a regulator of online content. Since British access operators (such as British Telecom) agree to block websites listed by the IFW, it becomes a *de facto* online censor [26]. In 2008, the IWF was involved in an overblocking process of Wikipedia, which called into question the regulatory role of a private stakeholder [75]. In fact, such private regulators often suffer from drawbacks such as a lack of technical competence, overreactions (which may lead to a "dictatorship of opinions"), promotion of certain ideologies, lack of control and lack of transparency. So, public independent regulatory agencies seem to be the best candidate as regulatory authorities so far as some rules are respected.

A first rule is that the composition of a public independent regulatory agency must ensure its independence, not only towards political authorities but also towards particular stakeholders (in particular, commercial companies or minority ideological groups). In the US, the Federal Communications Commission (FCC), which regulates the US telecommunication sector (one aspect of internet regulation), "*is directed by five commissioners appointed by the US president and confirmed by the US Senate*"[45]. France chooses a broader representation for the CNIL which is directed by a multipartite college of 17 people composed of "*4 parliamentarians* [...], *2 members of the Economic, Social and Environmental Council, 6 representatives of the supreme courts* [...] [and] *5 qualified board members appointed by*" the public authorities[46]. Ideally, the *board of trustees* of public independent regulatory agencies should extend CNIL's openness to other stakeholders of the internet. They must include representatives from internet users (appointed by civil society organisations such as a consumer organisation) and academic (appointed by national or regional scientific organisations). This is necessary not only to ensure the required independence with respect to governments and industrial interests, but also to guarantee a "democratic regulation" of the internet (through the representation of citizens). An open question is the presence of representatives of economic stakeholders who could be appointed by employers' organisations. Proponents could argue that economic stakeholders provide useful information about technical and practical constraints when discussing regulation mechanisms. Opponents could report that economic stakeholders are judge and jury if they participate in the regulation mechanisms applied to them. Moreover, they fear that, in practice, only big commercial companies will be represented. Anyway, goings and comings between independent regulatory agencies and those who are regulated (except the representatives of the latter, if any) must be avoided, and all appointed members should first disclose potential conflicts of interests (funding sources, stocks of a company active online, etc.). In particular, former members of the regulators must be banned from becoming lobbyists for the regulated (or vice versa).

A second rule is to guarantee the transparency of the appointment process and the publicity of the debates. Since the topics managed by public independent regulatory agencies (infrastructure access, privacy, etc.) affect large industrial interests, lobbying is pervasive. It is therefore essential that each stakeholder must access all the information involved in a given decision (legal arguments, reports, scientific references, experimentations, etc.).

A third rule is to precisely define the scope of a particular public independent regulatory agency. One could argue that a single agency should be responsible for all matters affecting the internet. But several countries already have different regulators with attributions linked to the three regulation principles (mainly concerning communication infrastructures, privacy and anti-competitive behaviour). So, in practice, it is perhaps more efficient to build upon existing public independent regulatory agencies rather than to create new ones. In fact, it makes sense to dedicate a public independent regulatory agency to a specific topic, whether it concerns online or offline situations (privacy is the best example). Moreover, the layered approach of regulation described previously imposes the coexistence of multiple independent regulatory agencies ruling the internet. This should not be a problem as long as the design of regulation mechanisms follows the "matryoshka principle": the more local the regulator, the more restrictive the mechanisms (as long as the general principles are respected, in particular the constitutional ones). Privacy illustrates this principle. Globally, no "real" agreement can be reached due to several states that want to control their citizens. Some transnational agreements can exist, such as in the EU. Next, a country like Germany, which must comply with the corresponding EU directive, can (for historical reasons) implement more protection for their citizens. Finally, the German Land of Schleswig-Holstein can be more protective than any other Land. When a given public independent regulatory agency has attributions covering several topics, it can be composed of several *policy committees* appoin-

[44] `http://tinyurl.com/pxp9bgk`, website of the IWF (consulted on February 28, 2014).
[45] `http://www.fcc.gov/leadership`, website of the FCC (consulted on February 19, 2014).
[46] `http://tinyurl.com/l47v9u4`, website of the CNIL (consulted on February 19, 2014).

ted according to the first rule discussed above, with each policy committee in charge of one particular topic. Policy committees are controlled by the board of trustees (upper part of Figure 3).

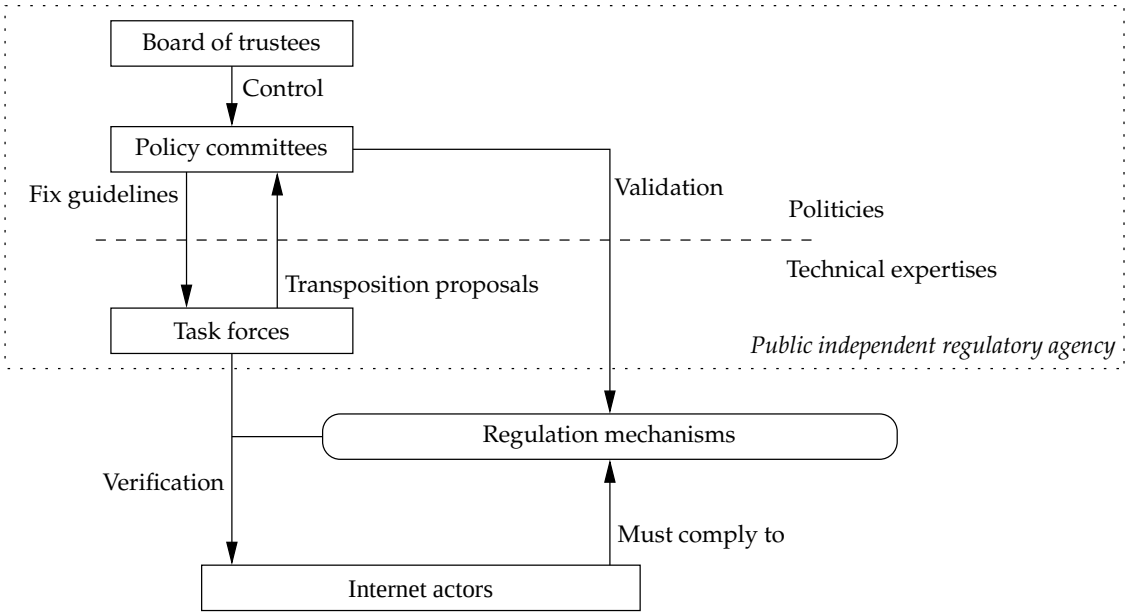

Figure 3: Public independent regulatory agency.

But public independent regulatory agencies cannot be made only from policy committees, it is also necessary that they have (technical) *task forces* (Figure 3). Technical expertise (legal experts, computer scientists, etc.) is necessary to conceive practical regulation mechanisms that find the right balance between the stalking of possible online abuses and the protection of the freedom to use the internet (including proposing innovative immaterial commodities). While the committees set the policy guidelines and the general regulation rules (for example the complexity of security systems depends on the degree of privacy of data managed), experts propose to transpose this into practical regulation mechanisms (such as setting security standards to implement). Such practical regulation mechanisms could include certification processes (for example for network equipment such as routers). It is important that regulation mechanisms take the context into account. Let us suppose that a company provides a new online service that helps internet users to track their health. When launched, this new service would probably be used by well-informed early adopters only and be limited to a few statistics. In such a context, the company should not be buried by regulation mechanisms. But, when the service develops, it is crucial to ensure that "normal internet users" are well informed, that privacy is scrupulously respected and that the company deploys state-of-the-art protection mechanisms for the data. Same company and same service, but different contexts. Of course, proposed practical regulation mechanisms must be validated by the corresponding policy committee before they become binding. Task forces are also needed to provide information (for the policy committees and for the public) about existing practices, policies, research issues and technical innovations. They could, for instance, run simulations to assess potential implications of regulation mechanisms discussed or to evaluate the actual extent of existing discriminations [5].

Another rule is to give to these public independent regulatory agencies the necessary resources and attributions, not only to build regulation mechanisms but also to verify that the various stakeholders respect them. This rule has two implications. First, provide funds to task forces, in particular if they must be composed of highly educated full-time employees (as is the case for agencies in charge of studying patent applications). Second, in order to verify that the defined regulation mechanisms are respected, public independent regulatory agencies must perform investigations or audits. Regulators make audits in other domains, such as to ensure food security: inspectors control restaurants without warning to verify that they respect the standards of hygiene and food preservation. Task forces must therefore be able to make unannounced inspections of some internet stakeholders (such as verifying that the filtering methods used by an access operator do not overblock). This is necessary to ensure that the regulation

process does not become a bureaucratic system. Indeed, investigations and audits avoid the need to systematically verify everything before it goes online, which would hinder innovation. In practice, it is more efficient to stop or block an ongoing service (eventually by forcing the providers to destroy data) once a problem is detected, rather than to systematically check that an online service (a new one or an upgrade of an existing one) verifies all the regulation mechanisms. Of course, policy committees can specify contexts where such a systematic check is required (for example when health data are involved).

A last important rule is the necessary relations between these public independent regulatory agencies, the judiciary power and state security stakeholders (criminal investigative and intelligence-gathering). Once a public independent regulatory agency discovers a practice of an internet stakeholder that does not respect the existing law and/or regulation mechanisms, it should take all the technical actions to stop it (such as asking network operators to block an IP address used to launch cyber attacks on strategic servers). But this is not enough. The public independent regulatory agency must pass the case to the corresponding inquisitorial system that must initiate an investigation. This may seem obvious at first glance, but in practice this is not always the case. Even for very serious offences, such as child pornography, regulators often limit their actions to blocking websites without launching a legal proceeding, while several websites are hosted in Europe or in the US [95]. We also see that many countries are sometimes reluctant to apply strong antitrust actions against big companies. This must change since regulatory mechanisms become inefficient if those who violate them remain largely unpunished. A zero tolerance policy is impracticable and the focus of regulation mechanisms and investigations must be on practices that hurt many people (for example breaches of privacy) or someone severely (such as child pornography). In practice, however, public independent regulatory agencies should be authorised to fine some stakeholders or order certain conservative measures in specific situations. For example, if a website hosting a large amount of personal data did not provide enough security measures, a public independent regulatory agency could decide to shut it down temporarily. Finally, public independent regulatory agencies should also help external oversight of how technologies are used by state security stakeholders. Every democracy has some kind of political control over these stakeholders. In some countries, they report directly to the government (such as in the US). In other countries, it is a committee of members of the Parliament that controls these stakeholders (for example in Belgium). In any case, public independent regulatory agencies must provide the technical skills needed to perform this control. In particular, it is crucial to assess that technical mechanisms deployed by the judiciary power and the state security stakeholders are proportional to the scope of the legitimate activities of the latter. For instance, if electronic surveillance of one individual can be legitimate with a corresponding warrant from a judge, the technical mechanisms should not store information regarding other people linked to this individual (such as a systematic surveillance of all of the person's contacts).

The question remains open whenever public independent regulatory agencies are involved in cyber defence. Cyber defence implies the protection of online resources, not only those managed by the states (such as governmental servers) but also those of the private sector (non-governmental organisations, private companies, etc.). If the protection of state online resources are the responsibility of state security stakeholders, it is not clear how the security of the private sector must be organised. This is an important issue since it is well known that many companies do not invest enough to be adequately protected against cyber attacks [57], including those which participate in strategic projects. Public independent regulatory agencies could first help the main stakeholders of the private sector to secure themselves (information, audits, technical tools, etc.). But they could also be responsible for verifying that the most sensible stakeholders take the necessary protection measures (hospitals, energy corporations, etc.).

Currently, no existing regulator or group of regulators has the characteristics described here: the openness of its political direction to ensure true democratic regulation, extended attributions to propose new regulation mechanisms and verify their applications, and the budgets to constitute the required team of experts to fulfil these new missions. Moreover, in the current economic context, most states are looking to cut spending rather than take on new costs. However, democracies have to make the necessary efforts to build the adequate internet regulation authorities to develop a more neutral network.

Throughout the rest of this essay, the term "regulators" refers explicitly to public independent regulatory agencies organised as discussed above.

5 Some internet regulation mechanisms

Once an extension of the internet regulation scope is accepted, it is necessary to study how to implement regulation mechanisms that respect the three principles proposed; in other words, which discriminations performed by internet stakeholders are acceptable and which must be regulated. This section discusses some regulation issues. I propose to address them following the three components of the internet: the communication infrastructures, the online contents, and the online services managing these contents. I finish this discussion with some cross-layer issues. My aim is to show how the three regulation principles presented above can guide the design of online regulation mechanisms and help to find the right balance between the liberty of internet stakeholders to act online while ensuring fair use for everyone. Some of the solutions proposed are easy to implement and others need a strong political will.

5.1 Communication infrastructure regulation

The communication infrastructures ensure that nodes communicate through the internet network by transmitting data through different physical telecommunication channels. The infrastructures must meet the bandwidth requirements of today's internet usages (such as video streaming) and be able to integrate future technologies (quantum cryptography, for instance, requires each node to have a "quantum repeater"). Currently, most network operators are commercial companies[47]: they invest in their networks and charge their clients for the use[48].

The central role of communication infrastructures explains partially why the debate on network neutrality has been focussed on network operators. I will first introduce some technical aspects before discussing several issues.

5.1.1 Technical insides

As explained in subsection 2.1, the communication of data on the internet is based on the transmission of datagrams. In practice, a datagram is composed of a data payload (a fraction of the actual data to transmit), and a header[49] that provides additional information needed to manage it. This header contains the IP addresses of the origin and destination nodes and is used to route the corresponding datagram through the networks. It also carries another piece of information, i.e. the *port number*[50], that distinguishes the type of data conveyed by a given flow of datagrams. For example, the port number 80 is conventionally associated with Web browsing. The port number allows the destination node to determine which application must manage the received data (such as a Web browser on a computer for the port number 80). From the traditional interpretation of the network neutrality concept, network operators must propose a best-effort delivery, i.e. no datagram should be discriminated against or favoured.

Several discrimination technologies exist. Deep packet inspections are filtering techniques that can discriminate a datagram based on its data payload and/or based on its header (for example, to block all datagrams emitted by a particular IP address). Stateful packet inspections are filtering techniques that can act based on a "global view" of a communication process (such as taking the state of connections into account). These filtering techniques work with discrimination rules (for example to block all datagrams emitted by a given IP address and which contain a given sequence of data). But determining the correct set of rules is difficult because new cyber attacks emerge every day. Moreover, some rules can block "legitimate" datagrams (overblocking). Firewalls are software or hardware-based security systems that implement discrimination technologies. Firewalls are implemented all over the internet (to protect the internet or the organisation's private networks).

Since the size of the internet network prohibits the computation of an optimal routing path for all datagrams, local heuristic approaches select the next node used to propagate a datagram. As explained, these heuristics discriminate by nature some datagrams, but this cannot be described as a deliberate discrimination: it is at most a technical bias (subsection 3.1).

[47] In some countries, public companies manage the networks or networks dedicated to science and education and are funded by public authorities.

[48] Some clients share their connections with their own clients and charge them accordingly.

[49] From a purely technical perspective, a datagram has multiple headers, with each layer of the TCP/IP model (IP, TCP and applications protocols) adding its own. A node that implements only the IP "understands" only the "IP header". But to simplify the discussion, I will assume that all these headers form a single one.

[50] The port number is information added by protocols just above IP (mostly TCP and UDP).

It should be noted that, since the beginning, the TCP/IP protocols propose mechanisms that allow a priority level for datagrams to be specified to intermediate nodes [55]. The first standard of the TCP/IP protocols (so called TCP/IP v4) provides additional "type of service" information, and the latest one developed in 1998 (so called TCP/IP v6) introduces a mechanism to treat some datagrams in real time. But these mechanisms, imagined before the network neutrality debate, have not yet been implemented.

5.1.2 Security issues

Everybody agrees that network operators should be allowed to discriminate against some data transmitted to protect their networks and their clients (safety principle) [110]. Examples include filtering packets emitted by computers used for distributed denial-of-service attacks[51] or blocking viruses and malware. Network operators can use discrimination technologies to decide to drop a datagram to prevent damages further down the line. Some IP addresses can be temporarily banned (for example blocking datagrams coming from non-friendly countries and targeting very sensitive servers), datagrams containing viruses, child pornography and malware can be dropped, or illegitimate connections blocked. As already mentioned, overblocking is a recurrent problem (subsection 3.2). It occurs when network operators use discrimination technologies which are too restrictive (mostly rules which are too general) when they want to legitimately block some datagrams. It can be seen as a collateral damage of the protection of networks and internet users. If network operators cannot, in practice, ensure that overblocking never happens, they should target zero overblocking. But, as already said, this is not always the case in practice [69, 75].

From the perspective of network neutrality, these discriminations are pointed out as acceptable in a traditional interpretation [110]. But many of its advocates are wary of inappropriate uses, in particular blocking datagrams for political or commercial reasons only. There is lack of empirical studies demonstrating that, in democracies, network operators regularly and willingly discriminate against contents or applications (exceptions are discussed below) [15][52]. It is the role of regulators to verify that discrimination is based on security and global management rules only (and thus do not produce any continuous overblocking). As mentioned in subsection 4.4, these regulators could make unannounced inspections of network operators, certify software and hardware that implements discrimination technologies, or certify those which install these software and hardware.

Cyber surveillance can be seen as a particular aspect of security. In fact, network operators, in particular access operators, can easily spy on the online activities of users if they gather all the corresponding datagrams transmitted through their networks. But I will postpone the debate on cyber surveillance until I have discussed the broader topic of online privacy (subsection 5.4.2).

5.1.3 Price discrimination against providers

Some network operators want to go further in discriminating against datagrams. They say that the internet stakeholders whose revenue (mostly online advertising) depends directly on rapid access to their online immaterial commodities should be forced to contribute more to global infrastructures. They find "*complaints about '5% of the users generating 50% of the traffic'*" [77], and argue that additional revenue is necessary to ensure innovation at the core of the network. They therefore defend price discrimination against providers. Some countries have adopted a similar idea for their road transport by adopting paid access to their motorways: lorries (big internet companies), which damage (gill) the motorways (infrastructures) and transport most goods (data), pay more than normal cars (internet users). This idea violates the traditional interpretation of the network neutrality concept [110], which explains partly why providers (such as Google) defend this traditional interpretation [99].

This is a hot topic, not only because network operators are asking for it, but also because of a decision by an American court in January 2014[53]. The court has ruled that the United States (US) Federal Communications Commission (FCC) cannot force network operators to respect two important constraints of the traditional interpretation of network neutrality (no blocking and no unreasonable discrimination)

[51] The principle of a denial-of-service attack is to clog access to a server with millions of unnecessary connections in order to prevent legitimate ones to pass.

[52] Of course, in authoritarian regimes, access operators are controlled by the government and systematically block many websites.

[53] `http://tinyurl.com/ktwzq8f`, decision of the United States Court of Appeals for the District of Columbia Circuit (consulted on January 28, 2014).

without considering them as common carriers. Since the FCC does not classify them as such, it cannot impose these constraints. It is therefore crucial that this question is being discussed and that some regulation mechanisms are implemented.

Many researchers have studied possible price discrimination. From an economic perspective, the communication market is a two-sided market: intermediates (network operators) offer connexion services to two distinct groups of clients (the providers and the internet users) [92]. As shown by the mobile telecommunication market, multiple providers should ensure real competition and should limit the risk of discrimination against clients [31]. While real competition exists in the *last mile* in many areas (but a geographical digital divide remains, as pointed out in subsection 3.2), the internet backbones are operated by a limited number of stakeholders. Many fear that these stakeholders form an oligopoly and take advantage of their position to set high fees. This limits innovation at the edge of the network [61]. Since some economic models indicate that the commercial viability of many providers is only possible with a mass of online customers [92], opponents of price discrimination fear that startup companies can be clamped by fees which are too high. Many therefore defend a *zero-price regulation* that prohibits network operators from charging providers. Some researchers recognise that a zero-price regulation solves possible discrimination problems, but feel that specific regulation mechanisms to manage these problems may be more efficient (in particular antitrust measures against network operators exploiting their dominant positions) [43]. In fact, some economic models tend to show that some form of price discrimination can be positive overall [92, 55]. Currently, there is a lack of empirical studies that evaluate the real impacts (positive or negative) of price discrimination [15, 31].

Two assumptions may justify price discrimination [55]: the "*internet traffic will increase at a rate which cannot be handled by the current technology and traffic management techniques*", and the impossibility for network operators to assume "*the costs for the necessary network infrastructure investments without tapping additional revenue streams*". To the best of my knowledge, no study validates or invalidates these assumptions. Some elements are in favour of the validity of the first assumption. The expansion of current practices (cloud computing, internet access with mobile devices, online video streaming, massive open online course, etc.) could need network bandwidths beyond those currently available. But new technologies (such as better data compression methods) or new traffic management techniques can decrease the congestion of the communication infrastructures. For example, maintaining local copies of highly consulted online contents (such as video streaming) avoids unnecessary multiple transportation of the same data over the whole internet. Concerning the second assumption, no evaluation exists for the costs of the future communication infrastructures. While the costs for deploying the current infrastructures decrease, new ones would probably be deployed in future (not only to increase the available bandwidth but also to provide new services such as enhanced cryptography). Moreover, it is not proven that revenue generated by the current business model (based mainly on subscription fees from internet users) is insufficient for these investments [77]. Two approaches are therefore possible. The first one postulates that at least one of these assumptions will never be true. In this case, the zero-price regulation must be maintained. The second approach is to allow some price discrimination by network operators as long as both assumptions are not invalidated[54]. The hope is that this new revenue will improve the quality of service of the communication infrastructures. The regulators then have a crucial role: to ensure that no monopolistic practices occur (competition principle).

I would favour the second approach if the regulators are built according to the model of the public independent regulatory agencies described in subsection 4.4. If the idea of some sort of price discrimination is accepted, its implementation deserves to be discussed. Foremost, a separation must be made between the *last mile* and the internet backbones. The first must be financed by the internet users only (the next subsection tackles this), while the latter may concern the providers. A comparison with road transport may be made once again: roads in cities are financed by local taxes while the motorways are managed by global authorities (at least regional ones).

If price discrimination presupposes that providers have to pay extra fees to network operators to transport their data, various implementations exist. Network operators can discriminate more or less against datagrams of every provider depending on a subscription fee: only the providers that paid for "premium" access will have their data transported at the current best effort. This will introduce a financial bias, discriminate against the small providers that cannot afford it, limit innovations at the edges and probably favour big providers that already have a dominant position. This is unacceptable (competition

[54] Regulators can make or coordinate research that studies these assumptions.

principle). Another solution is to propose a "fast lane" to some providers: network operators favour datagrams for some providers by transporting their data on a specific part of their communication infrastructures. The deal between the American network operator Comcast Corporation and the online video streaming provider Netflix reached on February 23, 2014 illustrates this solution [111]. This approach ensures that all providers benefit from the current best effort, some of them (probably again the big ones) will see a boost in the transportation of their data. This advantage can be lowered if the "fast lane" is limited to specific applications vital for established providers (such as the replication of distant servers). But, in addition to introducing a financial bias, in the medium term, if the second assumption is true, this solution equals the first one. Indeed, if the internet traffic increases heavily, the "fast lane" will become the only way to propose decent access to online immaterial commodities, and the small providers (especially the young ones generating innovations) will again be completely discriminated against. So this second solution is also unacceptable (competition principle). The best solution is perhaps to put a financial charge on the providers that generate the largest part of the online traffic[55]. It has the advantage of protecting the small and young providers and to redistribute some of the profits generated by the highest layer (the most profitable one) to all the stakeholders of the layers below (where many investments must be made). Such a financial charge is not incompatible with the competition principle as such. This financial charge could be additional fees paid to network operators or a new tax collected by governments[56]. In this latest case, such a *bandwidth added tax* can then be entirely allocated to innovative projects such as financing new infrastructures and research grants[57]. One could argue that such a tax allows some governments to discriminate against providers for arbitrary reasons (such as to be non loyal). But authoritative regimes discriminate already, and many constitutional rights (free speech, freedom of assembly, freedom of the press, etc.) prevent democratic governments from acting that way. In any case, the regulators must ensure that this financial charge remains legitimate (constitutional and safety principles) and reasonable (competition principle).

Of course, such a financial charge is only acceptable if certain guarantees exist. First, transparency must be absolutely necessary. As explained, nothing but research and development investments or online public politics[58] justify a new financial charge for some providers. In particular, the corresponding windfall cannot serve purposes other than these investments (such as additional dividends to financial shareholders). In the case of additional fees, regulators may require some obligations for network operators (such as providing a minimum service quality for every customer). Second, this financial charge must be proportional. It seems natural that, among the charged providers, those who contribute most to internet traffic are charged more (like the polluter pays principle in environmental law)[59].

Price discrimination against providers poses a problem for non-commercial companies. *Wikipedia* is, for instance, one of the most visited websites (and thus a big generator of online traffic) without generating revenue. The Wikimedia Foundation (which supports and hosts *Wikipedia*) and similar providers must be protected from financial charges that endanger their existence, in particular if their online immaterial commodities are of public interest (safety principle). A solution can be found around the concept of public interest organisations recognised by many countries. When high online traffic is generated by such public interest organisations only, they are exempt from any additional contributions while the others are not. In order to avoid an administrative burnout for these organisations, transnational agreements can be reached to extend the status of public interest granted by an organisation in one region (for example by the US) to other countries (such as in Europe). Regulators can always challenge this recognition locally in court.

[55] An approach is to work with the 50-quantiles of the total internet traffic. In practice, all the providers are ranked in descending order of their internet traffic. This series is divided into 50 groups, called quantiles, with each group corresponding to the same proportion of the traffic. The top quantiles contain a few providers generating a lot of traffic, while the bottom quantiles contain a lot of providers generating a small amount of traffic. Only the providers appearing in the first two or three quantiles are charged.

[56] In both cases, enforcement actions can be taken by network operators (possibly ordered by regulators in the case of a tax) against "big" providers that refuse to pay (e.g. for discriminating against some of their datagrams).

[57] There are several examples of taxes allocated to specific purposes. For example, France collects a contribution sociale généralisée ("general social contribution") which contributes to the financing of social security.

[58] Of course, public politics in the context of the internet must be understood in a broader sense. For example, financing educational programmes related to the internet or ensuring universal internet access are legitimate public policies.

[59] If a 50-quantile approach is used, different rates can be set for the first two or three quantiles corresponding to the providers to charge.

5.1.4 Price discrimination against internet users

The business model of commercial network operators in the *last mile* presupposes that only clients subscribing to them use their communication infrastructures. Historically, some price discrimination against internet users has always existed, since multiple subscription schemes provided different connection qualities. An investment in an infrastructure of a given bandwidth is therefore mostly done if there is a potential market for it. In other words, no reliable infrastructure will be built without a sufficient market, and the access to existing infrastructures is conditioned by the capacity to subscribe.

This business model is incompatible with an interpretation of the constitutional principle that guarantees "universal access" to each citizen with a minimum connection quality. Where this interpretation prevails, regulation mechanisms must ensure a complete cover of the area and a subscription aid when necessary. For the coverage problem, one could argue that public authorities can fund the required communication infrastructures. But in these difficult economic times, it would be absurd if the public sector paid for the less profitable infrastructures while leaving the most interesting ones to commercial companies. Accordingly, two solutions exist. The first presupposes that states participate directly either by funding high-quality connexions or by creating a fiscal policy to reimburse households with the lowest wages for their internet costs. This can be financed by new state revenue (such as the bandwidth added tax presented above or a valued added tax on high-quality access subscriptions). A second approach consists in associating some conditions to broadcast licences (not limited to radio broadcasting), such as specifying which area should be covered by any operator. The network operators are thus partly considered as common carriers. The question of "universal access" will become increasingly important since we know that the *last mile* infrastructures probably have to be updated (in particular for mobile devices).

As previously explained, several countries require that network operators provide all the necessary information to evaluate the connection qualities they propose [30]. This information should not be limited to tariffs and to data maximum quantity and bandwidth downloaded and uploaded. In the European Union (EU), for example, EU directive 2009/136/EC specifies that it should also include information such as a description of the discrimination mechanisms used for security reasons or the way network operators handle privacy [44]. Regulators should force all network operators to provide the necessary information and present it clearly (safety principle).

5.1.5 Data discrimination

As explained in subsection 5.1.1, datagram discrimination cannot only be based on the emission and destination IP addresses (i.e. provider and internet user discrimination discussed above), but also on the data payload transmitted. Network operators have legitimate and illegitimate reasons to discriminate some online immaterial commodities. One legitimate reason is the security issues described in subsection 5.1.2. I will consider some other reasons next and examine their legitimacy.

Everybody knows that illegal contents are transmitted over the internet (child pornography, copyright protected materials, etc.), and that technologies exist[60] to identify some of these contents. It therefore seems legitimate to block the data related to these contents. But the safety and competition principles require that any overblocking approach must be banned. The abovementioned blocking of the entire *Wikipedia* website in December 2008 because of a particular image hosted (the cover of the 1976 album *Virgin Killer* by Scorpions) illustrates what must be avoided [75]. More recently, one of the weaknesses of the Stop Online Piracy Act (SOPA) and of the PROTECT IP Act (PIPA) was precisely that network operators were forced to block entire websites that contain some copyright protected materials [112].

Sometimes legitimate reasons are used as an excuse to discriminate illegitimately against certain data. The case of peer-to-peer (P2P) architectures, mostly presented as responsible for the highest traffic, is exemplary. Some network operators want to block datagrams corresponding to P2P architectures for content distribution (such as *BitTorrent*) arguing that they are used to share copyright protected materials. But there are also legal uses of P2P architectures (several *GNU/Linux* distributions can be downloaded this way), including integrating them among data centres to reduce bandwidth costs when vast

[60] These technologies have built some characteristics of the data transmitted to create a sort of fingerprint and use it to categorise the data and/or compare them with a database. The mobile application *Shazam*, for example, is based on acoustic fingerprints to recognise a song through a smartphone.

amount of data must be delivered [64]. Moreover, online streaming becomes the biggest bandwidth consumer [60]. In reality, the goal of providers is probably to limit the traffic on their current infrastructures and to delay investments in new ones. P2P filtering should therefore be banned (competition principle). In 2007, for instance, it was shown that the American network operator Comcast Corporation degraded the traffic generated by *BitTorrent*. After a decision of the FCC [23], Comcast Corporation stopped this practice. That said, regulators should accept technical solutions that limit the bandwidth impact of P2P architectures without discriminating against their users[61] (safety principle).

In addition to limiting traffic, network operators sometimes have other motivations for illegitimate data discriminations. In 2005, it was established that the American network operator Madison River Communications was blocking the traffic associated with the VoIP[62] application *Vonage* seen as a competitor of its own telephony service [106]. Again, after a decision by the FCC, Madison River Communications was ordered to stop blocking *Vonage* and to pay a $15,000 fine. The competition principle must ensure that no online immaterial commodity is discriminated against by network operators, either to favour another one or because it generates network management difficulties (except of course security issues).

What is more, some engineers argue that it makes no sense to always treat each datagram equally [88]. They suggest that, when the network is congested, the priority for network equipment (such as routers) must be to drop datagrams that are related to non reactivity dependent applications. For example, datagrams associated with emails (where a delay of a few minutes is not a problem) will be dropped more than those related to VoIP (where fluid data flow is necessary for good communication). Since this approach does not discriminate against a particular stakeholder (specific IP addresses) or a particular content (a voice communication criticising the government with respect to another one), but relies on the application type only, it cannot be seen as a social or political control mechanism. In other words, this approach does not violate any of the three regulation principles presented. In fact, while protecting the most reactivity dependent applications, it can be seen as an application of the safety principle. Moreover, it would increase economic efficiency and consumer welfare.

While some existing or new data discriminations are legitimate, many are not. It is therefore necessary that regulators control how discrimination technologies are implemented. This can be done with a certification process for the network equipment used by network operators to build their infrastructures. Since a limited set of companies build this network equipment (in particular Cisco Systems), it is more efficient to control these companies than all the existing network operators. However, regulators may also conduct audits of network operators to ensure that no illegitimate discrimination technologies are implemented.

5.1.6 Infrastructure regulation

So far, I have discussed the regulation of the communication infrastructures from the perspective of the network operators. I want to address briefly now two more particular aspects.

First, among the different telecommunication channels, electromagnetic radiations have a particularity: it is a shared channel. While each network operator can build its own optical fibre infrastructures, those which use electromagnetic radiation (mostly through mobile telecommunication and Wi-Fi) share frequency bands. For short distance channels (such as Wi-Fi or Bluetooth), no regulation is needed and everybody can create their own infrastructure[63]. But for "middle or long" distances, national regulators set the available frequency bands (spectrum), and those that want to use them must generally buy some sort of licence. In practice, "*the vast majority of the spectrum remains encumbered by use restrictions that limit the technologies that can be deployed in any particular band*" [112][64]. This can be especially limiting for wireless mesh networks (WMNs). To put it simply, WMNs are wireless decentralised infrastructures formed by the peer-to-peer interconnection of devices with wireless antennae. Some of these devices can be connected to the internet (mainly through a subscription to an access operator), and share their connection with the others (eventually splitting the cost with every user). While multiple challenges remain, there is an increasing deployment of these WMNs by citizens, non-governmental organisations

[61] One of these technical solutions is "P2P caching". Its idea is to store the most accessed file parts locally to avoid multiple unnecessary transmissions over the internet.

[62] Voice over IP (VoIP) is a category of applications that allows internet users to phone each other through the internet.

[63] Many people build a private Wi-Fi network at home to share internet access between multiple devices simultaneously (computers, smartphones, tablets, etc.).

[64] The researcher describes the situation in the US, but similar ones exist in most countries.

and even government agencies [93]. In this context, it is important that regulators open more spectrum to unlicensed users to boost innovations and citizen-based applications (competition principle).

Second, following the revelations on the surveillance programme PRISM, some countries like Germany and Brazil plan to "secure their part of the internet" [10]. Currently, when a datagram must be transmitted between an emitter and a destination located in the same country, it is routed through a series of nodes that can be located in other countries. The idea is to force network operators to maintain a list of all the IP addresses located in a given country, and avoid that the path taken by datagrams to be transmitted between two of these addresses passes by nodes outside of that country. Several people criticise such an approach, arguing that the internet will partially lose its "*ability to take advantage of global cost and* [its] *capacity opportunities to route traffic*" and that it would not solve all the security issues [10]. While it makes sense to protect citizens from massive online surveillance (safety principle), a careful cost–benefit analysis is needed to evaluate the true positive and negative impacts before deciding. Regulators could, for example, simulate if any additional congestion problems occur if the possible paths taken by datagrams are limited to a (national) subset of the internet[65].

5.2 Online content regulation

While the communication infrastructures transmit datagrams, it is only the reconstruction of the whole content from them by an application (for example an email server that sends or receives a message) that gives meaning to the communication. In particular, online services that manage data (email servers, blogs, video sharing websites, social networking applications, etc.) can analyse all of a given content (an email, a picture, a video, etc.). The corresponding providers therefore have the opportunity to discriminate against certain contents (such as removing an uploaded video). As explained in subsection 3.4, discrimination always occurs through the existing biases (mostly technical) and is not necessarily the provider's will (for instance a search engine that does not index a whole website). But a provider can also consciously discriminate against a given content. I will first present a general discrimination framework, and then discuss some particular cases of content discrimination practices.

5.2.1 Content discrimination framework

In democracies, (offline and online) contents remain largely unregulated except when specified by the law: those resulting from crimes (such as child pornography), those related to criminal investigations when warrant is issued by a judge, those distributed without the permission of the creators (in particular copyright protected materials), and, sometimes, those conveying very outrageous talk (for example racism in the EU). For these exceptions, content discrimination could be allowed. But the diversity of situations makes it difficult to systematically differentiate contexts where content discrimination should be allowed compared to those where it should not be.

One approach might be to distinguish private and public communication processes. A *private communication process* presupposes that the identities of the communicating parties are important for both. This is of course the case when two people exchange an email, but it can also encompass an authorised connection of an internet user to a website. A *public communication process* occurs contrariwise when at least one communicating party is not interested in the identity of the other one. A blog falls clearly in this category since its writer wants to be read by everyone, as well as P2P architectures for content distribution since they do not choose who a given content is shared with. It is important to note that this distinction is not linked to a particular application but to specific contexts of its use. An email sent to a mailing list dedicated to questions regarding software is, for example, a public communication process. Some online applications that need an authentication to be used, such as social networking applications (for example *Facebook*), are more tricky. On the one hand, since internet users must identify themselves in order to enter the application, they can be seen as private communication processes. On the other hand, when an internet user shares his or her messages with hundreds of "friends" (the friendship with most of them being limited to a click on an "OK" button), the process cannot be considered as private any more. Again, the distinction depends on the context: communication within an online social network limited to well-known people only remains in the private sphere, while communication processes in large online social networks should be considered more as public. This distinction will probably

[65] In particular, regulators could verify if a "large subset" (such as Brazil or the EU) ensure the same level of flexibility as the whole internet when transmitting datagrams.

need some clarifications in several practical cases (for example by courts), but it should work for most situations.

It is then possible to define the acceptable discrimination approaches based on this distinction. For private communication processes, no content should be controlled and discriminated against (safety principle), except when they fall in the scope of the legal exceptions described above (constitutional principle). For public communication processes, software which performs some well-defined analyses automatically on online content can be accepted. Such analyses must be based on a set of clear rules (content and context) that respect the three regulation principles. In a limited number of situations, providers should be allowed to filter or remove some contents (for example blocking spam or removing uploaded copyright protected material from a website). As in the case of network operators, regulators must control providers to avoid any inappropriate uses (overblocking and illegitimate cyber surveillance). This regulation framework can manage complex situations characterised by legal heterogeneity. Consider, for example, that a Belgian posts a message containing racist words on his *Facebook* profile that has hundreds of followers. The number of followers makes the message public. Since *Facebook* is hosted in the US where free speech is absolute, Facebook should not be forced to remove it (constitutional principle). But a lawsuit must be filed against the sender in Belgium where racist words are illegal (constitutional principle). Moreover, a Belgian court could ask national network operators to block the particular *Facebook* page containing the problematic message (constitutional principle), but not the whole *Facebook* website (safety principle).

5.2.2 Discrimination against online contents

However, some providers, whose business model depends on a huge (and increasing) number of customers, discriminate more strictly against online contents. In particular, they tend to remove "problematic" content. Facebook, for example, does "*not permit individuals or groups to attack others based on their race, ethnicity, national origin, religion, sex, gender, sexual orientation, disability or medical condition*"[66]. A discussion on the "problematic" character of a given content is outside the scope of this article. It suffices to say that it varies from one place to another and that it is a mix of social norms: legal constraints (as described above), social and cultural conventions (for example nude pictures must be avoided on websites accessible by young internet users) and contexts (such as private and public communication processes).

The detection of "problematic" contents can be manual. For example, several websites propose a mechanism to their users to report that a particular content appears to be problematic (typically a form in a webpage). Some providers hire employees who scan the contents in order to detect the problematic ones. Since manual detection is impossible to deploy on a large scale, technologies were developed to automatically distinguish problematic and non-problematic contents. Since the nineties, technologies can, for example, identify online images containing naked people [32]. Recently, researchers have been working on automatically detecting online harassment that pollutes many online applications today [22]. It is also known that some providers (such as Google or Microsoft) use such technologies to scan some data stored (such as emails) for detecting illicit material [100]. All these technologies use some "features" of the content analysed (such as combinations of words in texts) to decide if it is categorised as problematic or not. Mostly, these features are automatically established by processing a given set of contents for which the categorisation is known (a so-called learning set). Most technologies have two weaknesses. First, it is impossible to obtain 100% accuracy in the results: some problematic contents will be classified as non-problematic and vice-versa. Second, as for datagram discrimination (subsection 3.2), choosing which features to use or which learning set to process can introduce biases.

Once a potentially problematic content is identified, the provider must decide if it discriminates against it (such as by removing it) or not. The easiest method is to combine automatic detection and suppression: once an algorithm has identified a content as problematic, it is automatically removed from the corresponding online application. Regarding the weaknesses of the detection algorithms mentioned above, there is a risk that non-problematic contents are removed by error (something similar to network providers that overblock datagrams) or that problematic contents remain online. This is the reason why several providers implement a human-based process. They define rules that are supposed to determine which contents must be discriminated against (violence, harassment, nudity, hate speech, etc.), and use some kind of moderator (an army of paid employees for "big" providers) to verify if a potentially

[66] `http://tinyurl.com/798528d`, *Facebook Community Standards* (consulted on August 26, 2014).

problematic content violates these rules or not. If a violation is found, the content is removed and/or the corresponding internet user is excluded. But interpreting whether a content violates a given rule is not an easy task, and humans always introduce a bias when interpreting a rule.

Regulators should have several roles. If the above discrimination framework is adopted, they could provide expertise in evaluating if communication in a given context must be considered as private or public. If it is normal that providers define discrimination rules to ensure a "fair" use of their services for all customers, regulators should verify that these rules follow the three regulation principles proposed (for example free speech must be preserved). This is especially crucial for "big" online services (*Google Search*, *Facebook*, *Twitter*, etc.) hosted by providers that occupy *de facto* monopolistic positions. In any case, regulators should force providers to precisely document the discrimination rules applicable to their services (subsection 5.3.3). In particular, it would be welcome to provide examples showing how these rules are interpreted in different contexts. One problem is a provider continuously changing the rules that internet users have to follow. It is difficult, and sometimes counterproductive, to prevent providers from changing these rules, but regulators should ensure that these changes do not harm internet users (constitutional and safety principles). In particular, regulators need to combat all forms of "lock-ins" (such as facilitating provider switching as discussed in subsection 5.3.2).

5.3 Online service regulation

The growth of the internet has given rise to a new digital economy based on information computerisation. Powerful new economic stakeholders have emerged (Google, Facebook, Amazon, etc.) and propose multiple online services to internet users (search engines, web-based office suites, online marketplaces, etc.). The development of cloud computing reinforces the role of the internet in our societies. Several providers occupy a position comparable to that of some network operators in the communication market. It is therefore not surprising that comparable problems can occur, and that similar regulation mechanisms should be implemented. Some issues related to online service regulation were already discussed when the constitutional, safety and competition principles were introduced. Here, after some general remarks on online services markets, I shall discuss four issues where regulation and regulators have a role to play: competition, transparency, certification, and the particular case of public online services.

5.3.1 Online service markets

I have proposed to consider the global online market as a collection of online service markets and online revenue markets (subsection 4.3.3). One can distinguish two categories of online service market: one-sided markets and two-sided markets. The first ones include the online services for which a transaction always involves the provider and only one particular stakeholder (an internet user or a group of internet users). Document storage in the cloud are an example: once a folder is created online, reading and storing documents (the transactions) imply only the stakeholder who has access to the folder and the provider. For two-sided online services markets, providers act as intermediates between different stakeholders. For example, a search engine links internet users who formulate a query and online immaterial commodity providers [101]. Similarly, a social networking application (such as *Facebook* or *LinkedIn*) connects different categories of stakeholders (friends, people and organisations that communicate online, employees and employers, etc.). Moreover, some online revenue markets are also two-sided markets and providers also act as intermediates between different stakeholders. Providers active in the online advertising market, for instance, target internet users for advertisers. In the same way, a cultural dematerialised goods provider (such as Apple through *iTunes*, Amazon or eBay) connects buyers and sellers.

Several online service markets are two-sided markets where the corresponding providers are a "'*gateway to cyberspace' for end users, i.e. the key passages towards applications and content that end users have to go through in order to fully enjoy the cyberspace experience*" [86]. Google with its search engine best illustrates this, but it is not the only one. For example, more and more internet users use *Facebook* to find information or play games. There is a risk that some powerful providers that occupy a *de facto* monopolistic position in two-sided online markets become exclusive "gateways to cyberspace". They can then either discriminate against certain stakeholders in some online service markets (in particular potential competitors), or capture all the revenue from a particular online revenue market.

One characteristic of the digital economy explains the possible emergence of oligopolies[67]: the tendency of the marginal operating cost to fall. The marginal operating cost represents the change in the total cost when the quantity produced increments by one unit. Let us take a search engine. Suppose that its infrastructure is capable of responding to 125 million concurrent requests. The marginal operating cost is the investment needed to respond to 125 million and one concurrent queries. It is obvious that, in this example, the marginal operating cost is zero. In fact, the marginal operating cost tends to be very low for online services. In other words, the initial investment costs to provide a given online service are generally much higher than the daily investment needed to meet an increased continuous demand. This reinforces the position of "big" providers and allows them to concentrate the necessary capital to increase the quality of their existing online services (such as building new data centres to decrease respond times) and to develop new ones. Since these "big" providers already have the necessary infrastructures to deploy online services on a large scale, they can easily provide new ones that are immediately available for millions of internet users and dominate new online markets. Moreover, the concentration of capital gives them the opportunity to buy other companies. This is not necessarily a bad thing, as it could result in better online services for internet users. But it also tends to create internet giants (such as Google or Facebook) that can potentially control who accesses which online immaterial commodities.

5.3.2 Online service competition

Everybody agrees that the development of the internet is largely due to constant innovations in the online services proposed. It is therefore essential for new services to emerge in future. On the other side, possible biases introduced by popular online services (section 3) may justify regulation mechanisms.

As mentioned earlier (subsection 4.3.3), the technological "lock-ins" that bind internet users to providers must be fought. The rumour that Bruce Willis sued Apple to ensure that its children could access his *iTunes* digital music library after his death illustrates the problem [1]. Many providers do not give their users all the management rights on their digital assets (emails, digital music, blog contents, calendar, etc.). This hinders them from changing providers or from passing ownership to someone else, as with common material assets. Two simple obligations for providers imposed by regulators could prevent this. First, each provider should propose an easy-to-master "export feature" to its users that restores all their data in a machine-understandable standard (XML, FLAC, etc.). This would make it easy for internet users to switch to another provider without losing all their assets. Such a feature would also allow the transfer of digital assets between internet users (through an export from a first internet user and a subsequent import from another one). Second, providers should delete users' data when asked by them to do so. This is not only a privacy issue, but also a commercial one. It guarantees a true termination of the contract that links an internet user and its old provider.

The whole Net neutrality debate emerges because people feared that network operators used their dominant position in a two-sided market to discriminate against some stakeholders and/or to favour others. Since some providers have similar positions in their own market, or even more dominant ones[68], the same kind of danger exists. In fact, in 2006, *Google Search* discriminated against webpages from Foundem, a company that proposes a potentially competitive search engine [39]. Similarly, in 2010, Apple censored a version of JAMES JOYCE's *Ulysses* for iPad because it contained a picture showing a woman's breasts [11]. And in 2014, Amazon, which has a 40 percent market share of all new books sold in the US, started to discriminate against writers (through delays in delivery or no pre-order options) from the book publishers Hachette and Bonnier-Verlagsgruppe with which it is involved in a commercial dispute [19]. American[69] and German[70] writers denounced Amazon's practices and what they considered as an abuse of a dominant position in widely publicised open letters. Of course, some forms of discrimination are legitimate. For example, a search engine can remove webpages containing illicit materials from its index, or a social networking application can block the account of an internet user who harasses others. But regulation mechanisms must ensure the neutrality of an online service

[67] An oligopoly is the domination of a given industry by a few stakeholders.

[68] *Google Search*, for example, concentrates more than 90% of the search engine market in some countries, a market share that not many stakeholders have on the communication market (except perhaps Cisco Systems).

[69] `http://authorsunited.net`, *A Letter to Our Readers* signed by more than 1000 American writers (consulted on August 27, 2014).

[70] `http://www.fairer-buchmarkt.de`, *Offener Brief an Amazon* signed by more than 1700 German writers (consulted on August 27, 2014).

related to a two-sided market (for example search neutrality for search engines) [55], in particular in online markets where dominant positions exist.

As for network operators, regulators must ensure that discrimination technologies are legitimate and do not overblock, and that providers do not favour particular online services or contents. Many of the proposed regulation mechanisms for communication infrastructures (subsection 5.1) can be applied to online services: the acceptance of some discriminations to protect internet users (safety principle); the verification that technologies implemented do not overblock (constitutional, safety and competition principles); and the surveillance of possible monopolistic practices (competition principle). The decision of the European Commission (EC) in February 2014 to force Google to propose three rival services whenever it promotes it own ones (subsection 4.3.3) is an example of the competition principle applied to online services. In August 2014, the results proposed by *Google Search* for the query "deep purple child in time"[71] starts with three videos hosted by *YouTube* (owned by Google), while the next video-sharing website ranks in eighth position (and is not visible without scrolling). Regulators must be particularly vigilant on service neutrality.

Monopolistic positions cannot only occur in an online service market, but also in an online revenue market. For example, in 2013, Google represented 33% of worldwide revenue from online ads [109]. Of course, one cannot prevent internet users from consuming online services that allow the corresponding providers to obtain most of the revenue from an online revenue market. But antitrust mechanisms can ensure that one company is not monopolistic in an online revenue market through the concentration of several online services, and they can block the takeover of a given company by another if the corresponding agglomerate becomes monopolistic in one online revenue market. In fact, these mechanisms exist and are applied to offline markets[72]. Internet regulators must be more reactive to these online situations (competition principle).

A particular example of the situation described above occurs when multiple providers compete in the same online revenue market, with some of them acting as intermediates for the others in a two-sided online service market. News aggregators are good examples: they select articles from several worldwide news sources, retrieve the title and the first sentence of the article (sometimes also a picture), and present the results on a single webpage. The problem may be that internet users settle for this small amount of information and do not go to the original website. The news sources, which create a great part of the added value, lose online traffic and therefore potential online revenue based on online ads while the providers of news aggregators increase their traffic and their revenue (either directly through online ads presented with the aggregated results or indirectly through other services proposed by the same provider). In a sense, we can speak of a sort of plunder of legitimate revenue. Again, it is difficult to define how internet regulators should react. One way to tackle this could be to consider that some sort of revenue sharing between stakeholders (that must be negotiated between them) limits a potential monopolistic position

5.3.3 Online service transparency

Guaranteeing competition between online services is not the only action that must be taken to limit biases. Ensuring greater transparency in their functioning is another one. Providing internet users with clear information can help them to evaluate the benefits and the drawbacks of a particular online service before using it (safety principle) or changing to another one (competition principle).

One could argue that, since multiple solutions exist for a given category of applications (search engines, email accounts, etc.), it is the responsibility of the clients to inform themselves before making their choice. This may be true for very specialised online services where clients have the skills and the resources to analyse the different solutions and choose the best suited one (such as a company that selects a given email provider). But there are at least three reasons which invalidate this for most general-use online services. First, providers are generally reluctant to provide enough information on their services or on how they use the data collected by their services (mostly because they believe their business needs secrecy). For example, many providers do not acknowledge that they collect certain information in server logs and, those that do, document too rarely how long the information is stored (sometimes "eternally") and how it is used (such as profiling internet users). Second, even when information exists, many internet users do not necessarily understand it because they lack technical skills or because the

[71] The query refers to the song *Child in Time* by the rock band Deep Purple.

[72] An example is the fusion of Holcim and Lafarge in the cement market where regulators forced them to give up certain assets.

useful information is found on many webpages. Third, there are categories of applications for which internet users lack, in practice, real alternatives. Only a few search engines provide an acceptable indexing of the Web, and most internet users have "no other choice" than to use them. Similarly, if all our "friends" use a particular social networking application, we are "forced" to use it too. Regulators should ensure that enough clear and easily understandable information is provided on how online services work and on the underlying practices of the corresponding providers.

As already mentioned, such regulation mechanisms exist in other industrial domains (subsection 4.3.2). Pharmaceutical companies must disclose the active substances of their medical treatments with all the contraindications. Similarly, food producers must specify the ingredients of their products and the amount of calories they represent. In food processing, they must also ensure traceability throughout the whole production process. In fact, many products must explain the risks encountered by consumers when used (for example that a prolonged listening with headphones can damage hearing). There are therefore no reasons not to adopt similar obligations for online providers.

Given the importance of search engines today, some researchers have suggested several transparency measures [104], in particular regarding their ranking methods. They propose that search engines justify why a given website falls in the ranking, that they establish a mediation process for those who feel aggrieved and that they provide clear guidelines about the ranking optimisation techniques they allow. While these measures may ensure fair competition between website owners for online visibility, they must be completed with transparency mechanisms vis-a-vis internet users (the other part of the two-sided market). In particular, a partial disclosure of the ranking criteria is needed to correctly interpret the results retrieved by search engines. To protect the exact ranking algorithm (an argument used by search engine providers to oppose any disclosure), regulators may act as intermediaries. They can sign a confidentiality agreement and write a "guide to online search" that helps internet users with guidelines without detailing the ranking algorithms.

As already discussed (subsection 4.3.2), search engines are not the only online services for which more transparency is needed. In particular, every time technical biases exist (such as for recommendation algorithms), the corresponding online services should document them. But transparency can also imply that providers specify which kind of protection they guarantee their customers. For instance, providers that organise economic transactions between different stakeholders should document how they protect them if problems occur. While some well-known online services (such as *Amazon* or *eBay*) propose reimbursement mechanisms, this is not systemically the case.

It is also crucial that the information provided is comprehensible by all internet users. This is a well-known problem in the offline world (unclear commercial contracts, not easily repayable loans, etc.). The economic crisis of 2008, for example, is partially rooted in the sale of financial products whose contents nobody understood. Regulators must therefore not only force providers to give more information, but also ensure that it is presented in a way that makes it understandable by an "average" internet user. Some researchers suggest that minimal contextual information is always provided by an online service to its users [25]. They illustrate this with a particular case: searching for restaurants near a given place with *Google Maps*. They notice that, for the default view map scale, only a few restaurants are shown (perhaps because they have paid for some advertising) while, when zooming, others appear. They propose that a comment is added such as "showing x out of y restaurants" accompanied by an icon that provides the internet user with a list of all restaurants. Some online services already propose such contextual information. For example, recommendation systems sometimes indicate why they recommend a given product ("people who bought x, y and z also bought a"), even if they do not clarify all the criteria used (such as the popularity of the suggested product).

The examples show that three kinds of information are needed to ensure a real transparency. I propose to call them *business information*, *infrastructure information* and *contextual information*. The first explains how online services generate revenue. Typically, a provider whose business model is based on online ads, should document which data collected by its online services are used and how they are aggregated to target its customers. This is particularly important for "free" online services when internet users do not necessarily understand how their interactions are valued by the corresponding providers. Business information is global for a given provider and must be referenced at a specific place (for example, detailed in a set of webpages). Infrastructure information concerns the general rules followed by a provider to run a given online service. A search engine describing its ranking method or an online retailer documenting its recommendation algorithms is infrastructure information. Similarly, providers should inform their customers what kind of protection they propose. Infrastructure information

is global for a given online service and must also be referenced at a specific place. Contextual information helps an internet user to interpret his or her interaction with an online service at a given moment. *Google Search* provides, for example, the approximate number of results for the query submitted. Similarly, *Amazon* gives a hint why it recommends a given product to an internet user (generally some of the user's past purchases). Contextual information is local to a particular interaction and must be directly visible to the internet user. In practice, for most online services, business and infrastructure information is almost never provided to internet users, while contextual information is shown more often (but not always completely). Regulators could force the most popular online services to provide the necessary business, infrastructure and contextual information.

5.3.4 Online service certification

The regulation of online service should not be limited to transparency. In particular, providers can lie or omit some important information to correctly assess their online services. Let us take the example of medicines once again. Indicating the active substances and the contraindications is not enough to sell a medicine, as regulators (such as the Food and Drug Administration in the US) must also allow it to be placed on the market after a "certification process". Something similar should also exist for some online services.

Due to their importance, the certification of online services based on cloud computing infrastructures has already been studied. First, they centralise many sensitive data[73], and it is necessary that the corresponding providers ensure that neither another client nor an internal employee can access the stored data [105]. Second, it is often difficult for potential clients to choose a particular online service based on cloud computing infrastructures: they lack information to evaluate the benefits of using the service and they have no metrics that help them to compare solutions offered by different providers (including the level of data security) [97]. Third, legal differences between the countries of internet users and of providers may lead to conflicts. The classical example is an EU internet user using a US online service: he or she may believe that the provider respects the EU privacy legislation while, in practice, the provider must only respect the US (less protective) one. Several criteria were proposed to certify online services based on cloud computing infrastructures, including data protection mechanisms, the security technologies deployed, the compliance with a given legal framework and the available information about any kind of potential lock-ins for their clients [97, 105]. The EU introduced a new framework to ensure that online services based on cloud computing will better respect the EU legislation on privacy, and has also proposed a pan-European certification process [56].

But internet users could benefit from the certification of other online services. The problem of different jurisdictions for internet users and providers is, of course, not limited to services based on cloud computing infrastructures. For example, it could also reassure internet users if an established third-party would certify that a given online marketplace (such as *Amazon* or *eBay*) provides some guarantees for their buyers (return an item, replace an item, etc.).

There is of course no sense trying to certify all online services. Not only is it impossible, but it would also be a real hurdle for innovation. Moreover, a certification process is generally costly: if it is done by a public organism, it constitutes a load on the state (and therefore on the citizens); if it is done by a private company, the company that must be certified has to pay for it. Certification should concern two categories of online services: the most popular ones, and those for which certification represents a competitive advantage. For the first category, the popularity of some online services (for example the five most used search engines, social networking applications, storage services, etc.) makes them of general interest. It therefore seems natural that regulators provide some kind of certification. Knowing, for example, that a search engine does not store the history of individual queries carried out by its users while others do, helps "normal internet users" to choose the one they want to use. The cost of a public certification process is offset by the increased safety for many internet users (such as for medicines). After all, *Facebook* has more than one billion users and *Google Search* processes three billion queries daily [71]. For the second category, certification can help a given online service to differentiate itself from its competitors. An email or storage provider that is certified to implement the highest level of security measures can attract new clients such as companies with highly sensitive data. Here, certification by

[73] Sensitive data include but are not limited to personal data. An organisation that uses an online storage service must be certain that confidential documents remain protected for unauthorised access.

private third-parties (paid for by the companies) is more adequate. The ISO 9000 series of standards can serve as templates.

5.3.5 Public online services

Public online services are a particular kind of online service provided directly or indirectly (though public funding or contracting) to citizens by states. I distinguish three categories of public online services.

First, is the availability of government data in an open and machine readable format. Examples include websites such as `http://data.gov`, `http://data.gov.uk` or `http://data.gouv.fr`. The use of *open government data* is supposed to ensure better public accountability and boost economic growth [68]. A survey shows that open government data also have a positive impact on innovation [58]. As already mentioned when the constitutional principle was presented (subsection 4.3.1), several states have a freedom of information legislation that should force them to adopt open government data widely. This process must be generalised and extended to every country[74]. To simply publish data online is necessary but not sufficient. In fact, if the data are not accessible through a structured and documented format (such as XML), civil society stakeholders will be unable to process it. Moreover, since the different data are stored in multiple places (mostly databases), if a coherent outline is not adopted, it will be economically impossible to cross data to extract useful knowledge.

Second, is the use of the internet to propose an *electronic government front-office*. States propose online access to public services to allow more direct interactions with their citizens and companies [63]. Belgium, for example, allows its citizens to fill in their tax return over the internet. Electronic government front-offices raise one issue: all citizens and companies must be treated equally (constitutional and competition principles). This has two implications. On the one hand, a public online service cannot entirely replace its offline counterpart while all citizens cannot easily access the electronic government front-office, either because some lack an internet connection, or because they are insufficiently educated to use the online service. Taking the example of online tax return in Belgium, since more and more Belgians use it, the government could be tempted to stop sending the paper form (in particular to save costs). But, to fill in a tax return, a citizen must first install a Java runtime and a particular plug-in for the browser. While the installation process is not particularly complicated, it is still too complex for many citizens. Therefore, today, public online services must mostly be seen as complementary services to the offline ones that must be promoted, rather than replacement services. On the other hand, public online services must be accessible by "standard applications". An example from 2007 is when the English government portal integrated non-standardised Web elements forcing every British citizen to use *Microsoft Explorer* to access public information [3]. Translated in the offline world, it would be similar to the police only accepting phone calls made with a particular telephone company. This is of course unacceptable, as are similar online situations.

Third, states can run or sponsor *online public interest services*, those that no private provider offers or those to be detached from commercial constraints. One example is `http://Europeana.eu`, a website that provides a portal for digitalised European cultural goods (books, films, paintings, etc.) and that is partially sponsored by the EC. But more application-oriented online services could also be funded. Since it is well known that search engines are the primary tool used to search for information online [83], it could make sense to develop a public search engine. It could disclose all the algorithms used (indexing, selecting and ranking methods), publish the list of the webpages indexed, and completely respect the internet user's privacy because its funding does not require revenue to be generated (and thus internet users to be profiled). In fact, the *Quaero* project launched in 2008 by a Franco-German consortium was an attempt to build such a public search engine. Although the scientific contributions of the project are indisputable, it fails to provide an alternative for commercial search engines (partially due to the lack of funding). But this does not mean that future similar initiatives will fail. States have shown in other domains that they can massively fund general interest research projects (take the CERN for example). The funding of online public interest services should be publicly debated.

It is of course crucial that public online services should exercise *technological neutrality*: they must "*prevent an unfair competitive advantage for existing technologies or for specific companies that produce or employ one or the other technology*" [45]. Open government data must be published in an open, machine-

[74] It is, of course, improbable that authoritarian regimes disclose government data. The difficulty to collect reliable information about greenhouse gas emissions and air pollution in China exemplifies this. But, again, it should not stop democratic countries from adopting open government data.

readable and coherent format. Electronic government front-offices must only use open standards and technologies in order to guarantee citizens and companies their freedom to choose whatever software they want. This involves three missions for regulators: to guarantee that governments make all data that should be public accessible online (constitutional and safety principles)[75]; to ensure that governments respect the same regulation rules (in particular concerning privacy) as every provider; and to verify that the public online services are technologically neutral.

It should be noted that many institutions facing a lack of legitimacy (such as the EU) hope that public online services may help them to gain an image of transparency and promote public debates [76]. In this context, unregulated public online services can create more harm than good (frustration when unable to find public information, unequal treatment, etc.).

5.4 Some cross-layer issues

All the regulation issues cannot be described according to the vertical approach used in the previous subsections. There are also horizontal regulation issues that concern every internet operator regardless of the layer on which it operates. These issues are independent of any particular technical discrimination approach. I will briefly discuss two of them: the cooperation between internet operators and states, and online privacy.

5.4.1 Internet operator cooperation

Ideally, internet operators, in particular commercial companies, should be independent of states as long as they respect the current legislation and regulatory framework. In democracies, internet operators cooperate with states in very few cases (for example, if a warrant is issued by a judge) or when they are under contract. In authoritarian states, internet operators, including private organisations established in democracies, are generally forced to some cooperation to be active in the corresponding "part" of the internet. So the question of the level of cooperation of internet operators with states is raised.

As explained in the previous sections, some internet operators (network operators or providers) have legitimate reasons not related to network management issues to discriminate against their clients (such as blocking a computer virus, spam, etc.). In these situations, should they cooperate with states, contact local authorities and provide them with information such as account data or IP addresses? Let us take the example of a video sharing website. If the provider detects that copyright-protected material is uploaded, it must remove it. But it probably makes no sense to contact the police for each such case. This position cannot be defended if the uploaded video contains child pornography. This question becomes more complex if the problematic video is uploaded by someone living in a authoritarian country. Intuitively, everybody would agree that nothing must be done regarding "harmless" problems, but actions are (ethically) obligatory regarding "harmful" ones. But these two categories are dependent on a given context. If a dictatorial regime asks for some information on political opponents because "they threaten national security" (a harmful problem in their view), it is unethical for an internet provider established in a democracy to respond positively. One way to tackle this complexity is to evaluate a situation in the light of the offline concept of the "failure to provide assistance to a person in danger" defined by several countries[76]. This concept presupposes that some kind of responsibility is engaged when someone does not try to avoid situations that he or she knows will harm another person (such as not helping someone who is drowning). In the online context, an internet operator must contact public authorities only when a non-denunciation leads to such situations. According to this concept, an operator must act in the case of child pornography no matter where the internet user uploads the video (constitutional principle), but not in the case of political opponents (safety principle).

Many internet operators are commercial companies whose main motivation is profit. Their growth is generally linked to an increase in the number of their clients, which leads them to cover a maximum of geographical areas including those under authoritarian regimes. Should they cooperate with these regimes to access the corresponding markets? In particular, should commercial companies established in democracies provide help (hardware, software and/or services) if they know that these regimes will

[75] Some data could be made accessible only for specific stakeholders (such as accredited researchers). But it is not the role of the regulators (at least not the ones described in this paper) to decide which data must be public and which must not.

[76] In many countries with common law, there is no general legal duty to provide assistance to a person in danger.

rely on or benefit from them to act in a way that would be considered as illegal in a democratic context[77]? Such a collaboration occurred in 2005 when Yahoo! gave information on a particular account it managed to the Chinese government, which led to a 10-year prison sentence for the Chinese journalist SHI TAO [49]. The debate regarding the complicity of companies in acts committed by states with which they collaborate is not new. It started after the Second World War and, since then, several commercial companies have been accused of violations of human rights laws (such as oil companies that hire military and police forces in authoritarian regimes to secure excavation activities by committing crimes against the local populations) [21]. Three categories of corporate complicity have be identified [**?**]: direct complicity (a company provides assistance to a stakeholder who violates human rights laws); beneficial complicity (a company benefits from the violations of human rights laws by a stakeholder); and silent complicity (a company is aware of violations of human rights laws by a stakeholder outside its sphere of activity but does nothing). While examples of companies prosecuted for direct complicity exist, beneficial complicity and silent complicity "*will rarely give rise to a case determined before a judge – but the organising and moral force of complicity allegations should not be underestimated*" [21].

Let us analyse two situations of cooperation between an internet operator and an authoritarian state: the sale of technologies to track internet users (for example to identify the real author of an anonymous online message), and the partial censorship of a search engine by its provider (such as removing some websites from the results). At first glance, no company established in democracies should agree to cooperate in any of these situations. But, in practice, there is a difference: while the first situation corresponds to direct complicity because it leads to harm for some citizens (because they can be arrested once they are identified), the second situation is beneficial complicity that does not change the reality of internet users (they will face censorship of websites anyway). These situations do not, of course, differ by nature (cooperation leads to undemocratic acts by states), but a difference in degree indubitably exists. As for other commercial companies, internet operators should therefore be treated differently depending on the form of cooperation with (authoritarian) states. Some researchers go further and call for the inclusion (in democracies) of "*a positive obligation to protect*" for commercial companies (something that could be in keeping with the safety principle) [107]. Such a legal measure would force internet operators located in democracies to refuse any form of cooperation in the first situation analysed, leaving them free to act as they wish in the second situation. One could argue that preventing national companies from being fully present in authoritarian markets could weaken them, compared to those established in other (democratic) states that do not have such restrictions (competition principle). But it is an ethical choice to consider that the competition principle is subordinated to the constitutional and safety principles.

Prosecuting an organisation established in a democracy, D, for its cooperation (for example, by selling technologies that it knows will be used "badly") or its lack of necessary cooperation (such as not denouncing child pornography) with a foreign state, F, is a complex issue. Some countries have a legislation that allows residents to be prosecuted for crimes committed abroad. The US Alien Tort Statute (ATS) was, for example, used by advocates to prosecute Yahoo! for complicity in torture in China by providing the authorities with the necessary information to arrest political opponents [20]. But, in practice, "*the ATS had played a modest role in the human rights movement*", and, through several decisions of the US Supreme Court, it has been narrowed [94]. Moreover, generally speaking, national jurisdictions cannot judge crimes in other countries. Let me present three naive ideas. Several countries have a universal criminal jurisdiction for international crimes. For example, in March 2014, a French court judged a Rwandan citizen for complicity with genocide [108]. Since most authoritarian states have a minimum of one hundred thousand internet users, it could be claimed that online harms are mass crimes. According to such reasoning, an internet operator can then be sued in D for certain kinds of cooperation with F. For countries without universal criminal jurisdiction, two other ideas exist. First, if a link of subordination can be established between an organisation and a local stakeholder that cooperates with F (for example, through the existence of a subsidiary, a contractor or a sub-contractor in F), that organisation can be considered as being involved in the decision and, therefore, be sued in D. A similar situation occurs if a Belgian citizen sponsors the murder of someone in another country: he or she can be prosecuted by a Belgian court. Second, once it is proved that the cooperation of an internet operator has led to "bad"

[77] As the recent disclosures on the surveillance programme PRISM have shown, some commercial companies already cooperate with democratic states in a probably illegal systematic surveillance of American citizens. But these companies can be prosecuted before national courts if they participate in illegal actions, which is far more complicated in the case of cooperation with foreign states (unless there is a prohibition of exports to these states in their home country).

uses by F, one may consider that all potential internet users present in F are victims, including those who are normally residents in D (residents of democracies visit authoritarian countries including Presidents and Prime Ministers). A D court can therefore consider that a prejudice exists regarding its own citizens, and sue the organisation. These three exploring ideas seem a little far fetched, but my point here is that some existing offline legal mechanisms can be also valid online. In other words, making organisations accountable for their foreign actions is more a problem of political will than a technical legal difficulty.

5.4.2 Online privacy

Online privacy is one of the most discussed impacts of the internet and mobile devices on our societies. It is of course impossible to address it in a satisfactory manner in the scope of this essay. Nevertheless, I shall briefly discuss three topics. Two of them have already been discussed in this essay: cyber surveillance and the ownership of the data uploaded by internet users. A third topic is the balance between freedom of expression and the privacy of individuals. These topics are crucial because of the ease of collecting online data and the low cost of storing it. States and private stakeholders are therefore tempted to gather everything and anything. Moreover, with the advent of a big data era and the emergence of the Internet of Things[78], a debate on privacy in the digital world becomes urgent. A company announced recently that it has developed a facial recognition technology for *Google Glass* that could "*identify any stranger walking down the street, without his knowledge or consent*" [53]. At this point, it is not clear how efficient the technology is nor if it would be approved by Google, but it shows why privacy is in danger today. I will summarise some elements already mentioned, but also introduce new ones.

The disclosure of the surveillance programme PRISM shows that cyber surveillance is a serious topic in democracies[79]. Almost nobody denies that states are responsible for the (cyber) security of their citizens, and can collect personal data under certain circumstances. For example, laws in several democratic countries force network operators to archive several months of the online activities of users in the case that an investigation needs these data. More generally, states could extend such measures to other internet operators. Let us take the example of a video sharing website once again. Once it removes an online video because it contains copyright protected material, must it archive this information for a given time in the case that it might be useful for an investigation? The topic of cyber surveillance has two facets: a political one (to what extent can state or private stakeholders spy on citizens?) and a technical one (how can cyber surveillance be controlled?). As already mentioned (subsection 4.3.1), I believe that in the name of the "war against terror", there are currently not enough constraints on what can be done in cyber surveillance. Large systematic automatic cyber surveillance projects (such as the one performed by US intelligence agencies) should be banned. Targeted cyber surveillance projects (individuals, specific contents, etc.) should always take place only if a precise control exists (such as a warrant issued by a judge). After a reasonable period of time[80], collected data should be destructed unless a warrant orders that particular data stay stored. This may sound bureaucratic, but I think it is the only way to ensure that only a small amount of collected data considered as valuable remain stored. In this context, as argued in subsection 4.4, regulators must contribute to the control of the different stakeholders (governmental agencies and commercial companies) and verify that established rules are respected (such as the period of storage).

Companies have reasons other than security to collect data on their users. In particular, many of them use data gathered to generate revenue (online advertising, sale of marketing information, etc.). In fact, it is highly probable that most online services archive all the past activities of their users (such as who was "friends" with whom or who used which keywords)[81]. There are legitimate reasons to gather

[78] The Internet of Things refers to the interconnection of multiple devices (in particular sensors) within the existing internet infrastructure. Smart thermostats, that allow the remote control of a home heating system, are examples of devices composing the Internet of Things.

[79] Cyber surveillance is also a major concern in authoritarian regimes, but the nature of these regimes makes a discussion about regulation meaningless.

[80] Which period of time is reasonable depends on the balance between the protection of privacy (very short period) and the time data collected may be needed by security services (very long period). At first glance, a period between 6 and 12 months seems a reasonable compromise. It is also possible to define different periods depending on the nature of the data collected. Data related to "very dangerous activities" (tax evasion, terrorism, global criminality, etc.) could be stored for several years, while data related to everyday activities should be stored for only 3 months. Ideally, laws should define the scope of "very dangerous activities".

[81] Providers of many popular online services (Google, Facebook, Twitter, etc.) are primarily hosted in the US where commercial companies have less constraints concerning privacy than other regions such as the EU.

data about online customers. For example, a newspaper website can track the articles consulted by its users to charge them accordingly on a monthly or annual basis. Moreover, it can also use these data to suggest new articles based on previously consulted ones. More generally, it is internet users' ownership of the data they provided directly (such as a message posted on *Facebook*) or indirectly (for example through a cookie) that is challenged. Currently, the general rule is that an internet user loses his or her ownership of personal data once uploaded. Regulatory frameworks are therefore limited to specify what companies can or cannot do with the data collected. While some regions such as the EU ensure enhanced privacy mechanisms [17], most regions propose less protective mechanisms or no mechanisms at all [50]. Even when internet users choose to destroy an account with a given online service, they have no guarantee that their data are not kept by the provider. Ideally, a regulatory framework should revert this general rule: ownership of data must be retained by internet users. Internet users accept the *loan of proprietary data* to a company only for the execution of a contract and only for the period covered by that contract[82]. Every online service should therefore explicitly describe what data will be collected and for what uses, and internet users must then accept it on a case by case basis (safety principle). Some argue that it is impossible to determine all the possible useful processing of privacy data when they are collected (for example with medical data) [42]. This means that if a company wants to make new uses (for example proposing a new service) or diffuse the data to another one (such as selling marketing data), it must ask explicitly for consent. If the company is bought by another one, the buyer must respect the past engagements. This idea of the loan of personal data also implies that internet users can at any time retrieve their personal data and/or ask the company to destroy them[83]. In such a regulatory framework, the regulators must verify that companies provide enough clear information for an "average" internet user about the uses of personal data, and that only the data necessary for the delivery of the accepted services are collected by the company. This implies that the safety principle (highest privacy level) subordinates the competition principle (companies develop new services or new revenues based on gathered data). Unfortunately, this is not always the case, as shown with the problem of cookies. While the EU tries to regulate their uses, the US prefers to consider them as an opportunity to extend the business [8].

As mentioned when presenting the safety principle (subsection 4.3.2), some researchers ask for a "right to forget" [70]. The idea is that some personal data cannot be maintained online endlessly without harming internet users. This is important since you are not always responsible for the online publication of some personal data. For example, someone can take a picture of you, and post it with a commentary on a popular social networking application without your consent. In other words, without being a "public person" about whom published information may be of public interest, your whole life can be described by others online. In May 2014, the Court of Justice of the European Union (CJEU) recognised that some kind of "right to forget" exists[84]. The case opposed Google and a Spanish man, the latter demanding that *Google Search* no longer index two press articles mentioning his past debts. The CJEU stated that, even if data were initially gathered legally, an internet user can ask a search engine not to use them in a second stage, considering "*that they are inadequate, irrelevant or excessive in relation to the purposes of the processing, that they are not kept up to date, or that they are kept for longer than is necessary unless they are required to be kept for historical, statistical or scientific purposes*"[85]. The CJEU mentioned that a balance must be found between freedom of speech (a collective right) and privacy (a personal right). In response, Google provide an online form where internet users can request to "*remove specific results for queries that include their name, where the interests in those results appearing are outweighed by the person's privacy rights*"[86]. However, it is not yet clear how Google will treat these requests, and if other providers of popular online services (such as *Facebook* or *Twitter*) will face similar obligations. If such a "right to forget" is enforced by laws in future (in particular in the EU), regulators should verify that requests are processed correctly by providers and in a "reasonable" amount of time. Of course, in democracies, a right to inform exists: one cannot prevent "useful" information from being disseminated. This is not a new problem. For quite a long time, traditional media (newspaper, radio and television) have had to deal with the balance between privacy ("nothing" should be disseminated) and the right to inform

[82] This does not mean that internet users should lend or sell all of their data, in particular personal data. The aim is simply for internet users to maintain control of the specific data they allow to be managed by online providers for well-defined purposes.

[83] This is equivalent to a breach of a contract. The consequence is that the company no longer has to provide the service to the internet user.

[84] `http://tinyurl.com/kfdz67r`, judgement of the European Court of Justice, Case C-131/12 (Consulted on May 30, 2014).

[85] *Ibid*.

[86] `http://tinyurl.com/lccacy6`, official website of Google (Consulted on May 30, 2014)

("everything" should be disseminated). Internet operators that produce online contents or mediate it (such as search engines) must comply with an equivalent framework.

Although digital privacy is a concern for every online service, the issue is especially problematic with "big" transnational providers because they transfer large amounts of personal data between different regions with different privacy protection frameworks (for example between the EU and the US). There is at least one technical reason that can be seen as legitimate. To provide fast responding time for their online services, some providers build multiple local data centres around the world, while having a main centre that collects all the modifications made to the data. But the legal status of the transferred data is not well defined. Ideally, a *favourability principle* should exist: data gathered from an Internet user from one region and transferred to another region is regulated by the most protective framework (safety principle). EU internet users should be sure that when they use online services provided by US companies, their personal data are managed with respect to EU directives. Similarly, in normal situations, personal data gathered from internet users in authoritarian countries by US online services must not be given to the regimes of these countries. But, as several examples discussed in this essay have shown, this favourability principle is currently an exception, and not the rule. Governmental authorities (states, the EU, etc.) should be more firm vis-à-vis companies, require them to meet this principle and take effective sanctions if it is not the case (which can range from substantial fines to the blocking of a particular service). As already mentioned, the EC has, in the past, forced a software giant such as Microsoft to change some of its practices to respect EU rules [28]. So, there is no objective reason that something similar cannot be done with major internet operators.

6 Conclusions

Many advocates of online liberty assert that in order to ensure network neutrality, regulation mechanisms should be limited to network operators only. This essay argues that plenty of examples demonstrate that biases are currently introduced at every layer of the internet, and that no real network neutrality exists. In fact, it is likely that it will never be reached (in particular because technical biases will always remain). However, I think that it is possible to limit the increase of network partiality in future through an extension of the current regulatory framework.

New regulation mechanisms should combine new laws and an extended role for independent regulatory agencies. They should also mix national, regional and international levels. While international agreements must be sought regarding global online problems (such as child pornography or spam), the difficulty to reach them implies that national and regional regulation mechanisms should be implemented whenever is it possible. In particular, democracies can introduce new regulation mechanisms that better protect internet users without threatening them through public independent regulatory agencies. Three general regulation principles were proposed to define the scope of internet regulation: a constitutional principle, a safety principle and a competition principle. I have tried to show that many offline regulation rules can be applied online if they are interpreted by regulators with enough technical expertise and resources to understand all the aspects of the internet.

This essay has discussed how new regulation mechanisms (through laws and regulators) based on the three suggested principles can be implemented at the different layers of the internet. Some ideas were proposed to guide the implementation of these new mechanisms (extended certification processes, private and public communication processes, favourability principle, etc.). The proposed regulation mechanisms are not definitive responses to all the existing online problems. Moreover, some of them may have drawbacks that I do not comprehend. They must be seen as an attempt to make my contribution to the inevitably continuous debate on internet regulation. The crucial issue is to find the right balance between additional regulation constraints for network operators and providers (which can sometimes consume time and resources), and their liberty to innovate. But it is even more important to regulate them, especially those which concentrate customers, contents, capital and personal digital data. If nothing is done, most internet users will drown in the network neutrality utopia.

The current proposal implies that national, regional and international independent regulatory agencies should be better funded to ensure their new missions. I am aware that this is a tough fight in the current budget restriction period, but I believe that this expensive path is the only road to take. There are also open questions that still need pubic debate. The discussion on internet regulation cannot be left to engineers, entrepreneurs, internet evangelists or researchers (including myself). In this context, it is crucial that citizens and politicians have a deeper understanding of the internet's technical, social and economic impacts (but this is also true for other fields such as genetic engineering), and that authorities which receive their legitimacy from a democratic process must also be more uncompromising when it comes to defending the general interest against big (online and offline) companies when necessary.

Finally, citizens are without any doubt the most important element of a true democratic regulation of the internet. Through their offline and online choices, they can influence politicians and force network operators and providers to change their current practices. This strengthens the role of education as the cornerstone of our modern democratic societies. More than ever, it is necessary for a public debate to focus on how humanity must deal with technology rather than always running behind the latest new technological "innovations".

Acknowledgements

This work was partly funded by the Région Wallonne under the contract 1017077 and by Innoviris under the project 2013-PFS-EH-1. I gratefully acknowledge helpful comments and suggestions by YANA BREINDL on earlier versions of this work. English proofreading was carried out by JANE CORRIGAN.

References

[1] Charles Arthur. No, bruce willis isn't suing apple over iTunes rights. *The Guardian*, September 2012.

[2] Judit Bar-Ilan. Google bombing from a time perspective. *Journal of Computer-Mediated Communication*, 12(3):910–938, 2007.

[3] Colin Barker. MP attacks government over microsoft policy. *ZDNet*, October 2007.

[4] John Perry Barlow. A declaration of the independence of cyberspace. 1996.

[5] Johannes M. Bauer. Dynamic effects of network neutrality. *International Journal of Communication*, 1:531–547, 2007.

[6] Wolfgang Benedek, Veronika Bauer, and Matthias C. Kettemann. *Internet Governance and the Information Society: Global Perspectives and European Dimensions*. Eleven International Publishing, 2008.

[7] Hal Berghel. Through the PRISM darkly. *Computer*, 46(7):86–90, 2013.

[8] Hal Berghel. Toxic cookies. *Computer*, 46(9):104–107, 2013.

[9] Chip Berlet. Toxic to democracy: Conspiracy theories, demonization, & scapegoating. Technical report, Political Research Associates, 2009.

[10] John Blau. NSA surveillance sparks talk of national internets. *IEEE Spectrum*, 51(2):12–14, January 2014.

[11] Julie Bosman. Apple finds 'ulysses seen' comic a little too graphic. *The New York Times*, June 2010.

[12] Pierre Bourdieu. *Le Sens pratique*. Editions de Minuit, Paris, 1980.

[13] Lisa M. Bowman. Net archive silences scientology critic. *CNET News*, September 2002.

[14] Sergey Brin and Lawrence Page. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30(1):107–135, 1998.

[15] Jerry Brito, Martin E. Cave, Robert W. Crandall, Larry F. Darby, Everett Ehrlich, Jeffrey A. Eisenach, Jerry Ellig, Henry Ergas, David J. Farber, Gerald R. Faulhaber, Alfred E. Hahn, Wayne A. Leighton, Robert E. Litan, Glen O. Robinson, Hal J. Singer, Vernon L. Smith, William E. Taylor III, Timothy J. Tardiff, Leonard Waverman, and Dennis Weisman. Net neutrality regulation: the economic evidence. Available at SSRN: http://ssrn.com/abstract=1587058 or http://dx.doi.org/10.2139/ssrn.1587058, April 2010.

[16] Fabián E. Bustamante. Broadly available broadband. *IEEE Internet Computing*, 17(5):3–5, 2013.

[17] Lee A. Bygrave. International agreements to protect personal data. In James B. Rule and Graham Greenleaf, editors, *Global Privacy Protection: The First Generation*, pages 15–49. Edward Elgar Publishing, 2008.

[18] Ewa S. Callahan and Susan C. Herring. Cultural bias in wikipedia content on famous persons. *Journal of the American Society for Information Science and Technology*, 62(10):1899–1915, 2011.

[19] David Carr. Amazon absorbing price fight punches. *The New York Times*, June 2014.

[20] Ariana Eunjung Cha and Sam Diaz. Advocates sue yahoo in chinese torture case. *The Washington Post*, April 2007.

[21] Andrew Clapham. State responsibility, corporate responsibility, and complicity in human rights violations. In Lene Bomann-Larsen and Oddny Wiggen, editors, *Responsibility in World Business: Managing Harmful Side-effects of Corporate Activity*, pages 50–81. United Nations University Press, January 2004.

[22] R. Cohen, D. Y. Lam, N. Agarwal, M. Cormier, J. Jagdev, T. Jin, M. Kukreti, J. Liu, K. Rahim, R. Rawat, W. Sun, D. Wang, and M. Wexler. Using computer technology to address the problem of cyberbullying. *SIGCAS Comput. Soc.*, 44(2):52–61, July 2014.

[23] Federal Communications Commission. Formal complaint of free press and public knowledge against comcast corporation for secretly degrading peer-to-peer applications. Memoranum Opinion and Order FCC 08-183, August 2008.

[24] Jean-Pierre Coustel. La déréglementation des télécommunications - le poids du démembrement d'AT&t dans la dynamique de diffusion du mouvement. *Revue d'économie industrielle*, 30(1):42–59, 1984.

[25] Andreas Dahn and Clemens H. Cap. Application transparency and manipulation. *Computer*, 47(2):56–61, 2014.

[26] Ronald J. Deibert, John G. Palfrey, Rafal Rohozinski, and Jonathan Zittrain, editors. *Access Controlled: The Shaping of Power, Rights, and Rule in Cyberspace*. MIT Press, April 2010.

[27] Valérie Durieux. Searching for healthcare articles on the internet: User-generated metadata VS librarians' descriptors. In *11th Annual International and Interdisciplinary Conference of the Association of Internet Researchers (AoIR)*, Gothenburg, Sweden, October 2010.

[28] Nicholas Economides and Ioannis Lianos. A critical appraisal of remedies in the EU microsoft cases. *Columbia Business Law Review*, 2010(2):346–420, 2010.

[29] Erika Eichelberger. Can you be denied a loan because you're unpopular on facebook? *Mother Jones*, September 2013.

[30] Gerald Faulhaber. Transparency and broadband internet service providers. *International Journal of Communication*, 4:738–757, 2010.

[31] Gerald Faulhaber. Economics of net neutrality: A review. *Communications & Convergence Review*, 3(1):53–64, 2011.

[32] D.A Forsyth and M.M. Fleck. Identifying nude pictures. In , *Proceedings 3rd IEEE Workshop on Applications of Computer Vision, 1996. WACV '96*, pages 103–108, December 1996.

[33] Pascal Francq. *Internet: Tome 1, La construction d'un mythe*. Editions Modulaires Européennes, 2011.

[34] Pascal Francq. *Internet: Tome 2, Le caractère fétiche*. Editions Modulaires Européennes, 2011.

[35] Batya Friedman and Helen Nissenbaum. Bias in computer systems. *ACM Transactions on Information Systems*, 14(3):330–347, July 1996.

[36] Jim Giles. Internet encyclopaedias go head to head. *Nature*, 438(7070):900–901, December 2005.

[37] Sharon L. Greene, Susan J. Devlin, Philip E. Cannata, and Louis M. Gomez. No IFs, ANDs, or ORs: A study of database querying. *International Journal of Man-Machine*, 32(2):303–326, 1990.

[38] Virgil Griffith. WikiScanner: List anonymous wikipedia edits from interesting organizations, 2007.

[39] James Grimmelmann. Some skepticism about search neutrality. 2010.

[40] Jürgen Habermas. *Strukturwandel der Öffentlichkeit*. Luchterhand, 1962.

[41] Michael Hauben and Ronda Hauben. *Netizens: On the History and Impact of Usenet and the Internet*. IEEE Computer Society Press, 1997.

[42] Jess Hemerly. Public policy considerations for data-driven innovation. *Computer*, 46(6):25–31, 2013.

[43] C. Scott Hemphill. Network neutrality and the false promise of zero-price regulation. *Yale Journal on Regulation*, 25:135, 2008.

[44] Bastian Henze, Florian Schuett, and Jasper Sluijs. Network neutrality and transparency. 2010.

[45] Mireille Hildebrandt and Laura Tielemans. Data protection by design and technology neutral law. *Computer Law & Security Review*, 29(5):509–521, October 2013.

[46] Chris Hoofnagle. Comparative study on different approaches to new privacy challenges, in particular in the light of technological developments - united states of america. Technical report, European Commission, May 2010.

[47] iProspect. Search engine user behavior study. White paper, iProspect, April 2006.

[48] Robert A. Joyce. Pornography and the internet. *IEEE Internet Computing*, 12(4):74–77, August 2008.

[49] Joseph Kahn. Yahoo! helped chinese to prosecute journalist. *The New York Times*, September 2005.

[50] Wolfgang Kilian. Germany. In James B. Rule and Graham Greenleaf, editors, *Global Privacy Protection: The First Generation*, pages 80–106. Edward Elgar Publishing, 2008.

[51] Els Kindt, Eva Lievens, Eleni Kosta, Thomas Leys, and Paul De Hert. Constitutional rights and new technologies in belgium. In Ronald E. Leenes, Bert-Jaap Koops, and Paul De Hert, editors, *Constitutional Rights and New Technologies*, pages 11–55. T.M.C Asser Press, May 2008.

[52] Gary King, Jennifer Pan, and Margaret E. Roberts. Reverse-engineering censorship in china: Randomized experimentation and participant observation. *Science*, 345(6199):891, August 2014.

[53] Erica Klarreich. Hello, my name is... *Communications of the ACM*, 57(8):17–19, August 2014.

[54] Yehuda Koren. Factorization meets the neighborhood: A multifaceted collaborative filtering model. In *Proceeding of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 426–434, Las Vegas, USA, 2008.

[55] Jan Krämer, Lukas Wiewiorra, and Christof Weinhardt. Net neutrality: A progress report. *Telecommunications Policy*, 2012.

[56] Nir Kshetri and San Murugesan. Cloud computing and EU data privacy regulations. *Computer*, 46(3):86–89, 2013.

[57] Nir Kshetri and San Murugesan. EU and US cybersecurity strategies and their impact on businesses and consumers. *Computer*, 46(10):84–88, October 2013.

[58] Erik Lakomaa and Jan Kallberg. Open data as a foundation for innovation: The enabling effect of free public sector information for entrepreneurs. *IEEE Access*, 1:558–563, 2013.

[59] Shyong (Tony) K. Lam, Anuradha Uduwage, Zhenhua Dong, Shilad Sen, David R. Musicant, Loren Terveen, and John Riedl. WP:clubhouse?: an exploration of wikipedia's gender imbalance. In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration*, WikiSym '11, pages 1–10, New York, NY, USA, 2011. ACM.

[60] Neal Leavitt. Network-usage changes push internet traffic to the edge. *IEEE Computer*, 43(10):13–15, 2010.

[61] Robert S. Lee and Tim Wu. Subsidizing creativity through network design: Zero-pricing anf net neutrality. *The Journal Of Economic Perspectives*, 23(3):61–76, 2009.

[62] Christophe Lejeune. {From Virtual Communities to Mediated Collectives}. In Pascal Francq, editor, *Collaborative Search and Communities of Interest*. IGI Global, 2010.

[63] Klaus Lenk and Roland Traunmüller. {Broadening the Concept of Electronic Government}. In J. E. J. Prins, editor, *Designing e-Government*, pages 63–74. Kluwer Law International, 2001.

[64] Xavier Leon, Rahma Chaabouni, Marc Sanchez-Artigas, and Pedro Garcia-Lopez. Smart cloud seeding for BitTorrent in datacenters. *IEEE Internet Computing*, 18(4):47–54, 2014.

[65] Lawrence Lessig. *Code: Version 2.0*. Basic Books, 2006.

[66] Steven Levy. *Hackers. Heroes of the Computer Revolution*. Penguin Books, 1993.

[67] Herbert Lin. Technology's limited role in resolving debates over digital surveillance. *Science*, 345(6198):728–730, August 2014.

[68] Gustavo Magalhaes, Catarina Roseira, and Sharon Strover. Open government data intermediaries: A terminology framework. In *Proceedings of the 7th International Conference on Theory and Practice of Electronic Governance*, ICEGOV '13, pages 330–333, New York, NY, USA, 2013. ACM.

[69] Christopher T. Marsden. *Net Neutrality: Towards a Co-Regulatory Solution*. Bloomsbury Academic, 2010.

[70] Viktor Mayer-Schönberger. *delete: The Virtue of Forgetting in the Digital Age*. Princeton University Press, 2009.

[71] Viktor Mayer-Scho?nberger and Kenneth Cukier. *Big data: A Revolution that Will Transform how We Live, Work, and Think*. Eamon Dolan/Houghton Mifflin Harcourt, London, 2013.

[72] Evgeny Morozov. *The Net Delusion: How Not to Liberate The World*. Allen Lane, 2011.

[73] Lev Muchnik, Sinan Aral, and Sean J. Taylor. Social influence bias: A randomized experiment. *Science*, 341(6146):647–651, September 2013.

[74] Milton L. Mueller. *Ruling the Root: Internet Governance and the Taming of Cyberspace*. MIT Press, 2004.

[75] Milton L. Mueller. *Networks and States: The Global Politics of Internet Governance*. The MIT Press, 2010.

[76] Johanna Niesyto. Virtualized campaigning for europe: Towards reinvigoration of european public sphere(s)? In Sigrid Baringhorst, Veronika Kneip, and Johanna Niesyto, editors, *Political Campaigning on the Web*, pages 93–120. Transcript Verlag, 2009.

[77] Andrew Odlyzko. Network neutrality, search neutrality, and the never-ending conflict between efficiency and fairness in markets. *Review of Network Economics*, 8(1):40–60, 2009.

[78] OECD. *Lutter contre l'érosion de la base d'imposition et le transfert de bénéfices*. Organisation for Economic Co-operation and Development, Paris, February 2013.

[79] OECD. *Plan d'action concernant l'érosion de la base d'imposition et le transfert de bénéfices*. Organisation for Economic Co-operation and Development, Paris, July 2013.

[80] ONI Team. Global internet filtering in 2012 at a glance, April 2012.

[81] Rufus Pollock. Is google the next microsoft? competition, welfare and regulation in internet search. *Review of Network Economics*, 9(4):76–104, 2011.

[82] Chris Preist, Dan Schien, Paul Shabajee, Stephen Wood, and Christopher Hodgson. Analyzing end-to-end energy consumption for digital services. *Computer*, 47(5):92–95, May 2014.

[83] Kristen Purcell, Joanna Brenner, and Lee Rainie. Search engine use 2012. Technical report, Pew Research Center's Internet & American Life Project, March 2012.

[84] Karl Rauscher. Writing the rules of cyberwar. *IEEE Spectrum*, 50(12):26–28, 2013.

[85] Priscilla M. Regan. United states. In James B. Rule and Graham Greenleaf, editors, *Global Privacy Protection: The First Generation*, pages 107–140. Edward Elgar Publishing, 2008.

[86] Andrea Renda. Neutrality and diversity in the internet ecosystem. SSRN Scholarly Paper ID 1680446, Social Science Research Network, Rochester, NY, August 2010.

[87] Howard Rheingold. *The Virtual Community: Homesteading on the Electronic Frontier*. MIT Press, 2000.

[88] Lawrence G. Roberts. A radical new router. *IEEE Spectrum*, 46(7):30–35, July 2009.

[89] Ira S. Rubinstein. Privacy and regulatory innovation: Moving beyond voluntary codes. {Public Law & Legal Theory Research Paper Series} {Working Paper No. 10-16}, New York University, School of Law, March 2010.

[90] Paul A. Samuelson. The pure theory of public expenditure. *The Review of Economics and Statistics*, 36(4):387–389, 1954.

[91] Jen Schradie. The digital production gap: The digital divide and web 2.0 collide. *Poetics*, 39(2):145–168, April 2011.

[92] Florian Schuett. Network neutrality: A survey of the economic literature. *Review of Network Economics*, 9(2), 2010.

[93] Alicia Solow-Niederman, Kevin Tsai, Andrew G. Crocker, and Jonathan Zittrain. Final briefing document: Public networks for public safety: A workshop on the present and future of mesh networks. Berkman center research publication no. 2012-12, harvard public law working paper no. 12-22, Harvard University - Berkman Center for Internet & Society, March 2012.

[94] Beth Stephens. The curious history of the alien tort statute. *Notre Dame Law Review*, 89(4):1467–1543, 2014.

[95] W.Ph. Stol, H.K.W. Kaspersen, J. Kerstens, E.R. Leukfeldt, and A.R. Lodder. Governmental filtering of websites: The dutch case. *Computer Law & Security Review*, 25(3):251–262, 2009.

[96] Cass R. Sunstein. *Republic.com 2.0*. Princeton Univ Press, New Jersey, 2007.

[97] Ali Sunyaev and Stephan Schneider. Cloud services certification. *Commun. ACM*, 56(2):33–36, February 2013.

[98] Douglas Thomas. *Hacker Culture*. University of Minnesota Press, 2002.

[99] Marcelo Thompson. In search of alterity: On google, neutrality and otherness. In Aurelio Lopez-Tarruella, editor, *Google and the Law*, volume 22 of *Information Technology and Law Series*, pages 355–403. T. M. C. Asser Press, 2012.

[100] Hayley Tsukayama. How closely is google really reading your e-mail? *The Washington Post*, August 2014.

[101] Hal R. Varian. The economics of internet search. *Rivista di politica economica*, 96(11/12):177–191, 2006.

[102] Fernanda B. Viégas, Martin Wattenberg, and Kushal Dave. Studying cooperation and conflict between authors with history flow visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 575–582, Vienna, Austria, 2004.

[103] Perttu Virtanen and Vili Lehdonvirta. A new frontier in digital content policy: Case studies in the regulation of virtual goods and artificial scarcity. *Policy & Internet*, 2(3):7–23, October 2010.

[104] Patrick Vogl and Michael Barrett. Regulating the information gatekeepers. *Communications of the ACM*, 53(11):67–72, 2010.

[105] David S. L. Wei, San Murugesan, Sy-Yen Kuo, Kshirasagar Naik, and Danny Krizanc. Enhancing data integrity and privacy in the cloud: An agenda. *Computer*, 46(11):87–90, 2013.

[106] Aaron Weiss. Net neutrality?: There's nothing neutral about it. *netWorker*, 10(2):18–25, June 2006.

[107] Florian Wettstein. The duty to protect: Corporate complicity, political responsibility, and human rights advocacy. *Journal of Business Ethics*, 96(1):33–47, September 2010.

[108] Kim Willsher. Rwanda former spy chief pascal simbikangwa jailed over genocide. *The Guardian*, March 2014.

[109] Brian Womack. Google is projected to expand lead in online-ad market. *Bloomberg*, June 2013.

[110] Tim Wu. Network neutrality, broadband discrimination. *Journal of Telecommunications and High Technology Law*, 2:141–178, 2003.

[111] Edward Wyatt and Noam Cohen. Comcast and netflix reach deal on service. *The New York Times*, February 2014.

[112] Christopher S. Yoo. Toward a closer integration of law and computer science. *Communications of the ACM*, 57(1):33–35, January 2014.