FREEDOM, AUTHORITY AND KNOWLEDGE ON LINE: THE DICTATORSHIP OF THE ALGORITHM[1]

Gianluigi Fioriglio
University of Rome "Sapienza"
gianluigi.fioriglio@uniroma1.it

Recibido: octubre de 2015
Aceptado: noviembre de 2015

Keywords: Search engine, search neutrality, web neutrality, search algorithm, liability.

Abstract: Web search engines are a very important mean to fight information overload's consequences, but this makes them the gate to digital information of any type and purpose. The cyberspace is accessed through such complex and automated tools: software agents execute secret and complex algorithms and make information easily reachable or hidden, but anomalies and bugs may have serious consequences. Private entities provide such services, on a global (e.g. Google), or local but relevant scale (e.g. Baidu). After the Google Spain case, literature on the right to be forgotten is growing. This paper aims at go further, investigating both the right to access on line information and to be correctly and neutrally indexed by web search engines. The law must regulate this topic, making them work in a neutral and non-discriminatory way, even if they work cross-borders and are private subjects. Otherwise, the Information Society will even more controlled by the dictatorship of the algorithm.

1. Introduction

Legal literature on web search engines[2] is growing, especially after a landmark decision by the European Court of Justice: the judgement in *"Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González"* (13 May 2014, "Google Spain case"; European Court of Justice 2014). It held that the right to be forgotten applies to search engines; in particular, in the European Union it is possible to ask for delisting of web site links from the SERP (Search Engine Results Page). Thus, even if a web page is on line and cannot be legally shut down, it may not be indexed if certain conditions are met. The ECJ judgement is a landmark decision because it makes a division among what can be easily found on the web, on the one

1. This is the revised version of the paper presented at XXVII World Congress of the IVR - International Association for the Philosophy of Law and Social Philosophy (Special Workshop on "From Net Neutrality to Net Profitability? Law, Politics and the Internet"), Washington DC, July 2015.

2. This paper uses the expression "search engine" to indicate automated web search engines.

hand, and what can hardly be found, on the other hand. A web content may be published, but it does not necessarily imply that everyone may rapidly find it, answering to a reasonable expectation of privacy.

This aspect is crucial in the Information Society, because search engines are the primary gatekeepers to digital information. Another aspect is perhaps even more important, but it needs deeper investigation: is there a legal obligation related to web site indexing? In other words, is there a right to be indexed?

In fact, if we take into account only what happens when a search engine responds to a query (as in the case of the right to be forgotten), we forget what happens before: how does a search engine build its SERP? How does it decide what pages can be shown and their ranking?

This is not only an antitrust matter, because it relates to all subjects whose content is published on the web and indexed automatically by search engines. "In an environment where consumers are no longer passive receivers of information, but increasingly active contributors to the information ecosystem, access also concerns the (controversial) debate about the entitlement of users (as creators) to be integrated into search indexes and ranking lists, or at least the possible remedies against discrimination in the indexing or ranking processes" (Gasser 2006: 232).

Today, a search engine can decide what can be accessed on the web; it can hide and/or rank information making it hardly be found ("By controlling the communication infrastructure of the Internet, they have become information gatekeepers"; Laidlaw 2008: 114[3]). It is fair: everybody uses a search engine to find something. Nevertheless, it is also fair to ask for *search* (or *web) neutrality*.

Before investigating this aspect, it is useful to make some preliminary observations also from the methodological point of view in order to clarify the perspective of this paper. Search engines are the key information retrieval systems of the Information Society and this field is not new for legal informatics[4]. By adopting such approach, law can be a meta-technology, as argued by Ugo Pagallo with reference to the "laws of robots". In particular, he proposes "to approach the laws of the law establishing the conditions of legitimacy for the design, production, and use of robots, conceiving the law as meta-technology, *i.e.*, as a means to govern other technological means" (Pagallo 2013: 10). Thus, "once such tecnique regulates other techniques and, moreover, the process of technological innovation, we may accordingly conceive the law as a meta-technology" (Pagallo 2013: 11).

With particular reference to search engines, it has to be considered that the cyberspace actually is a complex digital maze in which it is possible to retrieve any information stored. As the library of Babel, it may contain the solution to many problems (Amato Mangiameli 2000: 36). A search engine plays a key role: it is not just an information retrieval system, but

3. Moreover, "Users have become dependent on search engines, viewing them as authoritative and reliable. Search engines have become the tools through which the democratic potential of the Internet can be advanced or hindered" (Laidlaw 2008: 145).

4. Legal informatics always studied information retrieval systems and their legal issues. Technological advancements led to growing studies in the field of artificial intelligence in the legal field (including, but not limited to, legal reasoning and automated application of the law).

rather it enables everyone to find any information, avoiding getting lost in billions of webnodes. It drives its users through paths that are not predetermined but instead dynamically generated in execution of many secret algorithms. The aforementioned paths are made of hyperlinks; hypertext provides a means of non-sequential reading and thus it mimics "the brain's ability to store and retrieve information by referential links for quick and intuitive access" (Fiderio 1988: 237). However, automated activities of search engines not only modify human interaction in cyberspace, but also control the information flow and build, in whole or in part, the digital identity of any person. In fact, despite the fact that the Internet and the Web are decentralized, some sites and services become essential nodes and shape the information society at a whole. As Yochai Benkler notes, "The emerging patterns of Internet use show that very few sites capture an exceedingly large amount of attention, and millions of sites go unnoticed. In this world, the Babel objection is perhaps avoided, but only at the expense of the very promise of the Internet as a democratic medium" (Benkler 2006: 10).

This is due to many factors, including the use of search engines. Thus, being at least listed in the SERP, firstly, and to be highly ranked, secondly, can lead to a virtuous or vicious circle.

It is a virtuous circle if (A) the search engine works properly, (B) listing is carried out in compliance with neutral rules, and (C) operations are correctly performed.

It is a vicious circle if the above conditions are not met (for example, a rule is neutral and a content should be shown in the SERP, but a content is omitted in the SERP due to a software bug). This involves a series of questions: does a right to be indexed (or listed) exist and, if so, to be *correctly* indexed? Is there a legal relationship between search engine providers and content owners? What liability rule should be applied? How can neutrality and correctness be checked? Is it possible to impose specific legal duties to search engine providers and, if so, under what conditions?

Although it is difficult to answer these questions, it is necessary to investigate them deeply because search engines are the most important gatekeepers to digital information (texts, images, music, videos, etc.) and they usually provide other services linked to search.

Any search engine should put order among information, to avoid its users to be overwhelmed by them. However, today we risk being overwhelmed by search engines if the law is unable to regulate them.

2. Search engine as the Information Society fundamental expert system

Search engines work both as generalist (e.g., news) and as specialized gatekeepers (e.g., academic papers) to digital information and are a perfect means not only to get informed, buy products, and find resources, but also to sell advertising spaces. Moreover, vertical searches may be integrated within generic web search.

In all these cases, users choose between results by selecting one or more hyperlinks (links) listed in the SERP. The more

a content is highly ranked, the more easily users will select it.

Each search engine responds automatically to queries by showing results listed in order of relevance and pertinence. This is done by executing hundreds of top-secret algorithms. Only generic information upon them is public and thus it is not possible to evaluate if a search engine works in a neutral and fair way; in other words, it is very difficult, and often impossible, to check if the search engine reasoning may be deemed as neutral and correct. This is a matter of particular importance from the legal informatics perspective. Given the secrecy of the IT platform, it is only possible to make *ex post* deductions based on the analysis of the results and on the comparison with the content indexed or indexable. As a result, it is hard to challenge its results and understand if it is censoring specific web sites or results, or giving unfair advantage to its customers (i.e. who pays for advertising services): "opaque methods of ranking and rating online entities make it difficult for those who feel (and quite possibly are) wronged to press their case" (Pasquale 2011: 382).

However, one author argues that legislators could censor the "historically unprecedented free search tools that help create enormous social value. It would be easy for regulators, even well-intentioned ones, to inadvertently eliminate some of this value through misregulation. That outcome is worth fighting against" (Goldman 2011: 109-110). Nevertheless, these "search tools": (a) are not really free (usually payment is done through personal information); (b) may (inadvertently or not) eliminate or censor content; (c) have social value but it does not imply that their providers should be exempted from any liability or accountability (should any social-oriented activity imply no accountability? For instance, should an ambulance driver have no liability even if he/she kills someone during his/her work?); (d) earn money indexing on line content not provided by them (sponsored search is related to organic search, whose content is provided by third party websites).

Search engines have to deal with many difficulties, ranging from the intrinsic ambiguity of syntax, semantics and grammar, to the evolution of natural language and of computer code aimed at getting a better ranking in the SERP. In particular, natural language may not only evolve in relation to specific domains of knowledge but also be adapted to the Information Society and the Internet. One example is given by the growing number of SEO (Search Engine Optimization) services: content is written taking into account both human and automated readers. More generally, language tends to simplify and the art of rhetoric seems to be replaced by the wise use of keywords, hashtags, abbreviations, and so on. This makes the natural language poorer and may damage conceptualization and reasoning.

In this framework, the search engine has the task to find and understand information, whether available in natural or computer language, making them easily accessible to its users. The ability in carrying out these tasks, and therefore search engine efficiency, may lead to the success or failure of a search engine service. However, if a provider is so successful to achieve a dominant or monopolistic position certain legal duties may be, or are, imposed to it. Before investigating further this aspect, it is worth mentioning that since the early stages of computerization it was noticed that computers could store large amounts

of data and use them as inputs for algorithm elaborations; *inter alia,* this made computer science focus also on the study of archives and databases management and use (Sartor 1990: 7).

With regard to search engines, the input stage consists in the acquisition of data not only by the so-called spiders or crawlers that travel through the Web, but also from other services in order to customize the SERP: search history, interactions with social networks, geolocation, calendar, etc.

A search engine, however, cannot be considered only as a complex database with a user interface that makes it possible to query it: it is an expert system instead. In fact, it is clear that a search engine runs tasks that require human intelligence, because it works both as an archivist and as a librarian and thus works as an expert (even if language and text understanding is a common sense task; Sartor 1990: 15). A search engine tracks, analyses, catalogs, and makes available links to, digital information of different types in response to users' queries. Therefore, it makes many choices and must faces many difficulties, due to the diversity of information and of their sources. It goes well beyond the concept of database because such system must increasingly be automated and "intelligent". It makes decisions and often seems willing to replace who queries it, going beyond a simple answer to a question through the provision of additional replies: search suggestions or automated conversion between currencies are good examples to show how a search engine service is now going beyond the mere context of information retrieval.

From another perspective, it seems to be confirmed what argued by Lawrence Lessig. He stated that "code writers are increasingly lawmakers. They determine what the defaults of the Internet will be; whether privacy will be protected; the degree to which anonymity will be allowed; the extent to which access will be guaranteed. They are the ones who set its nature. Their decisions, now made in the interstices of how the Net is coded, define what the Net is" (Lessig 2006: 79). Therefore, they define its real – albeit immaterial – structure to be used by present and future generations. As stated by Palombella following Vico and Dilthey, generally speaking, any individual is housed in a novel written by others in which may write his/her chapter (Palombella 2007: 399). In the information society, however, it becomes more difficult because part of the novel is now written in an automated and sometimes unpredictable manner from some tools created by man: search engines.

3. Automation and subjectivity: some thoughts

Before investigating not only search neutrality but also the consequences of errors and decisions made by search engines, some notes on automation and subjectivity may be useful.

Today, a search engine affects many aspects of everyday life. Autonomous software agents are a key component: they retrieve and analyze information, and answer to users' queries. They are very important in the philosophical perspective of legal informatics also because semi-legal relations are established among them (Faralli 2012: 81); their actions produce consequences for information society at a whole.

We must distinguish between electronic agents that operate in the material and the virtual world. Robots act in the former and they have an artificial body. Perceptions and actions are related only to computer messages and data in the latter (Sartor 2009: 19; also in Id. 2009); the same applies to search engines.

Several scholars denied the possibility of attributing legal personality to software agents because no rule would allow for this. Liability is thus traced in the sphere of the user or of its manufacturer *a fortiori* (Biasiotti, Romano, and Sagri 2002).

The assessment of the consequences of intelligent agents is much more difficult when they have no body, because while actions done by a robot result in acts that can be perceived, actions done by software agents are necessarily intangible and are invisible unless not displayed by an output device. However, search engines are an even more difficult case. Given the secrecy that hides every line of their code, it is only possible to make deductions based on their results and comparisons with the content indexed or indexable. Indeed, it is very difficult, if not almost impossible, to perceive any incorrect or illegal actions; it may happen only when their consequences are so obvious that cannot be hidden even if illegal operations are carried out. "The search algorithms are protected as trade secrets, and the reasons for manual manipulation of rankings, particularly in any given case, are not publicly or privately revealed. This results in a quagmire wherein businesses rely on the search results, yet have no access to understanding changes in such results, even when the changes have a marked impact on the company's sustainability" (Laidlaw 2008: 137).

Several cases may occur: for example, a software agent may decide to exclude most of the pages of a site from the SERP without a valid motivation or, hypothetically, due to errors in the design and/or implementation of one or more algorithms (and/or of their interactions).

Two main issues arise. On the one hand, it is necessary to understand who should be liable for the actions of the software agent. On the other hand, it must be understood if the law can impose a duty upon the service provider and, if so, what characteristics it should have; furthermore, it is necessary to understand if rules in force are sufficient or if lawmakers must act. It also worth mentioning that such issues have a wide scale and apply to several countries because search engine service is usually provided globally.

The second issue will be dealt with in paragraph 5, while the first can be approached making reference to Sartor's theory of software agents' liability. "It seems that the guardian's liability for the action of a SA cannot be grounded only upon the fact that a damage could be foreseen according to the "normal" laws of nature (or of technology). We need rather to consider whether the SA intentionally or negligently produced the damage. If we have indeed to draw this conclusion, then the liability of the user of a SA would be similar, rather than to liability of a custodian of a thing, to vicarious liability (the liability of the employer for the employee). This form of liability is not based upon the fact that the employer could foresee the behaviour of the employee, but rather on the fact that the employee accomplished a tort, when acting in the course of the employment" (Sartor 2006: 22). After all, software programmers determined the system's will and thus the effects of the actions of the latter must still be referred to

its manufacturer (Borruso 1988: 253-255; see also Id. 1978).

Eventually, it is necessary to highlight another issue related to subjectivity: the impact of automation on subjectivity and personal identity building. Search engines draw the profile of a person by means of automated processing of data retrieved on the Internet. In such cases, one could argue that the search engine would be an isomorphic system because it knows only information provided by third parties and it does not provide new knowledge. Indeed, information is dynamically processed and built by means of software agents even in the SERP and this is even more true when additional services are offered, such as automated suggestions. This led to several proceedings[5], in which the plaintiff sued the search engine provider alleging offensive expressions suggested with the name and surname of a person[6], or the name of a company or an association[7].

There is another issue. As search engines become more complex, they also may provide content and services within each SERP depending on the query (e.g., traffic, weather predictions, calculations, currency conversions, etc.) and it can impact on organic search (and the neutrality of search engine providers that could advantage their services using their position).

4. Algorithms and software agents: choices, decisions, and errors of the search engine

Users who query a search engine expect it to respond correctly, i.e. listing the results in order of relevance[8]. If a legitimate expectation to the above-mentioned correctness may be argued, how should it be shaped? Would it be sufficient to execute algorithms without software bugs? Or could nothing be claimed because no relevant legal provision could be expected by the service provider?

To respond appropriately it is necessary to quickly remember the concept of algo-

5. "Based on comparable assumptions, courts have upheld claims for inducement of copyright infringement and dismissed allegations of damage to reputation"; however, "no solid judicial trend can be assumed" (Karapapa and Borghi 2015: 263).

6. One good example is the "Bettina Wulff case" (settled in 2015): "Former German first lady Bettina Wulff has taken on Google over search terms that link to false rumors that she used to be a prostitute. The company argues that it generates such terms based on "objective factors," but it's not that simple. Google has suppressed undesirable results before in response to powerful lobby groups" (see Lischka 2012).

7. For example, the Tribunal of Milan uphold one claim for damage to reputation and ordered Google to remove the association of two words with one plaintiff's name when using the autocomplete feature (Tribunal of Milan, 23 May 2013, as appeal judgement of Tribunal of Milan 25 March 2013 – in which the allegation was dismissed).

8. Obviously, "Any public interest in relevancy must be realistic. Businesses must expect fluctuations in rankings, and not every website can be highly ranked, nor is every undesirable link going to be caught and removed. Relevance is hard to measure, because it is as much dependent on the users search terms as it is on the algorithm of the provider, and results can be manipulated by google bombing, search engine optimisation and the like. For relevance to have meaning, the key is consistency: consistency in algorithms, consistency in decision-making regarding any manipulations of search results, and consistency in the values that drive such manual manipulation" (Laidlaw: 2008: 139).

rithm. Indeed, a computer can perform only tasks that are reducible to an algorithm: it is a sequence of prescriptions or 'instructions' that indicates precisely and unambiguously the steps to take to resolve correctly a certain problem, starting from certain information, in a finite time.

In the case of software agents, their typical characteristics (proactivity and reactivity, behavioral flexibility, persistence over time, mobility, intelligence, ability to communicate) lead to difficulty or inability to predict their behavior. In fact, the combination between the complexity of software agents and the richness of the environment in which they operate, makes it very difficult, if not impossible, to predict their behavior accurately. Efforts to predict them could "contradict the very reason for using an SA: delegating cognitive tasks" (Sartor 2006), but this may lead to more risks if they behave illicitly.

However, if we move from generic software agents to those used for providing search engine services, we see not only how problematic may be the profile of their choices and decisions, but also how easily mistakes and anomalies may occur. This has a severe impact on the growing debate on search (and web) neutrality. Before going deeper into these profiles, it is necessary to emphasize that search engine providers are in a role in which many benefits and risks coexist. It is objectively difficult to process billions of information properly by returning relevant results in response to billions of queries[9]. There is too much information that would make impossible finding them without using automated means; thus, software agents are the delegates to carry out an activity that, by its very nature, needs be automated because there is a practical limit due both to the excessive number of information and their rapid growth. However, great care must be taken. Vittorio Frosini noticed that computers act as an artificial eye built by people for people (Frosini 1988: 75), but many times that eye is now provided by search engines: they decide what information is visible and therefore they may control and censor them.

As mentioned above, providers must delegate the execution of this service to software agents that operate in an environment far too rich in information to be analyzed; furthermore, it is often optimised for search engines using SEO techniques that become more extreme as time goes by. These automated tools are based on highly complex algorithms to provide relevant and useful results to a plurality of persons: how may they understand what is actually sought by each user and how to interpret his/her intentions?

In any case, the engine must make choices that involve several evaluations affecting not just the scope of mere information retrieval (now automated) (phase 1) but rather that of their analysis (phase 2) and dynamic generation of each SERP in response to each query (phase 3). While current technologies do not meet particular obstacles in carrying out information retrieval operations (phase 1), it is still extremely difficult make a system evolve from a purely formal knowledge to a substantial one. Although technological advancements are well known, much has yet to be done with regard to phases 2 and 3. This is due to the fact that

9. The Google case is emblematic: the size of its index amounts at more than 100,000,000 of gigabytes, it customizes increasingly search results based on a multiplicity of parameters, it is used in more than 90% of web searches within the European Union (where it is a *de facto* monopolist), and so on.

modern search engines perform syntactic operations and semantic analysis of texts running multiple algorithms, in addition to contextualize them within the framework of cyberspace and therefore to assess their "popularity" using special algorithms (generally referred to the quantity and quality of links to a particular resource). A fundamental activity is therefore related to the analysis of each information and, in particular, of text trying to hopefully ensure neutrality at the same time. However, even the interpretation of a text may not be a neutral and objective activity, because interpretation requires a prior understanding, as stated by Gadamer. Thus, neutrality may be deemed as a utopia; if neutrality cannot be ensured, exemption from liability should never be claimed and thus service provider could not benefit from it.

According to Eric Goldman, search neutrality cannot be achieved: "the term "search neutrality" implies the existence of neutral search engines, but those are entirely mythical. Every search engine design choice necessarily and unavoidably reflects normative values. Thus, the term "search neutrality" implies a Platonic ideal of a search engine that cannot be achieved. Naturally, then, Google's practices fail to conform to this Platonic ideal, but so does every other search engine in the real world" (Goldman 2011: 107). Search engine service providers would never be held liable because neutrality could not be reached; in addition, it may also happens because some rules may be interpreted to give them a (too) wide protection (Supreme Court of California 2014). In addition, it could be argued that, in this task, the consequences of the actions of the agents are unpredictable because not only texts but also queries are not foreseeable and are beyond any control of the provider. Moreover, one could argue that there is no right to be indexed and that a search engine provider would not have any legal obligation towards users who query the search engine as well towards subjects whose information is analyzed and indexed.

However, search neutrality is not a Platonic ideal and the whole issue can be analyzed in another perspective: such reasoning would exempt an entire category of subjects from any liability making them *legibus solutus* in the provisioning of a crucial service of the Information Society; furthermore, such service is provided for profit by private entities.

Thus, it is possible to answer to the mentioned opinions arguing, firstly, that free speech protection cannot be taken on an extreme level, making a subject always exempt from liability despite any possible mistake and that many legal systems balance free speech and *neminem laedere*. Furthermore, until now free speech has been crucial in exempting search engine providers from liability in the U.S., but the European framework is more complex and it may shape the near future, as the Google Spain case shows.

Secondly, even in the case of unpredictability, any damage must be compensated from the provider who controls the service (as already happens with reference to several regimes of liability: for instance, custodian's liability (Sartor 2006) or strict liability (Pagallo 2009).

Thirdly, a right to be listed, and to be correctly listed, in organic search may be argued if the concept of "correctness" is investigated; in particular, any search engine provider claims to index and rank content objectively and neutrally, without discriminating anyone in the organic

search, while sponsored search obviously works in a different way. Thus, such provider accepts autonomously a legal obligation to answer to users' queries in a neutral and non-discriminatory manner. Only a search engine provider that does not have a dominant or monopolistic position may choose to provide a non-neutral service, but this business model is not used today and it should be make clear to its users.

It is necessary to emphasize that any search engine depends on the content that it indexes, because organic search is the "Trojan horse" to sell advertising spaces used in the sponsored search (i.e., the results shown together with the SERP). After all, even sponsored search depends on the SERP composition.

From a legal point of view, there are no problems in the structure of the aforementioned model if neutrality and impartiality are guaranteed in the organic search and the sponsored search is not deceptive. However, "The problem would be with search engines that mix the two strategies and hide the mix, or with a monopolistic search engine" (Castells 2009); moreover, search engines contribute to shape "the black box society": this metaphor refers both to recording systems and ones whose workings are mysterious (Pasquale 2015).

Therefore, it is crucial to ensure the neutrality of the choices and decisions of the search engine (and therefore of the neutrality of indexing tasks), especially when its provider has a dominant or monopolistic position, with obvious consequences from a legal point of view. The first issue will be analyzed in paragraph 5 and it can be anticipated that the only way to ensure neutrality consists in the possibility to evaluate the reasoning made by the search engine. This problem is connected to the second, because it is going to affect a private space that is generally not challenged by third parties or by a court unless there are no clues or evidences of an illicit behavior of the engine. Nonetheless, the legal approach to the issue should change if an entity has a dominant or monopolistic position, especially if it is reasonable to believe that the same entity may abuse or is abusing its position. In such case, competition law plays a primary role and antitrust authorities (and sometimes the courts) are the only power capable to oppose to an economic and technological power. In fact, if a search engine has a dominant or a monopolistic position, it also has enormous powers, not only economic, because it has the power of techno-exclusion. Those who control the gateway to the Cyberspace have a power that is paradoxically real albeit intangible. It may not only exclude at will but also build a digital profile suitable of damaging a person or an entity; furthermore, it may orient opinions and beliefs on certain issues simply showing certain links in prominent positions and hiding or low-ranking others.

The last profile, in particular, highlights the issue of so-called SERP ranking. Its utmost importance is well known: if a web site is low ranked, probably it will not be found by users. Moreover, in many cases the search engine decides that specific pages are not shown by default even if they have been indexed: this is the case of omitted results.

Besides, the search engine may decide to match one or more terms to others. However, it should be done in an automated, objective, and neutral way to avoid that the search engine service may translate into a dictatorship in the control

of information. It can even lead to their partial censorship and therefore to affect fundamental rights. For example, many web pages can be reached difficultly outside of a specific country if queries are done by another country, because the search engine assumes that they are of no interest to them (this is done by executing an algorithm that locates the user using various parameters such as IP address and browser language). It is legitimate to ask, therefore, whether the extreme customization of web search can bring back the same concept of space that was made obsolete and evanescent by the Internet and the Web, and one may wonder whether this is good or bad.

Eventually, another issue must be added to this framework: the consideration of search engine errors and of their effects. A computer is by no means infallible, above all due to its software rather than due to its hardware (but both can be defective). Daily life, however, is increasingly pervaded and controlled by a variety of software, which also show to be likely to make clear mistakes. This does not only refer to ranking manipulation for political activism (as in the case of "Google bombing": e.g., in 2004 the query "miserable failure" returned the biography of the U.S. President George W. Bush as the first result), but also to bugs and malfunctions that can affect the correct operation of the search engine. As a result, a search engine can wrongly rank web pages or even hide them because it has decided so.

In such cases, is the error due to an algorithm or to its implementation? Or may it rather arise in the interaction among a multiplicity of algorithms that lead to unexpected results or abnormal behaviors? Were errors or abnormal behaviors abstractly predictable?

There are no doubts about the liability of the service provider for any harmful consequences, but it is clear that the overall analysis of these issues goes far beyond the scope of statutory damages and touches a fundamental aspect of contemporary society because it is related both to access to an essential infrastructure (the Web) and to freedom of information. In fact, "In a democratic society, those who control access to information have a responsibility to support the public interest. By dint of their power over such an important resource, these gatekeepers must assume an obligation as trustees of the greater good. Indeed, barring some clear showing that they are bearing this burden voluntarily, government should impose it upon them" (Shapiro 1999: 225). In fact, the idea that search engines are merely 'businesses' or 'promotional services for website owners' is untenable. If information is becoming a critical commodity in modern society, then such bodies that manage access to information, that are tools for public discourse and democracy, should be accountable to the public. When the structure for access shapes meaning for its users, and influences public opinion, this need for a public-interest obligation is magnified" (Laidlaw 2008: 137).

5. Neutrality or dictatorship of the algorithm?

Search engine providers may control and censor digital information but the legal framework is still focused on liability exemption and de-listing of certain contents (see European Court of Justice 2014). In the U.S. context, the legal debate starts from the choice of considering a search

engine provider such as a "conduit" or an "editor": "on search bias claims […], the conduit theory is a recipe for regulation, while the editor theory offers a First-Amendment get-out-of-jail-free card. Google instead ought to be an active and opinionated *editor*, sifting through the Internet and using expert judgment to identify the important and the interesting. These two theories form the rhetorical backdrop to the ongoing legal battles over search. But when the issue is defamation, the conduit theory holds Google harmless for the sins of the websites it unknowingly connects users to, while the editor theory calls down the vengeance of the heavens on Google for its editorial decisions" (Grimmelmann 2014: 871).

However, once again it is necessary to emphasize that they are the most important intermediaries between online information and users. In fact, the provision of this service provides a huge power: control of information through the faculty to make them available and to decide how easily make them available. There is no doubt that the algorithms are increasingly sophisticated and that current techniques are extremely fine, making it possible even to customize each SERP on the basis of many criteria, not just geographical, even with *ad hoc* automated customization.

To date, search engine providers are practically exempt from liability, almost as if the algorithms used for providing the service were not realized by the same entities who earn relevant profits from them.

As anticipated, owners of web sites and data indexed have the right that this operation is carried out correctly because they provide the content that the search engine provider uses not only to provide such a service, but also to sell advertising space. In fact, without organic results, sponsored research would not make any sense because it can be reasonably assumed that nobody would use that specific search engine (the model of pay per click engines is not used anymore). In particular, given the central role of search engines in the Information Society, rules must be set to balance freedom of information and of economic initiative (and industrial property rights).

It is necessary to define policies aimed at protecting individuals and businesses harmed by automated decisions of search engines, going beyond the recent decision on the right to be forgotten. It is unacceptable that search engine providers are practically exempt from liability due to errors of their agents (so presumably due to errors in algorithms or abnormal behaviors resulting from their implementation and/or their interaction). In particular, control mechanisms should be implemented to verify the correctness of the work of software agents (such as, but not limited to, the ability to verify the reasoning). The secrecy surrounding such systems is extremely risky for the impossibility of protecting persons and business damaged by them. It is very difficult to prove intent or negligence of a service provider for the conduct of its digital agents, except in case of clear errors. Even in such case, absolute secrecy may make difficult to obtain the proof that a problem has been truly solved (e.g., not simply in its external manifestation in the SERP) or that hypothetically the SERP is built to promote its own services or its customers. Furthermore, "Particularly in a field as dynamic and complex as search, it may prove beyond the institutional competence of courts unable to deal with rapidly shifting business practices occluded by

trade secret protection. All these factors point toward the development of a public opinion in search, or a more regulatory approach, including teams of lawyers, engineers and programmers, that would complement existing litigation and competition" (Pasquale 2011: 405).

A paradox is clear. ICTs may make society more transparent also because a diffuse control upon power is more easily achievable. However, when the algorithm becomes the very foundation of the power exercised by a subject, as in the case of Google, and everything it is enveloped by the utmost secrecy, then we are really in front of the new version of *arcana imperii*, which protect not only the business activity, but also control life itself, directly or indirectly. In addition, the consultation of one or more profiles, or otherwise of any content type, now depends from algorithms that decide what can be shown in response to users' queries. Persons should not be subject to the dictatorship of the algorithm, in which the decision maker is replaced by automated procedures and the person disappears, turned into an object of uncontrollable powers (Rodotà 2014: 37-38). This is even truer when the person affected or excluded does not have the strength to make his/her voice heard, put in between the market and the technology. "In this market of markets, there is likely to be little incentive to ensure inclusion of these small markets and only a small cost (in loss of participation) for their exclusion" (Introna and Nissenbaum 2000: 177).

Therefore, ensuring neutrality in the provisioning of search engine service is a fundamental goal. Indeed, "Web-search mechanisms are too important to be shaped by the marketplace alone" (Introna and Nissenbaum 2000: 176)[10]. Emily B. Laidlaw proposes a framework of accountability for search engines' practices: "Three values should be present in such a framework: the value of relevant and unbiased search results; the value of a degree of transparency concerning algorithms and reasons for manual manipulation, and respect for the dignity of the users recognizing that how information is presented on indices can cause harm" (Laidlaw 2008: 145).

Legal issues are different and touch extremely delicate aspects, with particular regard to liability for the conduct of third parties, to the potential abuse of a dominant or monopolistic position, the protection of freedom, and the potential conflict with freedom of economic initiative of the service provider. Legislators cannot remain idle. It is true that the law (tries to) regulate cyberspace, with mixed fortunes. Nevertheless, the new technological challenge of avoiding the dictatorship of the algorithm has not been effectively taken.

10. "They provide essential access to the Web both to those with something to say and offer and to those wishing to hear and find. Our concern is with the evident tendency of many of the leading search engines to give prominence to popular, wealthy, and powerful sites at the expense of others. This they do through the technical mechanisms of crawling, indexing, and ranking algorithms as well as through human-mediated trading of prominence for a fee. As long as this tendency continues, we expect these political effects will become more acute as the Web expands. We regret this tendency not because it goes against our personal norms of fair play but because it undermines a substantive ideal –the substantive vision of the Web as an inclusive democratic space. This ideal Web is not merely a new communications infrastructure offering greater band– width, speed, massive connectivity, and more, but also a platform for social justice" (Introna and Nissenbaum 2000: 181).

This topic is likely to stimulate the debate between those who do not want any limitation (the service provider and those who believe that any restriction is an attack to the freedom of the net) and those who want to put severe limits (many subjects having different interests). The fragmentation of the opponents and the unity of providers are accompanied by a society in which individuals are usually not skilled enough to understand both legal and informatics profiles, being them complex and challenging.

However, the power resulting from the use of digital agents and the profits earned by using them are benefits that must be accompanied by accountability for their use, as well as by the need to ensure fairness and neutrality. A constant and careful consideration is therefore necessary to prevent search engine providers to consolidate their dictatorial and censorial powers; they provide a fundamental and essential tool for accessing the vast and growing amount of digital information, regardless of their form. However, any search engine depends on the content that it indexes, because the organic (free) search is a Trojan horse for the sponsored search (paid by advertisers) and for other services. This is perfectly legitimate, as long as neutrality, transparency and objectivity are guaranteed, while the presentation of sponsored results must not be misleading or intended to mislead users.

Moreover, organic search must work properly, to avoid excluding both businesses (that would be forced to pay for the sponsored search in order to appear) and people who use digital tools outside of their profession.

To ensure that the conditions set out above would be necessary, firstly, to impose a duty of transparency about the algorithms used and their concrete implementation (see Introna and Nissenbaum 2000: 181)[11]; secondly, to make possible the analysis of the reasoning that led a search engine to make its decisions. This does not mean the possibility to use the code by third parties, but only to study it and make sure that the neutrality of the search engine is effective and not only claimed. Any objections related to security and to the ability to exploit this knowledge to illegal or illegitimate purposes can be overcome by mentioning a well-known fact: the remarkable spread of open-source software commonly used in many areas and for many years without that security issues could be considered more serious than the ones of proprietary software.

Moreover, the person cannot be overwhelmed by the entity or the entities who control the digital tools. In 1964 Norbert Wiener said that "the future offers very little hope for those who expect that our new mechanical slaves will offer us a world in which we may rest from thinking. Help us they may, but at the cost of supreme demands upon our honesty and our intelligence. The world of the future will be an ever more demanding struggle against the limitations of our intelligence, not a comfortable hammock in which we can lie down to be waited upon by our robot slaves" (Wiener 1964: 69).

11. According to Emily B. Laidlaw, "the negative consequences of disclosure of algorithms on the market as sufficiently concerning that blanket transparency should not be the solution. Further investigation is required to determine the potential market fallout of full algorithmic transparency. However, disclosure should be required regarding manual manipulations" (Laidlaw 2008: 139). In addition, it may also be argued that disclosure should be imposed at least when errors and malfunctions are reasonably clear.

However, new computer systems (*rectius*, those who control them) sometimes bring persons not to think. So, provocatively and almost as a joke, Larry Page (co-founder of Google) stated that the objective is to ensure that Google can answer a question before it is asked, while Hal Varian (Google's chief economist) has basically stated that, albeit partially, a similar function is already offered through the convergence of certain services[12].

An additional contribution may result from constant cooperation in a multidisciplinary development of search engine. The development of search engines should be made with the contribution, in particular, of philosophers and jurists who may help ensuring respect for fundamental freedoms and rights (Introna and Nissenbaum 2000: 181). Moreover, search engines should be built following a neutrality by design principle.

In conclusion, it is clear that we cannot focus solely on reflections on de-listing and on the right to be digitally forgotten, but instead investigate the matter further. An obligation to properly index on line content should be expressly regulated. From a legal point of view and on the basis of the arguments presented in this paper, a general duty to correct indexing is already in force; search engine service providers must act in a neutral and non-discriminatory way. This obligation becomes even more compelling if such provider has a dominant or monopolistic position, so certain limitations may be more easily imposed thanks to antitrust laws.

12. "We all thought he was joking but Larry's vision has been realized by Google Now, an application that runs on Android phones. One day my phone buzzed and I looked at a message from Google Now. It said: "Your meeting at Stanford starts in 45 minutes and the traffic is heavy, so you better leave now." The kicker is that I had never told Google Now about my meeting. It just looked at my Google Calendar, saw where I was going, sent my current location and destination to Google Maps, and figured out how long it would take me to get to my appointment given current traffic conditions. Some people think that's the coolest thing in the world, and others are just completely freaked out by it. The issue is that Google Now has to know a lot about you and your environment to provide these services. This worries some people. But, of course, I share highly private information with my doctor, lawyer, accountant, trainer, and others because I receive identifiable benefits and I trust them to act in my interest" (Varian 2013).

Bibliography

Amato Mangiameli, A.C. 2000. *Diritto e Cyberspace. Appunti di informatica giuridica e filosofia del diritto*. Torino: Giappichelli.

Biasiotti, M.A., Romano, F., Sagri, M.T. 2002. "La responsabilità degli agenti software per danni prodotti a terzi." *Informatica e diritto*, 2: 157-167.

Benkler, Y. 2006. *The Wealth of Networks. How Social Production Transforms Markets and Freedom.* New Haven and London: Yale University Press.

Borruso, R. 1978. *Civiltà del computer*. Milano: Ipsoa.

Borruso, R. 1988. *Computer e diritto*. Milano: Giuffrè.

Castells, M. 2009. *Communication Power,* New York: Oxford University Press.

European Court of Justice. 2014. *"Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González"*, case C131/12, Judgement of the Court (Grand Chamber) of 13 May 2014.

Faralli, C. 2012. *La filosofia del diritto contemporanea*, Roma-Bari: Laterza.

Fiderio, J. 1988. "A grand vision." *Byte*, 1988, 10: 237-244.

Frosini, V. 1988. *Informatica diritto e società*, Milano: Giuffrè.

Gasser, U. 2006. "Regulating Search Engines: Taking Stock and Looking Ahead." *Yale Journal of Law & Technology*, 8: 202-234.

Goldman, E. 2011. "Revisiting Search Engine Bias." *William Mitchell Law Review*, 1: 96-110.

Grimmelmann, J. 2014. "Speech Engines." *Minnesota Law Review*, 98: 868-952.

Karapapa, S., Borghi, M. 2015 "Search engine liability for autocomplete suggestions: personality, privacy and the power of the algorithm." *International Journal of Law and Information Technology*, 23: 261-289.

Introna, L.D., Nissenbaum, H. 2000. "Shaping the Web: Why the Politcs of Search Engines Matters." *The Information Society*, 16: 169-185.

Laidlaw, E.B. 2008. "Private Power, Public Interest: An Examination of Search Engine Accountability". *International Journal of Law and Information Technology*, 1: 113-145.

Lessig, L. 2006. *Code. Version 2.0*, New York: Basic Books.

Lischka, K. 2012. "Blaming the Algorithm: Defamation Case Highlights Google's Double Standard." *Spiegel Online.* http://www.spiegel.de/international/germany/defamation-case-by-bettina-wulff-highlights-double-standard-at-google-a-854914.html. Retrieved November 9, 2015.

Pagallo, U. 2013. *The Laws of Robots. Crimes, Contracts, and Torts*, Dordrecht: Springer.

Palombella, G. 2007. "Ragioni di giustizia, diritti e generazioni future." *RIFD. Rivista Internazionale di Filosofia del Diritto*, 3: 399-436.

Pasquale, F. 2015. *The Black Box Society. The Secret Algorithms That Control Money and Information*. Cambridge (MA) and London: Harvard University Press.

Pasquale, F. 2011. "The troubling consequences of trade secret protection of search engine rankings." Pp. 381-405. in *The Law and Theory of Trade Secrecy: A Handbook of Contemporary Research*, edited by Dreifuss, R.C., Strandburg, K.J. Cheltenham: Edward Elgar.

Rodotà, S. 2014. *Il mondo nella rete*, Roma-Bari: Laterza.

Sartor, G. 2006. "Cognitive automata and the law." EUI Working Papers. European University Institute. San Domenico di Fiesole.

Sartor, G. 2009. "Cognitive automata and the law." *Artificial Intelligence and Law.* 17: 253-290.

Sartor, G. 1990. *Le applicazioni giuridiche dell'intelligenza artificiale. La rappresentazione della conoscenza*. Milano: Giuffrè.

Shapiro, D. 1999. *The Control Revolution. How the Internet is Putting Individuals in Charge and Changing the World We Know*, New York: Public Affairs.

Supreme Court of California. 2014. *S. Louis Martin* vs. *Google Inc*. (No. CGC-14-539972, Cal. Sup. Ct. Nov. 13, 2014).

Varian, H.L. 2013. "Beyond Big Data." Working Paper. NABE Anual Meeting. San Francisco, in http://people.ischool.berkeley.edu/~hal/Papers/2013/BeyondBigDataPaperFINAL.pdf. Retrieved November 9, 2015.

Wiener, N. 1964. *God and Golem Inc.* Cambridge (MA): The MIT Press.