



UvA-DARE (Digital Academic Repository)

Search engine freedom: on the implications of the right to freedom of expression for the legal governance of Web search engines

van Hoboken, J.V.J.

Publication date
2012

[Link to publication](#)

Citation for published version (APA):

van Hoboken, J. V. J. (2012). *Search engine freedom: on the implications of the right to freedom of expression for the legal governance of Web search engines*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter 8: Search Engine Freedom

8.1 Introduction

The aim of this chapter is to develop a general framework for the implications of the right to freedom of expression for the governance of search engines, a framework that is shortly denoted by 'search engine freedom'. Drawing on the conclusions in the previous chapter, the analysis will focus on the ways in which the right to freedom of expression may be invoked in the context of search engines by a search engine provider, on the one hand, and the information providers and end-users it mediates between on the other hand.

Search engine governance involves a great variety of public and private interests related to the effective and free dissemination of information and ideas, and the restrictions on such free dissemination. The three primary stakeholders, namely search engine providers, information providers, and end-users, have competing interests in the governance of information flows through the medium. And amongst information providers and end-users the types of interests vary as well. Moreover, third parties, such as private individuals, rights holders or public authorities have demonstrated their interests in the governance of search media through litigation and calls for regulation.

A complexity arises from the character of the decisions that are made by a search engine provider. A search engine combines a passive instrumental role with a role of independent communicator about information. It facilitates accessibility and mediates between information providers and end-users, but it does so by adding value through its selection and ranking decisions, which have considerable impact on what ends up being found by end-users. Most search engines, and selection intermediaries more generally, do not make individual determinations of the value of each source of online information in their index but predominantly use automated procedures, web data mining, and end-user data processing and modeling to decide which sources of information to prioritize over others.

This chapter will mostly focus on the question about the scope of the right to freedom of expression for the governance of search. This scope of the right to freedom of expression and information will be addressed in relation to the existing and possible restrictions that could stand in the way of the freedom of expression interests of the primary stakeholders to materialize. A range of existing, proposed or hypothetical legal restrictions on any of the communications between the primary stakeholders aimed at the governance of accessibility will be discussed to develop the argument in this chapter. Whenever feasible, reference will be made to specific proposals for the adoption of a specific type of restriction on search engine operations. Notably, some of the legal restrictions on the right to freedom of expression in the context of search will be addressed in much more detail in the third part of this dissertation, which focuses on specific regulatory issues relating to access and quality in search engines.

The focus will be placed on restrictions of a legal nature and the role of the state or of government is of central concern. Moreover, the analysis is, like in the rest of this study, restricted to issues that are relevant from the perspective of the free circulation of information and ideas, and in particular the governance of accessibility, which lies at the core of governance of information flows in the Web search context. The discussion of the scope of freedom of expression will be tied to the question about the implications of Article 10 ECHR. The ECtHR has not addressed search engines in any of its rulings yet, so

one have to draw analogies to other contexts to be able to address the implications of Article 10 ECHR for the possible legal governance of search engines. This has informed the analysis of freedom of expression in the context of the press, access providers and libraries in Chapters 5 to 7. These chapters have provided a broader picture of the way in which the implications of the right to freedom of expression are informed by the role of media and communication providers in the public information environment. The conclusions of these chapters about the relevant elements of freedom of expression doctrine will be used in this chapter and in the next part of this thesis.

Article 10 ECHR literally describes the freedom to receive and impart information and ideas as communicative actions which are protected. However, as was demonstrated in the previous chapters, it protects a broader range of communicative actions, including the right to transmit information freely, the right to use communicative means. The same is true for the First Amendment. Article 10 ECHR also implies a right to be able to effectively exercise one's freedom to impart or receive information and ideas, a right that can become relevant in horizontal relations. This focus on the freedom to communicate in different ways and capacities implies that to study the implications of the right to freedom of expression one could zoom in on the various communications that are taking place, identify the possible interferences with these communications and discuss their legal legitimacy under the European Convention. However, such an endeavor would be both too limited and too broad for a number of reasons.

First, not only communicative actions but also actions that are indirectly linked to communication can be protected under the right to freedom of expression. For instance, the right to freedom of expression also protects the freedom to decide how to use the means to communicate. For example, a restriction on the freedom of a theatre proprietor to decide to whom the theatre will be rented for public performances can be seen as an infringement. It can also protect media against special restrictive legal treatment, such as discriminatory taxation. In other words, the scope of freedom of expression also includes actions that are only indirectly related to communication but facilitate or are a necessary part for various actors to exercise it. Trying to discuss all these possibilities does not seem to be a valuable approach.

Second, not all kinds of communicative actions are meaningfully protected by the right to freedom of expression. The right to freedom of expression tends to protect the actions that have become to be seen as legally meaningful from the perspective of the ideals underlying the right to freedom of expression.⁵⁶⁵ Typical examples of such actions that are considered to be of particular importance are the right of citizens to speak publicly about matters of public concern, the editorial freedom of newspapers to decide which articles to print, the right to publish information without asking the authorities for permission, or the right of citizens to use electronic communications networks freely. Thus, conceptualizing the scope of freedom of expression in the context of search, it may be more productive to look for a characterization of search engine freedom that encapsulates the 'typical' scope in the context of search.

⁵⁶⁵ See e.g. Schauer 2004.

As mentioned, the typical scope of the right to freedom of expression is connected to the normative theories underlying the right to freedom of expression. The end-user's freedom to receive information and ideas freely and become an informed citizen is considered to be worth protecting because of the importance of informed citizens in a democratic society and their individual autonomy. Similarly, the analysis of freedom of expression theories for the press, ISPs and libraries in Chapters 5 to 7 demonstrate that the protection of these entities under Article 10 ECHR strongly takes into account the 'societal role' of these media. For instance, the function of the press as public watchdog and as platform for debate about matters of public concern informs strong protection of the press. This freedom related to the production and selection of information and ideas is meant to guarantee that public debate can be 'uninhibited, robust and wide-open', as formulated by the U.S. Supreme Court in *Sullivan*.⁵⁶⁶ In the case of Internet access providers, freedom of expression theory stresses the protection of the interests of end-users and information providers, precisely because of the relatively passive intermediary role of ISPs, to provide access. In short, the right to freedom of expression helps to protect the 'public freedom of expression interest' in the free dissemination of information and ideas. And similar to the case of access providers, the press and libraries, if search engines can be demonstrated to have a particular role related to the right to the public interests in freedom of expression, they could receive more protection.

When answering the question about such public interest, it has to be kept in mind that it would be wrong to make freedom of expression fully instrumental to a particular conception of the public interest in the free circulation of information, knowledge and ideas. There is no agreement about the definition of such a conception in freedom of expression theory. The discussion in Chapter 6 of the possible claim of access providers under the right to freedom of expression in the United States when asked to act as common carriers may serve as an example of this disagreement. The law and freedom of expression doctrine have referred to different public interests underlying freedom of expression, but they have also protected the right to freedom of expression independently. In the search engine context this protection independent of the public interest means that freedom of expression simply protects the communications between the information provider and the search engine and the communications between the search engine and the end-user against undue interference by public authorities or as a result of other legal restrictions.

The chapter will proceed in the following order. First, a number of general starting points for a discussion of the implications of the right to freedom of expression in the context of search engines will be discussed (Section 8.2). Drawing on a comparison between the entities discussed in Chapter 5, 6 and 7, Section 8.3 develops a conceptualization of the societal role of search engines, in which the implications of the fundamental right to freedom of expression can be incorporated. Section 8.4 discusses the implications of the right to freedom of expression for the most important communicative actions of search engine providers. After this, Sections 8.5 and 8.6 provide additional perspectives on search engine freedom, from the perspective of information providers and end-users respectively.

8.2 Search engine governance: starting points

8.2.1 Introduction

⁵⁶⁶ *New York Times Co. v. Sullivan*, 376 U.S. 254 (1964).

The Web search medium is a new and complex phenomenon in information law and policy. To be able to address the implications of the right to freedom of expression for the governance of search in this chapter, this section will first address some of the general characteristics of Web search governance. The following characteristics will shortly be addressed, as well as the way and the extent to which they will be addressed later in this and the next chapters: First, the structure of the search engine market and the prevailing business model for general purpose search engines. Second, the fact that online search services tend to be publicly accessible on the one hand and could be argued to be part of the public information environment accordingly, but they operate close to and often in the private realm of the end-user on the other hand. Third, the existing legal and regulatory environment search engines act in is in its infancy. And finally, search engines like many online services act in the global information environment made possible by the Internet, thereby presenting a number of issues relating to legal jurisdiction.

8.2.2 The search engine market and its business model

In the first part of the thesis the general structure of the search engine market was analyzed and the prevailing business model for general purpose search engines discussed. Monetization strategies can vary, but most services have an advertisement-based business model. This entails the auction of sponsored search results, which are placed next to organic search results. Publicly funded general purpose search engines are absent and over the last two decades the market has developed into an oligopoly of a few globally operating services, in which Google is by all measures the dominant player. Nonetheless, it was observed in the first part of this study that there remain many other entities, including specialized search portals, directories, Web publishers and new phenomenon like micro-blogging sites and social networks, that strongly contribute to the function of facilitating access to, and the findability of, information online. It has to be acknowledged that linking and referencing is a core function of the web environment more generally and which cannot be monopolized by a single search engine or single type of service.

Most search engines and other selection intermediaries are provided by private actors on a commercial basis. The dominant business model, discussed in more detail in Chapters 2 and 3, involves the sale of targeted reference space. The commercial nature of the services implies that they operate in a vertical relation with the state and a horizontal relation with end-users and information providers. As publicly funded search engines remain the exception, the analysis will focus primarily on the legal governance of commercial search services. A publicly funded search engine would have a comparable relation with the state as the public library. It would act in a vertical relationship with end-users and information providers, which would imply that the latter could have a more straightforward claim under freedom of expression with regard to the governance of information flows.

The advertisement based business model brings with it a number of regulatory issues relating to the law and regulation with regard to commercial advertising and the possible impact of advertising on the governance of the medium. The separation of advertising and editorial content in print media has served as a model for the legal governance of search media in the context of the demarcation of organic and sponsored results. The resulting regulatory framework will be more thoroughly discussed in Chapter

10. Further below in this Chapter, the connection to the development of a body of professional ethics of search engine governance and its relation to the right to freedom of expression will be explored.

Search media provide findability in the public networked information environment which not only provides for access to sources of knowledge, information and ideas, but also to a variety of other service providers and e-commerce sites. As a result, search engine do not only play an important role in the market place for information and ideas but also, and not in the least, in actual markets for actual goods and services. The commercial nature of many of the communications in the search engine context can become relevant for the protection under the right to freedom of expression. The right to freedom of expression does not protect commercial and non-commercial communications to the same degree. Generally, governments have more leeway to regulate commercial communications. It could be argued that the persuasive presence of commercial communications in general purpose search results, both in sponsored and natural results, gives more room for an active regulatory role and specific restrictions on the freedom of search engine providers. For example, in a recent judgment the U.S. Ninth Circuit characterized the selection decisions of specialized service for finding housing opportunities as commercial speech. This meant it could apply the less restrictive test for restrictions of commercial speech to a legal restriction on the operation of search engines pursuant to the Fair Housing Act.⁵⁶⁷

8.2.3 Public and private nature of Web search communications

This study has focused on Web search services which are offered to the public at large through a publicly accessible location on the Internet. Moreover, they index and refer their users to publicly accessible online destinations. Thus, search engines can be said to operate in the public networked information environment, alongside others such as online publishers. However, it is clear that the use of search engines often takes place in the privacy of the home, or in the relative privacy of the workplace. In addition, the interactive nature of the service implies that the actual communications that are taking place between end-users and search engine providers are primarily of a private nature.

To summarize, search media are arguably public, in the sense that the service is offered to the public and its search results refer to publicly accessible destinations. But their actual offering are quite private, since the actual output of search engine consists of a specific individualized set of search results offered to a particular end-user, in response to a particular query at a certain place and time. The increased personalization of search engine services shifts the nature of these services and their offerings more into the private, individualized sphere.⁵⁶⁸ There is no longer such a thing as ‘the’ references offered by a search engine. The offering of ranked collections of search results have started to depend so much on private interaction with specific end-users that the output of search engines becomes ill-defined without reference to their specific queries under specific circumstances.⁵⁶⁹

From a constitutional perspective and in particular from the perspective of the important constitutional dichotomy between public and private communications, the private nature of the actual output to end-

⁵⁶⁷ For further discussion, see Section 8.4.3.

⁵⁶⁸ On the questions, from a German Constitutional perspective not addressed in this thesis, see Schulz et al 2005a, p. 26.

⁵⁶⁹ See Zuckerman 2011, in reaction to Feuz et al 2011.

users, and the dependency of the output on particular end-users, could have implications for the role of government with regard to the communications that are taking place in the context of search. It could be argued that the reading of a newspaper or a book is also of a private nature. The difference, however, is that the actual output of search engines has become hard to define, without taking into consideration the actual user behavior. A newspaper or a book does not change its content depending on the choices of its specific reader.⁵⁷⁰ The relative private nature of the communications between end-users and search engines constitute an additional barrier for involvement by public authorities, as the communications are taking place in the private realm, where it would be less appropriate to regulate. As in the case of Internet access providers, not only Article 10 ECHR but also Article 8 ECHR is relevant.

8.2.4 A legal and regulatory framework in its infancy

Search engines operate in a regulatory framework that is still in its infancy.⁵⁷¹ Search engine law and regulation consist of a mix of the maturing application of general applicable laws, a growing body of self-regulation, and in some jurisdictions a few search engine specific legal provisions such as the DMCA safe harbor for search engines mentioned in Chapter 6. This mix of different elements of search engine makes it much harder to tackle search engine law and regulation as a whole.

As mentioned, for the most part search engine law consists of the application of generally applicable laws. Hence, when restricted to issues relating to the free dissemination of information and ideas, the study of search engine law consists of the study of tort law (including defamation law, privacy law and unfair competition law), property law, copyright and database law, trademark law, advertising law, consumer protection law, and data privacy laws as they apply to the search engine context.⁵⁷² This does not mean that the application of these general laws does not involve issues that are specifically related to search engines, but it simply means that the specific rules that have been applied in this process are judge made.⁵⁷³ As will become clear in greater detail throughout the remainder of this study, the right to freedom of expression – through the doctrines of direct or indirect horizontal effect of fundamental rights – has sometimes played a significant role in the debates about the proper application of these general laws and there are several judgments that specifically address this question.⁵⁷⁴

The statutory safe harbors for third party liability are currently the only example of search engine specific regulation. And these search engine specific safe harbors only exist in the United States and not at the EU level.⁵⁷⁵ As was shown in Chapter 6, these safe harbor provisions internalize the right to freedom of expression into the standard for intermediary liability for third party activity to some extent.

⁵⁷⁰ Due to personalization, a somewhat similar development can be observed in the context of in electronic media more generally. See also Section 10.2.1.

⁵⁷¹ For an overview of the absence of European search engine law, see Van Eijk 2006. See also Jakubowicz 2009; Van Eijk 2009, Van Hoboken 2009; Valcke 2008. For an overview of U.S. search engine law, see Grimmelmann 2007.

⁵⁷² For an overview, see Grimmelmann 2007. See also Gasser 2006.

⁵⁷³ For an (unconvincing) argument that search engine law should be further developed through a specialized judicial forum, building and extending the judge made character of search engine law, see Moffat 2009.

⁵⁷⁴ Such as the BGH Paperboy judgment and related rulings in Germany, BGH 17 July 2003, I ZR 259/00 (*Paperboy*).

⁵⁷⁵ Some EU Member States, for instance in Austria and Spain, have also adopted search engine specific safe harbours. For a discussion, see Section 9.3.4.

Except for these safe harbors, specific government involvement with search engines has mostly followed the self-regulatory paradigm for Internet content regulation, which was discussed in Chapter 6.⁵⁷⁶ Many of these proposals will be discussed in detail below, in particular how they relate to the right to freedom of expression.

It is important to note here that the call for search engine specific self-regulation often overlaps with the debate about the proper application of generally applicable laws. Generally applicable tort law, for instance, is indeterminate enough to argue both ways. It may be argued on the one hand that search engines should remove illegal information from their index upon request to escape general tort liability. On the other hand one may argue that search engines have no such obligation and that any removal of sources of information is thereby voluntary.⁵⁷⁷ Moreover, what could be considered self-regulation at one point in time could become an expected duty of care under general tort law principles if it is widely adopted. As was discussed in more depth in Chapter 6, Internet self-regulation often involves a high level of involvement by public authorities, which warrants the question of to what extent self-regulation merely informalizes the regulatory role of the state instead of minimizing it.⁵⁷⁸ This question is particularly relevant for the legitimacy of interferences with the right to freedom of expression under Article 10 ECHR. Any interference by public authorities with the right to freedom of expression, including informal ones, has to be ‘prescribed by law’.

8.2.5 Search engine governance, a global Internet and jurisdiction

The establishment of jurisdiction over online services and information providers is an important and complex issue. Compared to Internet access providers, most Web search engines have relatively weak jurisdictional ties. General purpose Web search engines like the one provided by Google operate on a global scale and, unlike access providers, they can easily be provided from abroad like any other Web site. Increasingly, Web search providers tend to diversify their services into country – or language – based versions. In addition, geo-location software is used to target search results based on location, implying that search engines tend to know the location of specific users. Still, end-users typically remain free to choose which of the language or country specific services to use. This presents a number of problems for any type of country-specific search engine law and policy.

To be able to regulate or apply existing general laws to search engines in specific circumstances, the A State should be able to claim jurisdiction in a way that is in line with fundamental principles of international law. These principles include the right to freedom of expression and information which specifically protects the free flow of information between jurisdictions.⁵⁷⁹ Considering the international nature of the offering of Web search providers and the design and the general governance of the Internet as a universal network of interconnected networks more generally, it is likely that different countries only have a limited claim of jurisdiction over the infrastructure and communications of a

⁵⁷⁶ See Section 6.4.1.

⁵⁷⁷ See Schulz & Held 2007.

⁵⁷⁸ See Tambini at al 2008, discussed in more detail in Section 6.4.6. See also Bambauer forthcoming 2011.

⁵⁷⁹ See Section 4.3.1.

specific search engine provider. This means that the scope of national law is only applicable or regulation could only be adopted with regard to that 'part' of the service.

The limitations this implies with regard to the effectiveness of any type of country-specific regulation that tries to restrict accessibility of content through search engines are obvious. Laws and policies which would aim to facilitate accessibility may end up being more effective. The search engine laws and policies that end up being adopted should not only respect the accepted principles of jurisdiction, but should also be enforceable. The enforceability of laws on the Internet used to be seen as *the* major obstacle for effective regulation of content flows by national law. It is now generally accepted that well-designed rules and policies can be enforced, even with regard to services and information providers on the Internet.⁵⁸⁰

8.3 Between access and quality: the societal role of search media

8.3.1 The role of search engines in the networked information environment: a comparison

As noted above, search engines combine an intermediary function, between information providers and end-users, with an independent function of communicator of information about information. In their capacity as intermediary, search engines are maybe better qualified as a *meta-medium*. In their capacity of independent communicator, search engines act as content (references) providers themselves, making choices about their communicative actions independent of information providers, end-users and others. As will become apparent in the following sections, both perspectives have their own merits and can help to properly answer the questions about the implications of the right to freedom of expression for the governance of search.

Search engines do more than just providing references to the information and ideas which are available online in return to user queries. This can be clarified by comparing the function of Internet access providers in the networked information environment, as discussed in Chapter 6, with the function of search engines. An Internet access provider is an essential but invisible part in any attempt of an end-user to access information online. In principle, an access provider - that does not block or discriminate⁵⁸¹ - makes every online destination similarly accessible to its subscribers.

The function of search engines is different from Internet access providers in two fundamental ways. First, they are not a necessary part in the communicative process between information online and end-users. An online destination can be reached without the use of a particular search engine, either directly by using its address on the Web or indirectly by using another search or recommendation service or following a hyperlink elsewhere.

Second, search engines do not facilitate equal access to everything on the Web. They do not connect all end-users to all online destinations but mediate between the two sides on the basis of range of content,

⁵⁸⁰ See generally, Goldsmith & Wu 2006.

⁵⁸¹ But as was discussed in Chapter 6, even ISPs sometimes make decisions that impact on the (relative) accessibility of information. For instance, ISPs could decide to block access to material. ISPs can also have a less direct impact through the caching of material and the reservation of special capacity for certain content providers. See Section 6.4.2 and Section 6.5.1.

destination, and end-user and query-specific characteristics. This mediation is carried out by its decisions and algorithms about the crawling of online material and their selection, ranking and presentation of references. In other words, search engines are not neutral by nature.⁵⁸² They actively match information providers (including advertisers) and end-users and this active role with regard to the valuation of different online destinations for its users is what helps to make one search engine more valuable than another.

While search engines do more than providing access, they also do less – as regards the valuation and selection of content – than an actual content provider. This can be clarified with a comparison of the function and the strength of editorial control that is asserted over information flows in the search engine context with the function of the press or the mass media more generally.

The function of search engines is different in two important ways. First, the press does not only select which information and ideas to present to its users, it actually publishes the information itself, thereby making itself into the source of publicity. While the press takes the responsibility for the publicity of information and ideas (as well as the non-publicity of what it did not decide to print), search engines merely refer to what is already published. In Chapter 5, it was shown that while the press has also been recognized as a platform for information and ideas to find their way to an audience, its primary role is considered to be to actively select what information to publish, using and relying on some and discarding others. By deciding about the trustworthiness of sources and the newsworthiness of the events they report about, they are actively helping to construct the boundaries of public debate a role that has given them the name gatekeepers.⁵⁸³

Second and related, because of its role to facilitate the navigation of the online information environment, a search engine's value greatly depends on the comprehensiveness of its index. Consequently, search engines tend to be far less restrictive in their selection decisions than content providers like the media. Search engines simply cannot be very restrictive in terms of crawling and indexing, because they would simply lose their navigational function for the Web if they would discard sources of online information that end-users are or might be looking for. 'Recall', one of the two traditional criteria for the quality of information retrieval systems, demands that the index of information retrieval systems should be as comprehensive as possible. On top of that, the sheer amount of available resources online makes it attractive and possibly commercially imperative to rely on others, including information providers, to decide what to index. In practice, the control over indexing and selection is therefore *de facto* shared and distributed. The power of search engines lies not in the exclusion of material but in making material more readily accessible, i.e. matching specific end-users with specific sources of information based upon decisions made by search engine providers and to some extent programmed into the search engines' architectures. To conclude, whereas a search engine does try to help end-users find the most valuable or relevant online destinations, it does not, except in certain circumstances, try to prevent its users from accessing online information and ideas of low quality or of other undesirable characteristics.

⁵⁸² See Grimmelmann 2010b.

⁵⁸³ See White 1950; Bass 1969; Shoemaker 1991; Gladney 1996; Sturges 2001.

To better understand the role of search engines in the public networked information environment better it is useful to compare the accessibility governance of search engines, with the collection and decisions about availability and accessibility that are made in the context of libraries and which were discussed in Chapter 7. Both the search engine and the library combine an intermediary function and an active function with regard to the selection of sources of information and ideas, but they do so in different ways. Simply put, libraries first select information and ideas which are allowed to enter their collections. These collection decisions reflect criteria relating to quality. Hence, decisions about quality are part of the governance of availability. This selective role of libraries is informed by their mission: public libraries, amongst other things, provide basic access to knowledge and our cultural heritage. A lot of media and information in the public information environment, such as commercial communications, illegal material or self-published content are typically not included in library collections in the first place. After the collections have been established, libraries provide an accessibility infrastructure for their patrons that makes everything *in the collection* findable. The library collection is made transparent for end-users through an accessibility architecture that is comprehensive and aims to be follow scientific criteria for the organization of material.

Like newspapers, or libraries, Web search engines could pre-select and exclude information and ideas from their indexes. In particular they could do so *on the basis of quality*. Generally speaking, search engines do exclude large amounts of material from their index, but they do so for different reasons than newspapers or libraries. For instance, all general purpose search engines exclude material because of legal restrictions, because of search engine manipulation and because of crawling instructions by information providers. All general purpose search engines have serious and ongoing problems with illegal or harmful publications by third parties in their indexes. The same search engine that can be used for looking up information about 17th-Century pre-industrial production processes or finding tickets for the 2010 World Cup in South Africa provides access to hardcore pornography, racial hatred, and illegal advertising. Most search engines have standard optional filters to prevent indecency from entering their search results. There are several possible legal restrictions on the lawfulness of including material in the index, for instance because of the existence of exclusive rights with regard to the underlying information, such as copyright law, database law, or because of content-specific restrictions such as defamation law, privacy law, data protection law, and child pornography law. The legal regime for search engine intermediary liability as well as their involvement in content regulation will be discussed in Chapter 9.

8.3.2 Search engine governance: between access and quality

It must be concluded from the above that search engines are more actively involved with content than access providers because of their role in selecting and valuating sources of information and ideas online. At the same time they are less actively involved as content providers like the press, because their selection decisions are relatively passive and merely refer to online destinations without taking responsibility for their existence. Moreover, search engines facilitate and exert power over the relative accessibility of information and ideas, not over its availability.

In one of the most thoughtful attempts to conceptualize the societal role of search media in the networked information environment, Dutch philosophers of science Marres and De Vries pointed out that search engines have to reconcile two conflicting ideals: the ideal of *universal access* on the one hand and the ideal of *information quality* on the other hand.⁵⁸⁴ The first ideal is to help end-users to navigate the *entire Web*, by ordering it making the material that is available, transparent and universally accessible. The second ideal is to prioritize valuable and accurate information and ideas over lesser ones and to exclude illegal or harmful material. Marres and De Vries point out that the resolution of these conflicting ideals in the 'old' public information environment used to be based on the conceptual separation between the production and legitimization of information and ideas on the one hand and the subsequent use of it on the other hand.

For instance, the New York Times traditionally claims to include "*all that is fit to print*."⁵⁸⁵ This implies the exclusion of stories that could not be confirmed by reliable sources, irrelevant information, and opinion of low quality. In other words, the legitimacy of the material ending up in a newspaper was, at least in theory, carefully guarded. The ideal of (universal) access was restricted to everything fit to print and the freedom and responsibility of newspapers to decide what was fit to print, i.e. their editorial freedom, was seen as an essential element of a free press.

Likewise, a library first selects the sources it subsequently makes accessible.⁵⁸⁶ Libraries apply quality criteria in the context of these selection decisions. They used to rely on established sources of information and ideas, and everything below a certain quality and relevancy threshold would and could be discarded. After having established a collection, a transparent accessibility infrastructure, informed by professional principles for the organization of knowledge, would ensure universal access to the materials that had been selected.⁵⁸⁷ Of course, library accessibility infrastructures such as their classification systems also have their biases, and accessibility is not always guaranteed in practice.⁵⁸⁸ The important point, however, is that in the print-based public information environment, access and quality were separated: libraries did not have to worry about quality standards in the context of accessibility, because of the simple fact that the book is available in the library means that it is a legitimate source of information for library patrons. Moreover, in their collection selection decisions libraries could rely on publishers which would only publish what they considered to be of high enough quality. Notably, the provision of Internet access in libraries, addressed in Chapter 7, has broken down the traditional separation of the governance of quality and access in libraries. Internet access in libraries extends the library collection beyond the sphere of materials with institutional guarantees relating to information quality.

⁵⁸⁴ Marres & De Vries 2002.

⁵⁸⁵ See e.g. The New York Times 1901 ("Only the news that is 'fit' to print finds place in its columns.").

⁵⁸⁶ See Section 7.4.

⁵⁸⁷ Although libraries may sometimes place restrictions on access to certain books, such as rare, valuable or pornographic books. See Bowker & Leigh Star 2000, p. 11.

⁵⁸⁸ On the social and political consequences of classification as a phenomenon, see Bowker & Leigh Star 2000.

Marres and De Vries point out that the traditional separation between the production and legitimization of knowledge and its use no longer exists in the context of the Web.⁵⁸⁹ Furthermore, and more fundamentally, they assert that *the legitimization of sources of information and ideas no longer takes place through the governance of availability but through the governance of accessibility*. In other words, competition in the marketplace of information and ideas has partly shifted from competition in the context of the platforms of publishing media to competition in the context of search media and selection intermediaries.

In general, the ideals of universal access and information quality lie at the core of many of the debates about the governance of information flows on the Internet.⁵⁹⁰ The networked information environment and the related possibility of self-publication and further dissemination to the public unrestricted by traditional knowledge institutions and the mass media has broken down the institutional encirclement of certain sources of information for societal consumption.⁵⁹¹ This 'disintermediation' is seen as one of the central promises of the Internet and the Web by those that lamented the gatekeeping power of mass media and traditional knowledge institutions.⁵⁹² At the same time, this may be seen as the central flaw of the Internet by those who place more emphasis on the public interest in restricting access to uncontrolled flows of information and ideas and on the value of shared platforms for debate and circulation of information and ideas which guarantee minimal levels of quality.

Freedom of expression theory also reflects these conflicting ideals of quality and universal access. Meiklejohn for instance has argued that the value of freedom of speech does not lie in everyone's freedom to speak but that *"everything worth saying shall be said."*⁵⁹³ In other words, freedom of expression law and theory should place emphasis on the creation of a framework that helps to foster a public debate of high accessibility *and* quality. Selection decisions with regard to information flows would be protected because they contribute to this ideal. Early in the Twentieth Century Walter Lippmann made a particularly strong argument in favor of information quality on the basis of the freedom of expression theory related to democratic self-governance.⁵⁹⁴ In *Liberty and the News*, Lippmann dismissed some of the most famous arguments for the freedom of the press as the expression of indifference over information quality and the formation of public opinion. He considered this indifference with regard to the functioning of the press wholly irrational:

⁵⁸⁹ Marres & De Vries 2002.

⁵⁹⁰ See Marres & De Vries, p. 188. on the so-called 'garage box' paradox. They compare this paradox to the classic paradox, debated by Lippmann and others, between the ideal of accurate representation of science and facts on the one hand and political democratic representation and public opinion on the other hand.

⁵⁹¹ On the dynamics of encircling and uncircling, from a perspective of search engines, see Duguid 2009.

⁵⁹² See e.g. Shirky 2008; Gillmore 2004.

⁵⁹³ Meiklejohn 1948, p. 25.

⁵⁹⁴ Lippmann's work also helped to shape the 20th Century ideal of separating the governance of the selection and legitimization of information in the public information environment on the one hand (quality) and its societal use on the one hand (access).

*“In a few generations it will seem ludicrous to historians that a people professing government by the will of the people should have made no serious effort to guarantee the news without which a governing opinion cannot exist.”*⁵⁹⁵

At least as prominent in freedom of expression law and theory, however, is the principle that the government through its laws and policies should not be allowed to restrict access to information or access to audiences but leave this to the ‘marketplace of information and ideas’ and the free choices of individuals instead. More speech is better, because it increases competition between speakers and exposes end-users to a wider range of information and ideas. In addition, choices about the relative value of information and ideas should be left to the free market. In the context of the press, controlling and restricting the media in view of information quality is considered an infringement with the press’ editorial freedom, which strongly protects the press against access regulation. As was discussed in Chapter 5 with regard to the implications of the right to freedom of expression for commercial communications, First Amendment doctrine contains a particularly strong assumption in favor of the public’s rationality.

The breaking down of the separation between quality and access governance in the networked information environment leads Marres and De Vries conclude that the societal legitimization of knowledge now *de facto* takes place through processes of opening up. This opening up of the networked information environment, understood as the process of connecting information and ideas online to their societal use, is not a technical undertaking but must be understood as a *societal and political process*.⁵⁹⁶ Web search engines in particular help to construct the boundaries of accessibility in the new networked public information environment. Interestingly, Marres and De Vries express particular interest in the possible role of government to resolve these competing ideals, and they point to the importance of considering a facilitative role of government. In their view, the opening up of the Web is process to which the government and public institutions could contribute.

If one follows this logic, the overarching public interest lies in the establishment of a rich and robust societal infrastructure for the opening up of the Web, one which fosters the societal project of the legitimization and contestation of information and ideas in the public networked information environment. This characterization of the public interest in the governance of accessibility of the Web is attractive precisely because it captures both perspectives, i.e. access *and* quality, which lie at the core of search engine governance. This conception clarifies that search engine providers have to make non-trivial choices with regard to the balance between access and quality. These non-trivial choices are a direct result of the dynamics in the public networked information environment they are acting in. Publicity is no longer restricted to entities that offer *a priori* institutional legitimacy to the information and ideas they make public. More fundamentally, the choices are not only non-trivial, they are of a political nature and involve the complex balancing of different public and private interests, including the interests of end-users and information providers.

⁵⁹⁵ See Lippmann 1920.

⁵⁹⁶ Marres & De Vries 2002.

Of course, an important underlying question is how different Web search engines have ended up dealing with the tension between the ideals of access and quality which is inherent in the governance of their service. From a legal and normative perspective, the question is in what way search engines can and should be able to assert the right to freedom of expression in view of possible restrictions on the freedom to decide about information flows through their services. In the next section, these questions will be addressed in more depth. A general distinction will be made between two types of decisions a search engine has to make about its governance of accessibility. The first type is about the inclusion and exclusion of information providers in its index, while the second type is about the selection, ranking and presentation of references from its index to an end-user in response to a query. Further in this chapter the implications of the right to freedom of expression for the communicative liberties of end-users and information providers will be discussed in relation to the freedom of search engine providers.

8.4 Search engine providers and the right to freedom of expression

8.4.1 Introduction

Van Eijk has already noted the difficulties of qualifying search engines in light of the right to freedom of expression because of the complex role discussed above.⁵⁹⁷ For reasons related to this double role of search engines with regard to online communications, Van Eijk asserts that the freedoms to impart or receive information, as explicitly mentioned in Article 10 ECHR, are not the main aim of search engines since the information is already present and accessible directly.⁵⁹⁸ A search engine does not and cannot change that. Van Eijk concludes that a search engines *facilitates access* to information and does not offer access to information itself. He asserts that this activity of *making accessible* should have a similar status under Article 10 ECHR as the activity of disclosing or disseminating information and ideas. In other words, it is the function of search engines to make specific information, out of the abundance of information online, more readily accessible.⁵⁹⁹ Web search engines impact the relative *accessibility* of information which is already publicly available for Internet end-users. Therefore, an analysis of the implications of freedom of expression in the context of search engines should focus on the governance of such accessibility.

8.4.2 The freedom to publish referencing information and the freedom to crawl

If we turn to the question about the implications of the right to freedom of expression as protected by Article 10 ECHR for search engine providers, an important starting point is that the publication of search results by a publicly accessible search engine is clearly protected under Article 10 ECHR. The publication of search results itself constitutes a publication of the search engine provider for its end-users. In other words, a search engine does not only facilitate access to information of third parties, but it produces information itself, namely information about information. The most important question in this context is not whether the function of referencing by search engines is protected under the right to freedom of

⁵⁹⁷ Van Eijk 2006, p. 5.

⁵⁹⁸ Id.

⁵⁹⁹ The shift of accessibility governance from availability governance runs parallel with a shift from a publisher economy to an attention economy. See Van Hoboken 2009c; Goldhaber 1997; Bermejo 2007. See also Section 10.2.1.

expression, but what level of protection should be attributed to search engines' practices of referring to information elsewhere. It is possible that the weight of protection differs because of the type of communication. Referencing information might receive its own level of protection. This status could depend on the public interest in the publication of these references, and the status of the Internet more generally. Linking lies at the heart of the networked information environment. The fact that search engines in particular rely on hyperlinking to others can play a role in the assessment of the weight that should be attributed to their freedom to do so. On the other hand, the types of content the search engine refers to may play a role in the assessment as well.

In *Times v. United Kingdom*, a case about the protection of publications against defamation lawsuits after a considerable lapse of time, the ECtHR for the first time qualified the importance of the Internet for the promotion of the values protected by Article 10 ECHR.

*"In light of its accessibility and its capacity to store and communicate vast amounts of information, the Internet plays an important role in enhancing the public's access to news and facilitating the dissemination of information generally. The maintenance of Internet archives is a critical aspect of this role and the Court therefore considers that such archives fall within the ambit of the protection afforded by Article 10."*⁶⁰⁰

But in the same ruling the Court also lowered the protection for the publication of old newspaper articles on the World Wide Web. It did so by on the one hand stating an undefined but seemingly large margin of appreciation for the Member States to impose restrictions on the publication of news archives online, and on the other hand heightening the duties and responsibilities tied to such 'ongoing' publications.

*"[T]he margin of appreciation [...] is likely to be greater where news archives of past events, rather than news reporting of current affairs, are concerned. In particular, the duty of the press to act in accordance with the principles of responsible journalism by ensuring the accuracy of historical, rather than perishable, information published is likely to be more stringent in the absence of any urgency in publishing the material."*⁶⁰¹

Referencing information does not in general qualify as a publication of public concern or as current affairs. At the same time, it could be argued that referencing information in search engines has become as necessary for end-users to inform themselves about matters of public concern as the underlying stories themselves. Without the appropriate freedoms to reference to stories about matters of public concern, many of those stories would not be effectively accessible.

In *Open Door* the Court touched more closely on the legality of restrictions on referencing information under Article 10 ECHR.⁶⁰² The Court's considerations contain a number of elements that are useful for discussing the implications of Article 10 ECHR in the context of search engines. The case involved an

⁶⁰⁰ ECtHR 10 March 2009, *Times v. United Kingdom*, § 27 (adding that "[t]he maintenance of Internet archives is a critical aspect of this role and the Court therefore considers that such archives fall within the ambit of the protection afforded by Article 10.>").

⁶⁰¹ *Id.*, § 45.

⁶⁰² ECtHR 29 October 1992, *Open Door v. Ireland*.

absolute prohibition under Irish Law against providing information about abortion facilities abroad. The Court concluded that the measures in question were disproportionate and were an infringement of Article 10 ECHR. The Court based its judgment on a variety of circumstances. First, it asserted that it was not illegal under Irish law to act on the information in question: it was legal under Irish law to travel abroad to receive an abortion.⁶⁰³ The Court further noted that the restrictions could have a negative impact on women's health and called for careful scrutiny because other Convention countries tolerated these activities.⁶⁰⁴ Above all, the Court pointed to the absolute nature of the ban on referencing information about abortion facilities, which was therefore overbroad and disproportionate.⁶⁰⁵

The Court's conclusion in *Open Door* was strengthened further by a number of other factors. First, the counselors that were informing pregnant women about abortion facilities were not advocating or encouraging abortion, but merely informing women about the available options:

*"[t]he decision as to whether or not to act on the information so provided was that of the woman concerned."*⁶⁰⁶

This circumstance, and the fact that some women were likely to decide not to have an abortion on the basis of the information provided meant, according to the Court, that the link between the referencing information and the decision of certain women to act on this information was not as definite as presented by the Irish authorities. The Court also took into account that the information was not provided to the public at large. Moreover, the Court considered that the referencing information was also available through other means, such as in magazines and telephone directories, and that these other means contained lower professional guarantees with regard to the health interests of the women involved.⁶⁰⁷ Perhaps, nowadays the Court would have added a reference to Web search engines or the Internet to this list. Finally, the Court deemed the ban ineffective in preventing abortions abroad from taking place and in fact suggested that this particular restriction on *reliable* information about abortion facilities abroad was likely to have an adverse effect on the health of woman.⁶⁰⁸

Open Door shows that absolute restrictions on the publication of certain information should at least be consistent to be legitimate. The fact that banned information can still be accessed through other legal means can render a ban disproportionate. This conclusion of the Court is consistent with its considerations with regard to the illegal nature of the prior restraint in *Spycatcher*.⁶⁰⁹ These conclusions also resonate in the search engine context, for the simple reason that a search engine cannot prevent Internet users from accessing the information directly. This leaves open the possibility that it would be illegal to access the information directly. But at the least, it raises questions about the proportionality of legal restrictions on referencing information without action being taken against the actual source of the

⁶⁰³ Id, § 72.

⁶⁰⁴ Id, §72. Note that the margin of appreciation in this case would normally have been wide because of the moral nature of the restrictions, in which Member States have more leeway than in other domains.

⁶⁰⁵ Id, § 73.

⁶⁰⁶ Id, § 75.

⁶⁰⁷ Id, § 76.

⁶⁰⁸ Id, § 76-77.

⁶⁰⁹ See Section 5.3.4.

material. More generally, legal restrictions on providing references to certain types or sources of information could incentivize search engines to exclude material the indexing of which would be lawful. Through the doctrine of chilling effects, such restrictions are also relevant under the right to freedom of expression.

The prohibition on search engine providers of publishing references to certain specific types of content or specific online destinations comes closest to what we would call a prior restraint. This is a common phenomenon in countries like Saudi Arabia and China.⁶¹⁰ In China, for instance, it is apparently prohibited to communicate references to publications relating to the Tiananmen Square protests in Beijing in 1989. In more democratic countries than China, including Germany and Argentina, there are court rulings in which search engines were ordered not to communicate any references to information about specific individuals.⁶¹¹ It goes beyond the scope of this chapter to discuss the legitimacy of legal interferences on the search engine's freedom to publish references in depth here or as the possible liability of search engines for referring to illegal or unlawful content. This will be addressed in Chapter 9.

Before discussing the protection of the communication of referencing information by search engines in more depth, it is important to address the communications that are a necessary to publish references on the basis of an index of online content in the first place, namely the crawling of the Web. Without crawling, the type of full-text search engines that have become the standard navigational tool in the networked information environment could not be offered.

Crawling consists of the use of specialized software that harvests the information which is available on the Internet through automated requests. In view of the ECtHR's decision in *Autronic*,⁶¹² this kind of automated requests can be considered to be an important and specialized means of reception in the context of the Internet, protected by Article 10 ECHR. The search engine's crawlers are a means of receiving information, in some ways a modern automated version of the researcher or journalist delving for sources of information by making phone calls, talking to people directly or looking through printed material. Article 10 ECHR does not explicitly refer to a right to gather information, but these communicative processes are generally understood to be included in its scope.⁶¹³

The automated and complex nature of crawling technology is directly related to the abundance of publicly accessible sources of online information. Crawling consists of the automated and repeated requests for content available online. The importance of search engines for the dissemination of information and ideas and the possibility of the public to inform itself could be one of the determining elements for the extent to which crawling and the subsequent indexing of references is protected and the question whether possible interferences are in accordance with Article 10 ECHR, second paragraph. Thus the protection against interferences with crawling by search engines could possibly be enhanced because of the importance of search engines in our information society.

⁶¹⁰ See Deibert et al 2007. See also Villeneuve 2008.

⁶¹¹ See Soghoian & Valle 2008; Bonim 2008. See also Van Hoboken 2009c.

⁶¹² See Section 6.3.1.

⁶¹³ See Section 4.4.3.

8.4.3 Search engine freedom: selection and ranking

Up to this point the analysis has focused on the protection of the search engine's freedom to control the composition of its index and publish references. But the governance of the composition of the index and of the freedom to crawl and publish references is only half of the story when it comes to the governance of accessibility of information and ideas by search media. The real impact of these media stems from their selection and ranking technology and their interaction with information providers and end-users: the selection, ranking and presentation of references in response to particular queries. To what extent can search engine claim protection under the right to freedom of expression for their decisions to select a particular set of references as relevant in relation to a query? And to what extent can search engines claim protection for the way in which they rank these references once they present them to their users? Before discussing the answers to these questions, it is important to discuss the character of decisions about selection and ranking more detail.

The very fact that search engines select and rank the references for end-users implies that they do make decisions that entail a valuation of information and ideas.⁶¹⁴ As a result, search media impose specific hierarchies on the relative accessibility of information and ideas through their services. It may be important to realize that a full understanding of the way in which search engines impose such hierarchies on accessibility and the impact this has on our public networked information environment is currently beyond our grasp. Seemingly, these hierarchies tend to be informed by a different mindset than traditional ordering mechanisms for knowledge, information and ideas, such as the ones would find in a library. Metrics of quality assurance by human editors have been automated and supplanted by relevancy and popularity algorithms. Moreover, commercial motivations have entered the domain of the organization of information and ideas on the Web due to the function of the Web in facilitating e-commerce and advertising. The incentives and subsequent strategies related to governance of the search platform are partly a result of its business model, i.e. a targeted marketing platform, where organic results are offered alongside sponsored results that provide for income.⁶¹⁵

Google's search engine algorithms, and its PageRank algorithm in particular which was the core of Google's search engine when it launched, are the best known example of the imposition of a particular hierarchy on the ranking of search results. Before Google, and in traditional information retrieval systems, the selection of a relevant set of search results from the index, before ranking, used to take place by imposing a binary measure of relevance between the query and the websites stored in the index. The subsequent ranking of these websites would take place by looking at the similarity between the query and the relevant websites. A website mentioning the University of Amsterdam, for instance, would be considered relevant for a user entering the queries [University of Amsterdam], [university] or [Amsterdam]. A website mentioning the University of Amsterdam 100 times would be considered, by the typical early Web search engine, to be more relevant than a website only mentioning it once. Not surprisingly, this did not always produce desirable results.

⁶¹⁴ See Finkelstein 2008 (pointing out that Google's ranking decisions are driven by popularity metrics). See also Rogers 2009a. Google tends to compliment itself for the disruptive but 'democratizing' impact of its services. See also Caldas et al 2008.

⁶¹⁵ See Marres & De Vries 2002. See also Finkelstein 2008.

Google's PageRank algorithm was one of the early inventions about the way in which search engines could use the Web environment as a whole to be statistically predict the relative value of different websites in the index. The reputational and commercial economies of hyperlinks were already present before the development of PageRank by Google's founders, as linking between websites were establishing the navigational paths of the visitors of the Web. Google cleverly harvested the value of information that was present in the linking structure of the Web by turning these links into the building blocks of a global ranking metric, in which online popularity and reputation is taken as a relevancy measure, and in which those websites that are more relevant are considered more likely to point to other relevant websites as well.

A decade later, Google's selection and ranking algorithms have undergone major changes. While the linking structure of the Web may still play an important role in its selection and ranking, new ideas of how to improve ranking or establish baseline relevance have been developed. For instance, the choice of users for particular websites, expressed by their clicking behavior, now plays an important role in the ranking of search results. Most importantly, the actual selection and ranking does not follow from the use of a single smart algorithm but the combination of a large variety of algorithms and algorithmic corrections, which together produce a search engine's output.

The seemingly dominant criteria for the valuation of information and ideas in commercial Web search engines, namely relevancy, popularity and marketing, can be contrasted with quality, independence, and institutional legitimacy, i.e. the traditional professional standards for the ordering of knowledge, information and ideas.⁶¹⁶ As a consequence of their ranking and selection criteria, search engines are helping to disrupt the traditional boundaries between different spheres and hierarchies of information, such as the difference between independent information about commercial products and advertising, publications by experts and non-experts, and relevant facts and research and related fantasies and pseudo-science. General purpose search engines do relatively little to discern between these categories.⁶¹⁷ Experts or laymen, politicians or activists, professional journalists or amateur bloggers, well known or little known people, companies or non-profit organizations, advertisers and independent product reviewers, all seem to enter the search index on the same footing and are allowed to acquire prominence on search engine results pages.⁶¹⁸

Although the centrality of major search engines in the networked information environment might lead one to conclude that they are a new type of powerful gatekeeper, search engines hardly exert strict gatekeeping power over the accessibility of information and ideas *independently* of information providers and end-users.⁶¹⁹ The search engine is an interactive navigational tool, both in relation to information providers that want to reach an audience as well as in relation to its end-users who are entering and reformulating their queries. The communicative process between a search engine and an

⁶¹⁶ See Marres & De Vries 2002, p. 188.

⁶¹⁷ See e.g. Zittrain 2008, p. 131. It has to be noted that, the erosion of these categories is a broader societal phenomenon. On top of that, many of these categories are simply replicated in the hierarchies in search media, with or without the help of search engine providers.

⁶¹⁸ Willingly or unwillingly.

⁶¹⁹ See Röhle 2010.

end-user is to a large extent shaped by the end-user's input and choices. Search engines tend to rely on end-users to decide for themselves what to search for and where to go. The statistical character of the predictions of relevance also implies that they must be characterized as the mere *suggestion* of possible relevance, not a statement that these references *are* in fact relevant to the user.

Search engines usually rely a great deal on the information providers they reference. Information providers tend to control whether they are included in the first place and can influence their ranking in major search engines in various ways and degrees. Webmasters receive specific guidelines how to optimize their rankings in Web search services. The somewhat open nature of the private governance of search engine selection and ranking decisions and the possibility to influence and manipulate rankings has also given birth to a thriving industry of search engine optimization (SEO). As a result, the accessibility of online destinations in a particular search engine is not all but carved in stone, but instead is the result of a variety of interrelated streams of communications flowing in and out of the search engine's service.

In view of the complexity of selection and ranking decisions by Web search engines, it may not be a surprise that information law and policy did not yet develop a proper understanding of the legal status of these decisions, let alone their protection under the right to freedom of expression. Most of the discussion of the proper legal treatment of these decisions by search engines in view of their impact on public and private interests has taken place in academia. Still, there are some legal cases and regulatory developments which directly implicate the protection of selection and ranking decisions under the right to freedom of expression, some of which will be discussed in the remainder of this chapter and the next chapters and which start to give an answer about the implications of the right to freedom of expression for their legal governance.

What is at stake in this discussion is the freedom of search engines to decide which selection and ranking algorithms to apply to the material in their index and their legal accountability for making specific decisions. Even if one accepts that there may not and should not be a legally correct way to select and rank references, it is possible that there are unethical as well as legally problematic ways of selecting and ranking search results.

There is very little case legal material to work with in Europe, but it is probable that general applicable law, such as tort law, privacy law or competition law, does limit search engines in their ability to apply certain selection and ranking algorithms at will. At the same time, the status of selection and ranking decisions under the right to freedom of expression should inform the proper application of these laws. Tort law, for instance would probably prevent search engines from deliberately ranking results in a way that damages certain parties without any proper justification, for instance by not selecting or devaluating a particular e-commerce provider while treating its direct competitors quite favorably. European data privacy and anti-discrimination laws probably may restrict the deliberate selection and ranking of results on the basis of sensitive personal characteristics, such as the ethnic background of the targets of search engine queries.⁶²⁰ More specifically, image search engines that would allow end-users

⁶²⁰ See Article 29 Data Protection Working Party 2008, p.13. For Google's reaction, see Fleischer (Google) 2008, p. 3-5.

to search for material on the basis of images of people, applying facial recognition technology, may be legally problematic in Europe.⁶²¹

Competition law may restrict the freedom to rank and select material in an anti-competitive fashion. In particular dominant search engines have a great impact on the underlying markets as well as the markets they operate in themselves. They could use their dominant position to harm competitors. Google for instance has been argued to have foreclosed competition in the market for online geographic services by starting to integrate its own service into its search engine (Google maps). In 2010, the European Commission has started to investigate a similar complaint relating to anti-competitive behavior with respect to other (vertical) search engines.⁶²² There are a variety of other ways in which Google could compete assertively in the way it ranks and selects search results. It could, for instance, promote the millions of websites in its search results which are using the AdSense service, thereby impacting the broader market for online advertising services or prioritize YouTube in video search.

In the United States there have been a few cases involving the application of general laws in which the defending search engine claimed, and in one of them won, wide discretion over its selection and ranking decisions on First Amendment grounds.⁶²³ These cases, including a case about the refusal to run particular advertisements in sponsored search results,⁶²⁴ sparked debate about fairness and access in the search context, issues which will be discussed in more depth in Section 8.5.⁶²⁵ Although it is questionable how much weight should be attached to these particular decisions, a discussion of them can help to clarify the issues that are at stake.

The most famous of these decisions involved a conflict between Google and SearchKing, an online intermediary that seems to have been a relatively successful business because of a particularly good ranking in Google's search results. SearchKing was not an e-commerce provider itself, but a paid-inclusion directory for local Oklahoma businesses. Sometime in 2002, SearchKing dropped in Google's results resulting in significantly decreasing revenues for its clients, and a blow to its business model. SearchKing reacted by suing Google for tortious interference with contractual relations, seeking injunctive relief. Google defended the deliberate devaluation of SearchKing in its rankings by referring to manipulative interference and the low quality of the information offering of SearchKing and its network of websites. Google also claimed that its rankings represented protected speech under the First Amendment. The U.S. District Court for the Western of Oklahoma agreed that ranking decisions were

⁶²¹ See Van Hoboken 2008a.

⁶²² See Dakanalis and Van Rooijen 2011.

⁶²³ *Search King, Inc. v. Google Tech., Inc.*, No. CIV-02-1457-M, at 6–12 (W.D. Okla. May 27, 2003). *KinderStart.com, LLC v. Google, Inc.*, No. C 06-2057 JF (RS), 2007 U.S. Dist. LEXIS 22648, at 13–15 (N.D. Cal. Mar. 16, 2007). In *KinderStart*, the court did not have to address Google's First Amendment defence. More recently *Habush v. Cannon*, Case No. 09-CV-18149 (Wis. Cir. Ct. June 8, 2011). In *Habush* the Wisconsin court concludes that “the use of a computerized system to sequence search results is not speech.” This is based on its consideration that “the hidden process which causes the link to appear at all [...] is content neutral”, which is unconvincing considering the way in which search results ranking and selection is triggered by content.

⁶²⁴ *Langdon v. Google, Inc.*, 474 F. Supp. 2d 622, 629–30 (D. Del. 2007).

⁶²⁵ See Bracha & Pasquale 2008; Grimmelmman 2008; Pasquale 2008, pp. 83-84 (2008).

protected under the First Amendment. It concluded, with reference to a 10th Circuit case (*Jefferson County*⁶²⁶) that

*“A PageRank is an opinion – an opinion of the significance of a particular web site as it corresponds to a search query. Other search engines express different opinions, as each search engine's method of determining relative significance is unique. There is no question that the opinion relates to a matter of public concern. [...] 150 million search queries occur every day on Google's search engine alone. [...] PageRanks do not contain false factual connotations. While Google's decision to intentionally deviate from its mathematical algorithm in decreasing SearchKing's PageRank may raise questions about the “truth ” of the PageRank system, there is no conceivable way to prove that the relative significance assigned to a given website is false. A statement of relative significance, as represented by the PageRank is inherently subjective in nature. Accordingly, the Court concludes that Google's PageRanks are entitled to First Amendment protection.”*⁶²⁷

The Court subsequently concluded that the sought injunction would chill protected speech and would be adverse to the public interest. Actually, it went even further. It concluded that search engine rankings cannot give rise to a claim for tortious interference with contractual relations because they cannot be considered wrongful, even if the speech is motivated by hatred or ill will. Hence, the central premise of the Oklahoma Court seems to be that because there is no correct way to select and rank references there can also be no legally incorrect way.⁶²⁸ Hence the decision how to select and rank references is seen as the expression of an opinion, which is strongly protected by the First Amendment. The societal impact and possible damage should play out in the free market place of ideas. This reasoning is similar as in *Tornillo*, in which the Supreme Court established the First Amendment protection of newspapers' editorial freedom. The Oklahoma Court could have added that it has yet to be demonstrated how government interference on the crucial process of selecting and ranking websites by search engines could be exercised consistent with First Amendment guarantees.

The legal consequences of the *SearchKing* judgment's conclusion of bringing selection and ranking decisions under the scope of the First Amendment may obscure the fact that to do so may be quite reasonable. Leaving the predominantly technological and statistical nature of search engine selection and ranking decisions aside, the selection and ranking of references from the index in return to a user's query seems highly comparable to an editorial activity, in many ways comparable to the decision of newspapers what to put in the paper and in which place. This may imply that these decisions should in fact be similarly protected, also in the European context under Article 10 ECHR. Notably, the result of bringing these decisions under the scope of the editorial freedom of the media protected by Article 10 ECHR, would in no way have the same implications for the possibility to limit this freedom to select and rank at will in specific circumstances, for instance when a search engine provably ranks a website lower for the sole reason it wants to harm the respective information provider.

⁶²⁶ *Jefferson County Sch. Dist. No. R-1 v. Moody's Investor's Services, Inc.*, 175 F.3d 848 (1999)

⁶²⁷ *Search King, Inc. v. Google Tech., Inc.*, No. CIV-02-1457-M, at 9 (W.D. Okla. May 27, 2003).

⁶²⁸ *Id.*, at 6–12.

Another interesting aspect of the *SearchKing* decision is that despite automating most of its selection and ranking decisions, Google did not to give up its right to select and rank individual sites manually and differently. More generally, the technological nature of the decisions about which selection and ranking algorithms to apply may lead some to conclude that these decisions are of a different nature than those made by humans and that this difference should impact the legal protection of these decisions under the right to freedom of expression.⁶²⁹ Search engines themselves tend to contribute to the view by claiming that their search results are mechanical, neutral and objective. Newspapers may have been expected to rely on certain well-respected organizational mechanisms for editorial decision making, in line with the ethics of journalism. However, the real criteria for the newspaper's contents often remained equally unknown. In fact it would be considered an interference with the newspaper's editorial freedom, if the law were to mandate complete transparency in these matters. Moreover, automation in Web search engine governance is to some extent a result of the massiveness of the Web and the related massiveness of the index. It is simply humanly and economically impossible to provide human evaluations of the relevance of websites for all the queries search engines respond to. And automated decision-making about the value of information is not only economical. Digital technologies have made it possible to incorporate the daily feedback from close to or over a billion websites in the editorial processes of the search engine. As a result, the automated nature of the selection and ranking decisions are directly related to the societal benefits of having modern accessibility technologies and services that are able to keep up with the abundance of information online.

A major problem in discussing the protection of ranking and selection decisions is that the actual ranking and selection criteria that search engines use are rather opaque and thereby impossible to evaluate specifically. As discussed in Chapters 2 and 3, search engines tend to keep details about their algorithms secret. Because of the lack of transparency about the operation of search engines, it is unclear whether the underlying principles for the governance of accessibility of particular search engines are informed by commercial, ideological or scientific principles. Arguably, by refusing to be transparent about the underlying principles for the selection and ranking of references, search engines refuse to publicly take responsibility for the search results they offer to their users. Thus, one could argue that search engines can hardly claim particular protection for the kind of selection and ranking decisions they make, unless they would be more explicit about those decisions.

Arguably, the comparison of the editorial freedom of the Web search engine to the editorial freedom of the press is problematic due to the commercial nature of search engines as well as the commercial nature of much of the speech they facilitate. This leads to another case about the First Amendment status of selection and ranking decisions, which involved the application of a general applicable law to the selection and ranking mechanism of *Roommates.com*, a platform and search engine for housing opportunities. The selection mechanisms were considered to be in violation of the anti-discrimination provisions of the Fair Housing Act (FHA). Roommates' interactive service was facilitating its users' individual and sometimes discriminatory sexual and racial housemate preferences by allowing them to express these preferences and subsequently matching users on the basis of these preferences. Roommates defended itself against the claimed breach of the Fair Housing Act by asserting the

⁶²⁹ On the question about protection of software under the First Amendment more generally, see Halpern 2000.

Communications Decency Act, section 230, and its First Amendments rights to free speech and intimate association.

The Roommates case is best known because of the 9th Circuit's dismissal of Roommates CDA 230 defense, on the basis that the service was considered a speaker itself as regards the selection technology.⁶³⁰ The next line of defense, a First Amendment claim against the application of the FHA, was addressed by the U.S. District Court for the Central District of California after remand. The District Court argued as follows about the status of selection decisions of the service:

*"it is far less obvious that Roommate's prompting of its users to provide personal characteristics in order to use its service, or its use of that information in its search and matching functions, is "speech" protected by the First Amendment. Instead, those functions performed by Roommate, and the questions it prompts its users to answer, could be considered mere conduct rather than either speech or a communicative act expressing a viewpoint to which the First Amendment applies. [...] However, even if Roommate's prompts for discriminatory information, searches, and matches could be considered speech, it would be, like Roommate's publication of its users' discriminatory preferences, speech of a commercial nature."*⁶³¹

After this conclusion, the Court proceeded by applying the test for the regulation of commercial speech,⁶³² which ended up in favor of the constitutionality of the application of the Fair Housing Act on the Roommates service.

It seems questionable whether the application of the standard for commercial speech is warranted on selection decisions by referencing services. Why would the commercial nature of a search engine as well as the commercial nature of many of the websites in its index have to implicate on its protection under the right to freedom of expression in principle? The U.S. District Court's conclusion seems at least inconsistent with the conclusion about the same questions in the context of the press. In *Sullivan* the Court explicitly considered that the paid-for nature of the advertisement did not matter for First Amendment purposes, while pointing to the contribution of the commercial business model to the ideals underlying the right to freedom of expression.⁶³³ Hence, the fact that Roommates provided its services for remuneration should not automatically have had implications for the protected status of its communicative actions.⁶³⁴

Both the *SearchKing* and the *Roommates* example relate to the application of generally applicable laws to selection and ranking by search engines and the question about the implications of the right to freedom of expression in such cases. However, legislatures could – in theory – decide to establish ranking obligations, for instance to increase the relative accessibility of certain information and ideas. Broadcasting law already contains obligations with regard to the carriage and minimal representation of

⁶³⁰ On CDA, Section 230, see Section 6.4.4. See also Section 9.3.3.

⁶³¹ *Fair Housing Council v. Roommate.com, LLC*, CV 03-9386 PA (C.D. Cal. Nov. 7, 2008).

⁶³² See Section 5.5.3.

⁶³³ See Section 5.3.2.

⁶³⁴ See also Ardia 2010; Wilemon 2009.

certain broadcasting products, such as local broadcasting or programs in the interest of specific national minorities in view of the demands of pluralism. The possible extension of these obligations to minimal rules for representation in electronic programme guides for audiovisual media services would amount to sector-specific search engine regulation. Recently, a proposal for a ranking obligation was put forward in France in the context of the issue of online copyright infringement.⁶³⁵ It would require search engines to promote online services indicated as constituting a legal offering of content online by a new public authority through higher rankings.

Clearly, the selection and ranking of search results by dominant search engines is of major importance in the relation between search engines and end-users on the one hand and information providers and search engines on the other hand. The function of the search medium is to help Internet end-users find valuable websites. They also help information providers to reach an audience.

From the perspective of search engines as intermediaries between information providers and end-users, some have expressed concern about the unfair exclusion or unfair treatment of information providers by dominant search providers.⁶³⁶ This debate has focused on the proper resolution of the conflicting interests of information providers to reach an audience without undue interference by search engines or other selection intermediaries and the discretion of search engines to decide about the composition of their index and the ranking of results. In Section 8.5 the idea of a fair ranking obligation that has been proposed will be discussed in more depth, in particular the extent to which such obligations could be based on the right to freedom of expression of information providers, as for instance in the context of common carriage obligations on Internet access providers.

Other restrictions on the freedom to select and rank at will could be informed by the interests of end-users. The selection and ranking of results is at the basis of making search engines work for end-users. They will evaluate the quality of search engines in light of the information they are presented in return to their queries. Any ranking obligation therefore, if not specifically informed by and enhancing the rights and interests of end-users, runs the risk of harming them in their interest to freely inform themselves. It is possible to imagine several types of obligations with regard to ranking and selection that would be informed by the interests of end-users. Some have argued, with reference to democratic theory, that search engines should become more open and allow end-users to control the ranking of search results.⁶³⁷ Not the search engine, but the end-user would become the true editor of search results.

End-users do have an indirect impact on the ranking of results. Increasingly, selection intermediaries personalize their offering on the basis of the end-user's location and behavior. They build profiles of their users to facilitate further personalization and better targeting of the advertisements that are sold. In other words, search engines have started to discriminate between end-users. If such discrimination

⁶³⁵ See Manara 2009.

⁶³⁶ See e.g. Chandler 2008, Paquale & Bracha 2008, Van Eijk 2006, Van Couvering 2009; Introna & Nissenbaum 2000. See also Grimmelmann 2007; Goldman 2006. For a discussion of a documented incident of the manipulation of specific rankings see Rogers 2009a.

⁶³⁷ See Rieder 2009.

would start to harm certain groups of end-users, this could be another reason for governments to intervene. One of the ways the law could possibly intervene is by granting end-users more control over the personalization of information services and, for instance, offering the opportunity to opt-out of profiling or to edit and share their profiles. At the same time, governments can try to use and courts can take account of existing personalization based on location to establish jurisdiction over search engines, maybe even mandating such personalization, thereby limiting end-users' ability to switch between different localized versions of search.⁶³⁸

There is already one legal obligation with regard to the ranking and presentation of search results which is widely accepted, namely the obligation to delineate sponsored results from organic search results. The typical search results page contains two types of references: on the one hand references to information providers that are shown without financial reward, and references that are shown for payment. Both in Europe and the United States it is considered a legal obligation that search engines be transparent about this difference. These and other possible obligations on search engines in view of the communicative freedom of end-users will be discussed in section 8.6 and in more depth in the Chapter 10.

8.4.4 Search engines: editorial control, freedom, duties and responsibilities

The ECtHR's arguments in *Open Door* discussed in section 8.4.2 imply that the reliability of referencing information and the professional context in which information is communicated strengthens the protection a communicator receives under Article 10 ECHR. This is in line with the Court's case law on the duties and responsibilities that are tied to the exercise of the right to freedom of expression in Article 10 ECHR and which will be discussed in more detail below. Before asking about the proper duties and responsibilities of search engines when selecting, ranking and publishing references, the question should be addressed how much responsibility search engines actually take in the context of the governance of the composition of their index and the selection and ranking of search results. This question can help to further qualify the way in which the selection and ranking decisions of the search engine should be seen as editorial ones.

In the context of the press, the newspaper's decision to include or exclude is conceptualized as editorial and protected by the right to freedom of expression. Editorial freedom is one side of the coin, editorial responsibility the other. By publishing a story, the newspaper not only exercises its control over the contents of its medium, but it also takes responsibility for the decision that the story is worth being printed. This kind of responsibility is typically absent in the case of a search engine's inclusion of references. Due to the navigational function of search engines for the Web, the fact that information is available online is itself a reason to include it in the index. Moreover, not only do general purpose search engines not try to control everything that goes into their index, they also do not tend to know and neither try to monitor everything that can be found with the use of their services.

⁶³⁸ Several European countries have established specific requirements for country specific search services but neither of them requires them to block access to other country-specific sites entirely. In China, Google's non china specific version was regularly blocked. See also Section 9.2.

It is important to note that editorial control by search engines over the contents and optimization practices of the websites they reference is not necessarily absent either. Search engines can easily exert control over inclusion and exclusion of references and actually do so in a number of instances. Search engines have to guard the quality of search results against so-called 'spam', meaning websites which unduly interfere with the automated processes of indexing and ranking. Spam protection is typically based on the enforcement of webmaster guidelines and minimal standard of relevance. It does have some parallels with the filtering of unsolicited e-mail by Internet access providers. Spam protection in the context of search engines is necessary precisely because of a lack of editorial control by search engines with regard to the inclusion and ranking of websites. If a library would allow the public to place books on the library shelves and insert references into the library catalogue, they would probably have to assert a similar type of oversight to delete references and manipulation of the system.

Of course, search engines also evaluate the quality of their search results more generally and change their algorithms and search engine rankings accordingly. General purpose search engines systematically analyze the behavior of their users by looking at massive amounts of user data. They also let human evaluation teams look at the quality of search results.⁶³⁹ Sometimes, the evaluation may lead to direct manual changes to search results, for instance in specific cases of search engine manipulation by content providers. Sometimes, the solution to what is considered a problem of search quality is solved by introducing new algorithms. Reported examples include the degrading of the rank of directories with paid-for references in view of the lack of original content, such as *SearchKing* discussed above, as well as the defusing of the so-called Google Bomb, where Google used to return the website of the President of the United States, George W. Bush in return to the query [miserable failure].⁶⁴⁰

Looking at dominant general search engine practices, they seem to aim for comprehensiveness of their index on the one hand, while on the other hand dealing with matters of information quality through their selection and ranking algorithms, which reflects the observation of Marres and De Vries discussed earlier in this chapter. This does imply that they cannot defend the availability of the references in their index with the argument that they are all carefully selected. Still, the decision of search engines to include websites in the index should probably be considered an editorial decision as much as specific decisions to change the selection and ranking algorithms. The question about the editorial nature of the decision to in or exclude is not only relevant when assessing the protection against restrictions on the communication of references. It is at least as relevant when addressing the possible conflicts of interest between information providers and search engine providers about access. In the context of Internet access providers this issue was addressed in detail, in particular the question of to what extent an access provider would be able to claim a right to freedom of expression not to provide access, as is the case with the press. This led to a similar discussion whether the decision of an access provider to block communications between information and end-users could be considered editorial.⁶⁴¹

⁶³⁹ See Google 2007b; Sullivan 2010a. See also Section 3.3.1.

⁶⁴⁰ See Lettice 2009. See also Sullivan 2007.

⁶⁴¹ On the question of 'editorial freedom' of Internet access providers, see Section 6.3.2 and Section 6.4.4.

Unlike in the case of Internet access providers, the relation between information providers and search engines, while also considering the interests of end-users, arguably entails a genuine conflict of protected interests under the right to freedom of expression. The reason is that there is a variety of reasons to exclude or disregard particular sources of information, based on the role of search engines in the networked information environment. This role does include consideration of the quality and relevance of information.

An example may help to illustrate the fact that there may not be a single right way to make these choices. Consider the relatively simple query for [travel insurance]. There is a variety of actors and types of information that could be argued to deserve representation in the set of results the end-user receives. Of course, travel insurers would like to be found. And if many end-users are using the search engine service, insurers would probably be interested in paying for additional representation in the form of advertising. There are many other sources of information that could be included in search results, such as independent consumer organizations which offer advice about the quality of travel insurance and independent bloggers and customers who have written something about the current market's offering. Some may argue that the second type of information should be considered more valuable for end-users. Others may respond to that prioritization with the complaint that this would be unfair for the first category, while it would also force insurers to buy advertising to reach end-users at all. And these types of information providers are only the beginning of a list of conceivably relevant sources of information that could be selected by a search engine. All sorts of other information related to travel insurance, and insurance more generally, could be included. Notably, some of these information providers can be expected to actively compete for user attention in search engines, whereas other information providers would remain passive.

To be able to make these choices at the scale at which general purpose search engines operate, search engines rely on automated decisions and machine learning involving a variety of statistical predictions about the meaning and value of the queries and the information in the index. This may sometimes produce results that would not have been selected by a human editor and could dissatisfy specific information providers or parts of the general public.⁶⁴² However, since the amount and variety of queries can simply not be handled by humans and the Web is constantly changing, Web search engines cannot be expected to make a careful *ex ante* and fair evaluation of the set and ranking of references for each single query, such as the query mentioned above.⁶⁴³

To a large extent, the legal debate about the possible obligations of search engines relating to the composition of their index has focused on obligations to remove references to unlawful or harmful material. Some commentators, however, have argued that dominant search engines should be restricted in their freedom to exclude material from their index and should act as common carriers. This issue will be discussed in more depth in the next Section.

⁶⁴² See also Section 10.2.4.

⁶⁴³ As mentioned above, major search engines do tend – and acknowledge – to have quality control by human evaluators. See Google 2007b; Sullivan 2010a.

Like any other entity that would assert protection under Article 10 ECHR, a search engine will carry duties and responsibilities that are tied to the exercise of its rights. The question is what these duties and responsibilities could be argued to entail in the search engine context. As was shown in Chapter 5, according to the ECtHR the duties and responsibilities, ex article 10 ECHR, depend on the technical means used for expression and dissemination. The Court has stated that:

*"[...] whoever exercises his freedom of expression undertakes "duties and responsibilities" "the scope of which depends on his situation and the technical means he uses. The potential impact of those means must be taken into account when considering the proportionality of the interference."*⁶⁴⁴

To which the Court added in its more recent judgment *Stoll*:

*"These considerations play a particularly important role nowadays, given the influence wielded by the media in contemporary society: not only do they inform, they can also suggest by the way in which they present the information how it is to be assessed. In a world in which the individual is confronted with vast quantities of information circulated via traditional and electronic media and involving an ever-growing number of players, monitoring compliance with journalistic ethics takes on added importance."*⁶⁴⁵

These last considerations do seem to resonate in the context of search engines. In line with the Court's reasoning in *Open Door*, it seems to imply that the more a communicator does to enhance the quality of information, the more it can claim protection. In particular, professional standards with regard to such quality and the mode of communicating serve as an extra line of defense against interferences. Most search engines, as noted above, do fairly little to guarantee a threshold of minimum quality of the information to which they provide references. Generally, the lack of editorial control and oversight over the inclusion of references could weaken a search engine's protection against interferences under Article 10 ECHR. And if the average quality of search engine references were to be deemed very low, they could hardly claim any protection against interferences with their right to freedom of expression under Article 10.

But there are good arguments in favor of less stringent duties and responsibilities as regards the quality of references. It is important to note here that the duties and responsibilities can cut both ways. On the one hand, they can be argued to imply responsibility on Web search providers to guarantee the quality of references. On the other hand, they could be argued to entail a duty on search engine providers to be comprehensive and not to exclude references lightly. From the perspective of quality, one could argue that entities that govern the accessibility of information will have to be the ones to sift the wheat from the chaff through filtering, the exclusion of references and the imposition of hierarchies informed by conceptions about information quality. But from the perspective of access, the entities that govern accessibility should be inclusive instead of restrictive.

⁶⁴⁴ ECtHR 14 June 2007, *Hachette Filipacchi Associes v. France*, § 42

⁶⁴⁵ ECtHR 10 December 2007, *Stoll v. Switzerland*, § 104.

In reality, the choices of search engines with regard to the inclusion of references fall somewhere in between the two poles. Although Web search engines tend towards comprehensiveness, a truly inclusive search engine simply does not exist. Large parts of the Web remain excluded from search engines, because of technical reasons such as file formats, crawling speed and the size of the Web, because of compliance with webmaster instructions not to crawl their content,⁶⁴⁶ because of the exclusion of manipulative results, because of various legal reasons, and because of self-regulation with regard to illegal and harmful content. At the same time, none of the major search engines exercises the kind of editorial oversight with regard to the quality of specific references as discussed in the example of the query [travel insurance] above.

And of course, the important question from a legal perspective is not whether and in which ways search engines could start to guarantee the quality of their references but whether they ought to do so. By exercising stronger editorial control over the composition of their indexes, search engines would end up being much more restrictive and it would be likely that much of the Web would become harder to find for end-users. In *Open Door* the ECtHR referred to the normative principle that a communicator that leaves the decision to act upon its communication to the receiver cannot – in principle – be blamed for those decisions. The more passive – or maybe better facilitative – the communicator is with regard to the decisional autonomy of its audience, the more protection it receives against interferences that seek to prevent a certain reaction by the audience. The protection search engines would receive would therefore be enhanced by their facilitative role as regards the interests of end-users.

These two perspectives lead back to the paradox of access and quality. From the perspective of access, search engines should help end-users to navigate the entire Web. Any exclusion of references from the index would interfere with this primary function. The ECtHR's case law suggests that search engines can defend a choice for inclusiveness with the argument that it is precisely their societal role, in the interest of facilitating accessibility, to provide references to all the sources of information on the Web. The scale at which general purpose search engines operate implies that they cannot check the lawfulness and quality of all the information they refer their users to. This may in particular be the case for the non-dominant search engine providers which most likely have fewer resources to proactively deal with issues of quality in their indexes. They could do so by prioritizing websites they consider of high quality, while still aiming for comprehensiveness of their indexes. That being said, a search engine that put more energy into reviewing websites and limited its index to references of established quality and relevance could arguably claim protection under the right to freedom of expression to defend its choice to exert editorial control over its index.

It is striking that while claiming their freedom of expression right to rank information providers freely, search engines do not readily take responsibility for the subjective nature of some of the decisions they make in the process of determining winners and losers in the networked information environment. For instance, in response to a favorable ranking of the anti-Semitic site Jewwatch in return to the query [Jew], Google responded that its

⁶⁴⁶ Notably, compliance with robots exclusion protocols is sometimes argued to be mandatory. See also Section 10.2.3.

*“search results are generated completely objectively and are independent of the beliefs and preferences of those who work at Google.”*⁶⁴⁷

Since 2007, the words “*completely objectively*” in this explanation about Google search results have been omitted. To be fair, in both statements Google does take some responsibility for its choice not to manually interfere with the favorable ranking of *Jewwatch.com*. Strikingly, however, the statement now contains an even more sweeping statement of disinterest in the quality of its results:

*“The beliefs and preferences of those who work at Google, as well as the opinions of the general public, do not determine or impact our search results.”*⁶⁴⁸

Considering the dominant position of Google in the search engine market, there may be some room for improvement here. In fact, the potential impact of the medium and the nature of the content that can be found through a search engine are considered relevant for its possible duties and responsibilities which exist in the European context. On several occasions the ECtHR clarified that the particular impact of media has to be taken into account when considering the permissibility of interferences by public authorities. For instance in its *Jersild* judgement, the Court concluded as follows with regard to the nature of audiovisual content.

*“In considering the “duties and responsibilities” of a journalist, the potential impact of the medium concerned is an important factor and it is commonly acknowledged that the audiovisual media have often a much more immediate and powerful effect than the print media [...]. The audiovisual media have means of conveying through images meanings which the print media are not able to impart.”*⁶⁴⁹

While the Court seems to refer to the impact of a certain medium, such as audiovisual media in this case, the question could be asked of whether the same logic could be applied to the context of search media, dominant providers such as Google in particular. Looking at Web search media, it could be argued that major general purpose search engines with a particular strong societal impact, would have enhanced duties and responsibilities based on their widespread use and their impact on the different value chains in which they operate. As a result of particular selection and ranking decisions, dominant search providers like Google may have a disproportional impact on the governance of accessibility. The need for particular legal restrictions on their freedom to make access, selection and ranking decisions which would harm the legally protected interests of others, including the communicative liberties of information providers and end-users, may be more easily established. Such restrictions could be argued to reflect enhanced duties and responsibilities based on their dominant position. However, the need to respect the freedom of search engine providers to make their mediating choices freely would remain and interferences with their protected freedom would need to be both effective and proportional.

⁶⁴⁷ See Google 2007c.

⁶⁴⁸ See Google 2007d.

⁶⁴⁹ ECtHR 23 September 1994, *Jersild v. Denmark*, § 31.

8.5 Search engine freedom and the interests of information providers

8.5.1 Introduction

The interest of information providers in the governance of Web search engines is to be present on the search engine's platform and to find their way to an audience. What is at stake for information providers can also be formulated in terms of representation. If no search engine includes a particular source of information, this would deprive it from an important means to acquire attention and legitimacy. Search engines, dominant search engines in particular, help to establish the winners and losers in the competition for end-user attention.

The impact of dominant search engine selection and ranking decisions goes beyond the mere accessibility of information and ideas. The impact of Google's search platform means that representation in Google is directly related to the degree of success or failure of many online activities, be they economic or political.⁶⁵⁰ Moreover, several studies show that end-users consider a high ranking in a popular search engine as an independent sign of information quality.⁶⁵¹ For many, inclusion in Google's results alone has become an important point of reference with regard to the mere 'existence' of online information. This leads to the belief that all that is excluded from Google and other general purpose search engines can be legitimately excluded or disregarded elsewhere. In fact, Google's ultimate goal, according to one of its founders, is to make itself indispensable for anyone interested in anything.⁶⁵² The search engine index is marketed to its users as a reliable copy of reality itself.

8.5.2 Search engine freedom and the ideal of unmediated mediation

In the previous section it was argued that the right to freedom of expression may actually protect the search engine operator to make selection and ranking decisions freely, thanks to the editorial character of its mediation between information providers and end-users. It was also shown that search engines do not readily take responsibility for the way they make those choices. When justifying those decisions to rank some information providers worse than others, search engines often refer to their role in relation to end-users. This reference to the interests of end-users can also be found in Jennifer Chandler's characterization of the speaker's interest in their relation to search engines and selection intermediaries more generally. Drawing on the ideal of an unmediated public sphere, Chandler stresses

*"the right to reach an audience free from the influence of extraneous criteria of discrimination imposed by selection intermediaries. If selection intermediaries block or discriminate against a speaker on grounds that listeners would not have selected, that speaker's ability to speak freely has been undermined."*⁶⁵³

The freedom of information providers to reach an audience thereby becomes based on the freedom to select of (specific) end-users. This principle as formulated by Chandler, comes closest to the ideal of

⁶⁵⁰ See the discussion of *SearchKing* in Section 8.4.3

⁶⁵¹ See Hargittai 2007b.

⁶⁵² See Levy 2011, pp. 66-67.

⁶⁵³ See Chandler 2008.

universal access: if discrimination or blocking of sources takes place, it should be transparent and end-users should be able to control it. The filters for adult content that are typically installed on major search services are an example of the adherence to this principle as formulated by Chandler.

But one can raise a number of objections to the application of this principle to the governance of search engines more generally. First, the role of search engines is to help users find valuable information and ideas. To be able to offer their service they have to make some choices about the relative value of information and ideas for end-users. Second, it is debatable whether selection intermediaries should not be allowed to prioritize information and ideas in a way that does not directly reflect end-user preferences. Why would it be wrong for selection intermediaries to try to represent a variety of different speakers and sources of information in their search results and *not allow* end-users to block or discriminate between results? In other words, it is possible for selection intermediaries to add value by making a determination independent of end-user preferences. The ideal that the public library adds value by confronting its patrons with its entire collection and presents its patrons with a carefully selected corpus of knowledge and cultural heritage which may not all be according to their preferences are examples of this.

Overall, search engines may be in a better position to decide which information is useful, relevant or valuable than end-users themselves, not excluding the end-users' possibility to critically evaluate these decision afterwards. In fact, this is what search engines are doing all the time. If research about information searching behavior is to be taken seriously, end-users are looking for relatively easy answers, not an incredible range of choices of high quality for every specific search.⁶⁵⁴ Generally, it seems fair to assume that end-users are relying on search engines to make intelligent selection and ranking decisions for them, instead of having to make all those decisions themselves. From the perspective of the public interests in the opening up of the Web, understood as the process of connecting information and ideas online to their societal use, it may be better to focus on the quality of the decisions Web search engines end up making than to deny them the freedom to make these choices in the first place.

8.5.3 Information providers' control over search engine governance

With regard to the interests of information providers to reach an audience through search engines, it is important to acknowledge the amount of control they have over search engine governance in practice. Dominant search engines have given information providers a lot of control over their indexing, ranking and the content of references. Information providers can typically control whether they are indexed in the first place by using instructions like robots.txt. They can influence how they and others are ranked by following 'white hat' search engine optimization guidelines, for instance by integrating their information offering through the establishment of links with the rest of the Web. They can control how they are presented when they are included by stipulating a title of their page or offering a so-called site map.

⁶⁵⁴ Research on information seeking behaviour shows that most users stop searching when they found a satisfactory answer, not the best answer. See Section 3.3.1.

In fact, search engine governance currently entails a degree of control for information providers over the governance of accessibility by Web search engines which may be antithetical to the interests of end-users.⁶⁵⁵ As mentioned above, webmasters *de facto* have complete control over inclusion of information in search results. Even a popular publicly accessible website can exclude itself from all major search engines with the simple placement of an instruction to search engines on its website. And the search engine optimization industry, which is a direct result of the reliance of search engine rankings on third party signals, is paid by information providers to reach end-users, not to enhance the fair representation of sources of information and ideas in search results.

This begs the question about possible legal restrictions on information providers to compete unfairly – for attention – through Web search engines. Currently, the formulation and enforcement of search engine optimization guidelines is left to search engine providers. The public interest in search engine quality may warrant regulation of so-called Web spam, i.e. websites exerting undue influence on search results through their deceptive publications. Web spam is perhaps not as visible as e-mail spam but it is a major problem.⁶⁵⁶ The possibility to influence search engine results and rankings is the main driver of the SEO industry.⁶⁵⁷ The problem is that unlike in the case of e-mail, the distinction between Web spam and legitimate publications is hard to make without restricting the right to freely publish on the Web. In general, there is no such thing as an unsolicited online publication. Of course, there are a range of legal restrictions on the lawfulness of online publications, but these are not directly related to the context of search engines. Defamation law, trademark law and a range of other laws restrict the freedom of information providers to impart information and ideas online, they similarly restrict the freedom of information providers to reach an audience through search engines. It may be very hard to formulate effective and legitimate additional legal restrictions on publications because of their negative impact in search engines, but do not restrict publication of lawful material online.⁶⁵⁸

8.5.4 Restricting lawful information from entering the search engine index

Increasingly, the effectiveness of search engines in opening up information and ideas online leads to legal and regulatory pressure on the responsibility of the actual publishers. Generally, the global nature of the Internet produces this effect. If national States would all be allowed to claim jurisdiction over all online publications, we could end up with a global lowest common denominator of legally permissible online speech.⁶⁵⁹ In the context of search engines, a similar tendency would be observed to have information providers not only carried responsibility for the decision to publish certain information and ideas, but also for the subsequent further dissemination of the publication by search engines. The result could be that in certain cases, it would end up being illegal to publish online what could have been legally published offline.

⁶⁵⁵ See e.g. Segal 2010. See also Singhal (Google) 2010.

⁶⁵⁶ See Section 3.3.2. See also Section 10.2.3 and Section 10.3.1.

⁶⁵⁷ See e.g. Websense Security Labs 2009.

⁶⁵⁸ The *Zwartepoorte* case in the Netherlands is a good example of this. A webmaster was ordered to change the legal content of its website because of the resulting suggestive effect of the content in the snippets of Google. Gerechtshof [Court of Appeals] Amsterdam 26 July 2011 (*Zwartepoorte*).

⁶⁵⁹ See Mueller 2010, p. 186.

From the perspective of the right to freedom of expression legal restrictions – on information providers – to allow for the further dissemination of *lawful* publications through search engines are problematic. Nonetheless, there seem to be some legal proposals and developments in this direction, in Europe in particular. An example is the restriction on the further dissemination of certain publications through search engines by the use of robot exclusion protocols. More specifically, the law would mandate that information providers cannot let certain publications be indexed, because they are considered to be unsuitable to be opened up through a search engine. European data protection authorities have started to interpret data protection law as posing some limits on the freedom of information providers to make publications involving personal data crawlable.⁶⁶⁰ Under certain circumstances, web publishers are expected to use protocols such as robots.txt to ensure that certain personal data – a legal category in EU law which includes all information relating to identifiable natural persons – cannot be found in search engines, except when visiting the actual site on which they are published.⁶⁶¹ The mandatory standard use of exclusion protocols by social network sites that guarantee that the users of such sites cannot be found in search engines is an example of adherence to this practice. Others have proposed that the safe harbor for third party material on websites, such as comment sections on blogs, should be made conditional on the use of robots exclusion protocols for these parts of the website.⁶⁶² Interestingly, these proposals all rely on the assumption that it would be illegal for search engines to disobey robot exclusion protocols. Although it is widely accepted that it is good practice to obey these instructions, it makes a difference if such good practice is turned into a legal obligation.

The idea that certain sensitive, controversial or possibly illegal information can be published online but should be excluded from search engines may be understandable from the perspective of the concerns relating to the wide publicity of personal and reputational information. Search engines greatly diminish the practical obscurity of the information they index and can facilitate the accessibility of information and ideas of questionable quality. Information that is traditionally made public mandatorily, such as public records containing personal information, is now more easily accessible. Information which was not traditionally published at all, such as conversations and debate between readers in reaction to news and current affairs, has found permanence on blogs, message boards and comment sections.

The low average quality and sometimes illegal nature of such end-user conversations might also warrant leaving these end-user driven conversations, which do not necessarily abide by the standards of responsible journalism, in the obscurity of the un-indexed Web. However, a legal regime which would mandate such exclusions would be problematic from the perspective of the right to freedom of expression. To state it mildly, the idea that public authorities have a role to play in obscuring lawfully published information is difficult to defend.⁶⁶³

It is possible that further self-regulatory practices will emerge with regard to the exclusion of information from search engines. Some argue that the press should exclude older material from their

⁶⁶⁰ See Dutch Data Protection Authority 2007. See also Van Hoboken 2008a.

⁶⁶¹ Dutch Data Protection Authority 2007.

⁶⁶² Kerr 2007.

⁶⁶³ Notably, an exception is the protection of minors against content that is considered harmful. See also Section 9.2.

Internet archives which could negatively impact on the reputation of persons.⁶⁶⁴ The line of reasoning could be that through the doctrine of duties and responsibilities undertaken by publishers of information exercising their freedom of expression under Article 10 ECHR, the absence of a robot exclusion instruction could impact on the determination of the lawfulness of a restriction on an online publication. There are two related objections to such self-regulatory practices, both grounded in the ideal underlying the right to freedom of expression that the societal circulation of information and ideas be “*uninhibited, robust and wide-open*.”⁶⁶⁵ First, it is questionable whether it is good practice for search engines to obey exclusion instructions if information providers use it to try to obscure legal and publicly accessible information which end-users would be interested in reading. Maybe it is ethical for search engines to do exactly the opposite. Second, there is a lot of pressure on search engines to limit the findability of information in their services, precisely because it is one of the primary effective means to find information and ideas.⁶⁶⁶ Information that is excluded is still ‘public’ in theory, but in many ways it simply ceases to exist. The pressure on search engines to prevent any perceived harm resulting from the opening up of the World Wide Web can easily result in the ‘self-censorship’ of perfectly valid online sources.⁶⁶⁷ To limit risk, information providers would simply exclude controversial information on their site from search engines, and search engines would obey these instructions to keep up good relations with the sites they need to crawl and escape possible litigation. The results would be a bias towards uncontroversial information and ideas in search engines.

8.6 Search engines and the freedom of end-users

8.6.1 Introduction

The freedom of expression interest of the end-user of any kind of information service or medium is typically characterized as the freedom to inform oneself freely. The use of search engines has become a daily activity for end-users. End-users rely on search engines to find information about news, products, services, diseases, travel destinations, political candidates and more. Knowledge workers like journalists and academic researchers rely on general purpose search engines for their daily activities.⁶⁶⁸ Ultimately, it is the fundamental interest of Web search engine users to become an informed citizen and consumer that is at stake. It may come as no surprise that there is disagreement about what is needed to facilitate this process.

8.6.2 Search engines and the end-user’s interests: access and quality as conflicting perspectives

The disagreement in the legal and regulatory debate about end-users interests in the context of search engine governance maps fairly well to the conflicting ideals of universal access and quality discussed earlier in this chapter. The ideals of universal access on the one hand and information quality on the other hand correspond with the following conflicting perspectives on the needs of search engine end-users.

⁶⁶⁴ See e.g. Hins 2008.

⁶⁶⁵ *New York Times v. Sullivan*, 376 U.S. 254, 271 (1964).

⁶⁶⁶ See Van Hoboken 2009b..

⁶⁶⁷ See Gerhart 2004.

⁶⁶⁸ See Machill & Beiler 2007, pp. 143-242.

On the one hand, end-users are portrayed as perfectly capable of navigating information and ideas on the Web, precisely because of effective information location tools, which, if possible, should be improved upon to increase the transparency of the corpus of online information for end-users even more. This side would hold that end-users have been liberated from traditional gatekeeping institutions and institutionalized paternalism. They can and should be allowed to decide for themselves what is useful, relevant, harmful, informative or entertaining. More particularly, they should not be hindered by restrictive indexing policies of search engines, possibly as a result of applicable laws and policies, in their freedom to find information online. In line with this perspective, proponents of end-user control have argued that end-users could and should have even more control over their search process, for instance by being able to choose between or completely control the different ranking algorithms and the possible personalization of search results.

On the other hand, end-users are often portrayed as lost, mislead, confused, injured, and otherwise negatively affected by the current state of affairs of accessibility of information through search engines.⁶⁶⁹ This state of affairs is typically portrayed as giving those end-users that want to do harm or commit crimes the tools to do so. For example, there have been a number of proposals and legal measures to limit the ability of end-users to use certain words and signs as queries. In 2007, European Commissioner for Justice and Home Affairs Frattini sought industry cooperation “*to prevent people from using or searching dangerous words like bomb, kill, genocide or terrorism*”.⁶⁷⁰ This particular call for self-regulation was not taken seriously. It’s actually hard to imagine how a search for genocide could be considered harmful or dangerous.

Obviously, both characterizations of end-users – one focusing on empowerment and access, the other on quality and harm – are simplistic projections of particular viewpoints. Still, this might fit the role of listeners, viewers or end-users in media law and regulation more generally, which often remains what Helberger calls ‘*a spiritual one*’.⁶⁷¹ The perceived interests of end-users were and remain decisive in regulatory debates about media freedom, but rarely does media law or policy involve them directly.

From a European regulatory perspective, the notion of the ‘media-literate viewer’, and ‘media literacy’ more generally, has become the point of reference in the regulatory debate about the interests of end-users of various media, including search engines. The concept of the media-literate end-user incorporates a variety of perspectives on media use, including the citizen and consumer perspective in the relation between end-users and the media.⁶⁷² Media literacy is generally defined as “*the ability to access the media, to understand and to critically evaluate different aspects of the media and media contents and to create communications in a variety of contexts*.”⁶⁷³ While media literacy is being promoted in media law and policy, the rise of the regulatory notion of the media literate end-user is also a signal of a diminishing role of government in media regulation. Regulators no longer need to decide what is in the interest of the public, but can rely on more active end-users instead. Notably, the

⁶⁶⁹ See e.g. Herring 2007.

⁶⁷⁰ See also Section 9.2.3.

⁶⁷¹ Helberger 2008, p. 136.

⁶⁷² See Helberger 2008. See also Scammel 2000.

⁶⁷³ See European Commission 2007c.

European notion of media literacy is rather ambitious, including with regard to search engines. In the European Commission's view the concept specifically includes *"feeling comfortable with all existing media"*, *"being able to assess information, dealing with advertising on various media, using search engines intelligently"* and *"understanding the economy of media and the difference between pluralism and media ownership"*. Paradoxically, besides placing a focus on end-user choice, the notion thereby also signals a high level of government interest in the end-users actual use of new media.

Search engines are used by a wide variety of end-users with different degrees of skills and knowledge. Many professionals use the same search services for their research as laymen do. Journalists have become mass users of Web search services.⁶⁷⁴ The Internet in combination with effective search tools is a tremendous tool for self-education. It is also a tremendous source of confusion, annoyance and controversy. Obviously, there are groups of end-users that simply lack the education or the necessary skills to turn the abundance of information online into an advantage.⁶⁷⁵ What is important to recognize is that many of the problems that the least skilled may experience will not be easily resolved by placing obligations or restrictions on search engine providers. Many of the problems of Internet end-users simply exist at the more fundamental level of education and general knowledge skills⁶⁷⁶ or are a direct result of the uncontrolled nature of Web publishing.

And it must be admitted that almost everyone uses search engines without having a very good idea about their actual functioning. The complexity of Web search technology makes it impossible, even for the best informed users, to know what is really going on if they type in their search queries in different search services.⁶⁷⁷ And most will recognize that, when receiving a ranked set of search results, it is quite difficult to judge the relative quality of different search results without considerable background knowledge about the possible sources of information and ideas that the search engine could have referred to. It is easier to find information online than to assess its quality.⁶⁷⁸ What remains is a set of functioning hyperlinks, which can be used or discarded. Simply put: every search of the Web in a search engine like Google is a new experiment.⁶⁷⁹ While this may be said from the perspective of end-users it is similarly true from the perspective of search engines themselves. At Google, *"[e]ssentially all queries are involved in some test"*.⁶⁸⁰

8.6.3 The end-user: consumer or citizen?

The freedom of expression interests of end-users in the context of the governance of search engines involves both consumer as well as citizen aspects. It includes the typical civic engagement with media, in

⁶⁷⁴ See e.g. Machill & Beiler 2007; Machill & Beiler 2008.

⁶⁷⁵ See e.g. Kodagoda & Wong 2008.

⁶⁷⁶ Of particular interest in this regard is the work of Eszter Hargittai. See Hargittai 2004; Hargittai & Hinnant 2005; Hargittai 2007a; Hargittai 2007b.

⁶⁷⁷ On the way people see and understand Web search engines, see e.g. Hendry & Efthimiadis 2008.

⁶⁷⁸ Bowker and Leigh Star 2000, p. 7.

⁶⁷⁹ I owe this characterization to a talk by David Gugerli at the 2009 Society of the Query conference in Amsterdam. Gugerli has explored the way in which search engines turn our world into a database. See Gugerli 2009. On the need for critical engagement with Google's offering, see Lovink 2008.

⁶⁸⁰ See Levy 2011, p. 61.

particular with an eye on forming an opinion about matters of public concern. The heavy use of search engines during political campaigns may serve as an example from this context.⁶⁸¹ But it also includes the freedom to become an informed consumer. In free-market based societies, consumer freedom has become a fundamental aspect of informational autonomy. Obviously, the consumer and citizen aspects of media users do not always align. In the context of mass media it is generally accepted that commercial motivations can create biases with regard to the editorial content.⁶⁸² To prevent commercial motivations from dominating editorial decisions, newspapers tend to have a formalized separation between editorial and commercial governance of the organization. European audiovisual media regulation still contains a number of restrictions on advertising, sponsoring and product placement, such as the legal constraint that news broadcasts cannot be sponsored.

The tension between editorial and commercial motivations and the parallel citizen and consumer aspects of end-users is quite present in the context of Web search services and possibly unresolved. It may simply be hard for search engines to balance the competition for consumers in their services with their role in facilitating culture and knowledge. The advertising based business model of major search engine providers may lead to additional tension between these two different value chains, since it will incentivize search engines to focus on the consumer needs of their user.

Most early search services, and search engine portals in particular, used to put quite some emphasis on commercial interests. The user-friendly search engine that Google started offering in 1998, with its improved ranking algorithms, could be seen as placing emphasis again on the interests of end-users to find online information. In their well-known academic research paper about the PageRank algorithm, Larry Page and Sergey Brin, at that point still working at Stanford University, wrote a special appendix about the tension between end-user interests and commercial incentives, which was in their view inherent in an advertisement based business model:

“[W]e expect that advertising funded search engines will be inherently biased towards the advertisers and away from the needs of the consumers.

Since it is very difficult even for experts to evaluate search engines, search engine bias is particularly insidious. [...] [L]ess blatant bias [is] likely to be tolerated by the market. For example, a search engine could add a small factor to search results from “friendly” companies, and subtract a factor from results from competitors. [...] Furthermore, advertising income often provides an incentive to provide poor quality search results. [...] In general, it could be argued from the consumer point of view that the better the search engine is, the fewer advertisements will be needed for the consumer to find what they want. This of course erodes the advertising supported business model of the existing search engines. However, there will always be money from advertisers who want a customer to switch products, or have something that is genuinely

⁶⁸¹ See Oldham & Leach (Google) 2008. See also Zeller 2006.

⁶⁸² For an overview, see Baker 2002.

*new. But we believe the issue of advertising causes enough mixed incentives that it is crucial to have a competitive search engine that is transparent and in the academic realm.”*⁶⁸³

According to some sources, Google’s founders remained reluctant to introduce advertising on their service. Reportedly, they were in general quite negative about the value of marketing and considered industry practices around the year 2000 unethical and contrary to the interests of the end-user.⁶⁸⁴ When they did introduce targeted advertising and perfected it into the ‘AdWords’ programme, it made Google into a revolutionary business success.⁶⁸⁵

It is fair to state that it is more commercially compelling for search providers to satisfy end-users with advertisements than with unpaid for results. Google’s advertisement space has been expanded and advertisement prominence has been improved ever since. Moreover, Google seems to aim for the perfect ad: that is the ad that actually satisfies the end-user better than any other search result. Similarly, when Microsoft stopped the development of search products for academic materials in 2008, it pointed out it would focus on search with high consumer intent.⁶⁸⁶ The debate about search engine quality and advertising will be discussed in more detail in Chapter 10.

8.6.4 End-user Privacy

The ambiguity of commercial search providers towards the end-user’s interest in accessing information freely may be most striking when it comes to end-user privacy. All major Web search engines are offered for free, but they make their users ‘pay’ with unprecedented amounts of personal data.⁶⁸⁷ Search engine providers, like most other Web-based services for end-users to be fair, tend to log every single detail about the use of their services. The educated end-user that has an idea what is going on in terms of user data processing is still presented with a give or take with regard to their user data when it comes to searching the Web. If a person chooses to use any of the major search services without taking precautions to limit the possibility to track themselves, they have no access to or control over the data that is being collected, nor are they able to limit its subsequent use in significant ways. This data is usually stored in various unidentified locations, accessible to national security agencies, law enforcement agencies, and third parties in accordance with a variety of local laws. Google’s CEO at some point even defended the extensive collection of user data with the argument that it can help government agencies combat crime and terrorism.⁶⁸⁸

The unparalleled amounts of data that are being recorded by a handful of search companies and the special role of search engines play as the first points to access information online, lead privacy regulators across the world to express concerns about the data collection and processing practices of major search engines. In 2006 two international policy documents on privacy and search engines were

⁶⁸³ Brin & Page 1998.

⁶⁸⁴ For a discussion, see Edwards 2011, p. 68, 307-308. See also Levy 2011, p. 89, 145.

⁶⁸⁵ See Section 2.2.2 and Section 2.2.3.

⁶⁸⁶ See Microsoft 2008.

⁶⁸⁷ See e.g. Zimmer 2008

⁶⁸⁸ For a discussion, see Section 10.4.2.

adopted, containing statements and recommendations on end-user privacy.⁶⁸⁹ In 2008, the Article 29 Working Party issued an official opinion as to the application of European data protection law to the context of search engines. Data retention periods for individual search log data were of particular concern. The Article 29 Working Group also noted the impact of detailed processing of end-users at an individual level on the fundamental right to privacy as enshrined in Article 8 ECHR and the right to access information freely, in particular without surveillance by public authorities, as enshrined in Article 11 of the European Charter of Fundamental Rights. The Article 29 Working Party pointed out that detailed profiles were being stored and possibly used by third parties and governments:

*“An individual's search history contains a footprint of that person's interests, relations, and intentions. These data can be subsequently used both for commercial purposes and as a result of requests and fishing operations and/or data mining by law enforcement authorities or national security services.”*⁶⁹⁰

Arguably, the user's privacy is a precondition for the fundamental right to search, access and receive information and ideas freely. Information seeking behavior could be seriously chilled if the main available options to find information online entail comprehensive surveillance and storage of end-user behavior without appropriate guarantees in view of intellectual freedom.⁶⁹¹ Chapter 10 will address the protection of user privacy from the perspective of the right to freedom of expression in more detail.

8.7 Conclusion

The aim of this chapter has been to develop a general theory for the implications of the right to freedom of expression for the governance of search engines, a theory denoted by 'search engine freedom'. The analysis of Chapter 5, 6 and 7 clearly showed that freedom of expression implications were informed by the dominant normative conceptualization of the societal role being fulfilled by the press, the Internet access provider and the library respectively. Hence, one of the underlying aims of this chapter was to demonstrate the way in which such a role for search media in the public networked information environment could be conceptualized. By comparing search media's function in the networked information environment with the functions of the entities studied before, a number of conclusions could be drawn.

Search media combine a passive (conduit/access) and active (editorial/selective) role in their production of meta-information. The search engines' role is related to the *relative accessibility* of information and ideas online. They do not, like publishers, produce content themselves, and compared to traditional editorial media they are more passive. At the same time, however, search media are inherently more active in their mediation than a passive conduit, as for instance an Internet access provider. It is not the role of search engines to facilitate access to everything equally. After having established an index, search engines rank and select information and destinations online in return to user input in ways that, if

⁶⁸⁹ See 28th International Data Protection and Privacy Commissioners' Conference 2006.

⁶⁹⁰ Article 29 Data Protection Working Party 2008

⁶⁹¹ Compare the privacy discussion with regard to the Google book settlement, <http://michaelzimmer.org/2009/09/08/google-book-search-privacy-policy-mirrors-web-search/>

anything, resemble editorial media like the press, albeit in respect to a much larger variety of information.

When looking deeper into the societal role of search engines, two conflicting ideals emerged: the ideal of universal access on the one hand and the ideal of information quality on the other hand. The first ideal for search engines is to help end-users to navigate the entire Web, by crawling and ordering it and making the material in the index transparent and the underlying sources of online information accessible. The second ideal is to prioritize valuable information and ideas over lesser ones. The analysis showed how the general purpose search engine, by definition, has to reconcile these conflicting ideals in its operations and that much of the regulatory debate about the responsibility of search media could be explained with reference to the tension between these different ideals.

The tension between information quality and information access that exists in the public networked information environment did not exist in the same manner in the public information environment predating the Web, in which the organization of access to information for the public tended to be separated from the organization of basic levels of quality and legal permissibility. Access was restricted to everything ‘fit to print’ and their editorial freedom and responsibility was seen as an essential element of a functioning free press and democracy. Likewise, the public library first selects the sources it will subsequently make accessible and applies information quality criteria in this context. After having established a collection, a transparent accessibility infrastructure, informed by professional principles for the organization of knowledge, it would ensure access to the materials that had been selected. For access providers to such a relatively controlled information environment, most of the discussion about liability and filtering in Chapter 6 would have been unnecessary.

Web search engines help to establish the relative accessibility of information and ideas in the networked public information environment. This lead Dutch philosophers of science Marres and De Vries to make the claim that the societal legitimization of knowledge in the networked information environment takes place through processes of opening up, understood as the process of connecting information and ideas online to their societal use. If one follows this logic, the overarching public interest in the legal governance of Web search engines, from the ideals underlying the right to freedom of expression, lies in the establishment of a rich and robust societal infrastructure for the opening up of the Web.

This characterization of the public interest in the governance of accessibility of the Web is attractive precisely because it captures both perspectives, i.e. access *and* quality, which arguably lay at the core of search engine governance. This reasoning also clarifies that search engine providers have to make non-trivial choices with regard to the balance between quality and access. In the networked information environment publicity is no longer restricted to entities that offer *a priori* legitimacy to the information and ideas they make public. And the choices are not only non-trivial, they are of a political nature and involve the complex balancing of different public and private interests, including the interests of end-users and information providers that depend on search engines as well.

When looking at the specific implications of the right to freedom of expression on the governance of search, the analysis showed that under Article 10 ECHR the search engine should be able to claim

protection for its publication of references on its website as well as the process of crawling that makes it possible to offer a search engine in the first place. The protection of the search engines selection, meaning the selection of a set of relevant websites related to a specific query and ranking decisions under the right to freedom of expression may be one of the most interesting legal questions discussed in this chapter. In the United States there is some early case law which establishes First Amendment protection for selection and ranking decisions by search engines. In *SearchKing* a U.S. District Court applied the editorial freedom standards as developed by the Supreme Court in decisions such as *Miami Herald* and *Sullivan* to the freedom of search engines to decide freely how to rank and select references in response to user queries.

Considering the fact that operating a search engine logically implies a fundamental choice about the way to value online resources, there is much to be said for this part of the Court's conclusion. In other words, the choice of search engine operators how to select, rank, and present should be considered an editorial process, which deserves protection under the right to freedom of expression. The predominantly technological nature of the way in which these choices are expressed says less about the nature of this underlying process than about the massiveness of the index and the way in which technological innovation has offered new ways to organize and provide access to digital information collections. A proper understanding of the societal role of search media points in the same direction; by curating the relative accessibility of information in their index, Web search engines reconcile the ideal of universal access and navigation of the entire Web with the ideal of information quality.

Notably, accepting that a search engine providers' decisions how to select, rank and present would be protected by the right to freedom of expression does not imply, at least not in the European context, that such freedom would be unlimited or could not be restricted in view of other fundamental values. The conclusion that there may not be and should not be 'one correct way' to select and rank search results does not logically imply that there cannot be any legally impermissible ways to do so when offering search media online. First, one can imagine certain editorial choices by search engine operators that could be unlawful in and of themselves, such as the choice to implement algorithms that are specifically directed at causing harm or that cause harm while having no justifiable purpose. Second, the right to freedom of expression as enshrined in Article 10 ECHR is not absolute and may be restricted in the interests and freedoms of others. It is possible to imagine legitimate restrictions being imposed on search engine operators, dominant search engines in particular, that aim to ensure that the fundamental interests of information providers and end-users remain sufficiently respected. This question, which points to a possible positive role of the state to safeguard the right to freedom of expression in the governance of search, will be more thoroughly addressed in the next and final part of this thesis.

Of special importance in the European context is the question about the duties and responsibilities of search engines which are tied to the exercise of the right to freedom of expression. The duties and responsibilities under Article 10 ECHR are tied to the exercise of one's expressive liberties, and need to be interpreted in the present day conditions with regard to the new media environment, in which, in the Court's view, they have taken on an added importance. The potential impact of the medium and the nature of the content that can be found through a search engine will play a role in the determination of

its possible duties and responsibilities. In other words, it is likely that major general purpose search engines such as Google, which has a particular strong impact on the public information environment, could have enhanced duties and responsibilities based on their widespread use. In this context it is worth noting that the Court's case law seems to imply that the more a communicator does to abide by professional standards with regard to quality and the mode of communicating, the more it will be able to defend itself against interferences. Most search engines do fairly little to guarantee a threshold minimum quality of the information they provide references to. Generally, the lack of editorial control with regard to the actual content referred to and the lack of professional oversight over the inclusion of references could weaken a search engine's protection under Article 10 ECHR.

But there are other arguments in favor of less stringent duties and responsibilities as regards the quality of references in search engines. Arguably, duties and responsibilities should cut both ways. On the one hand, they can be argued to imply a professional responsibility on Web search providers to care about the quality of their references. On the other hand, they could be argued to entail a duty on search engine providers to be comprehensive and not to exclude references too lightly. In fact, an important normative principle in the Court's case law (*Open Door*) is that a communicator that leaves the decision to act upon its communications to the receiver cannot – in principle – be blamed for those decisions. The more facilitative the search engine would be with respect to the decisional autonomy of its users, the more protection it would receive against interferences which seek to prevent a certain reaction by the audience.

The protected interests of information providers under the right to freedom of expression can be best understood as the freedom to be present in the search engine's index and thereby to find their way to an audience. What's at stake for information providers in the context of dominant general purpose search engines could also be formulated in terms of representation. Hence, de-indexing by a dominant search engine is particularly problematic from the perspective of legitimate information providers. Legal issues related to the de-indexing of websites will be discussed in Chapter 9. The same may be said about an unfavorable treatment through selection and ranking decisions. However, it is impossible to argue that all information providers could have a legal claim to enter a dominant general purpose search engine index, as well as receive a favorable treatment by selection and ranking algorithms. Besides being impossible in practice, this claim would overlook a variety of legitimate grounds upon which a search engine could have for de-indexing or unfavorable ranking, including grounds directly related to the protection of search engines providers under the right to freedom of expression and the interests of the end-users.

With regard to search engine users, ultimately, their interest under the right to freedom of expression in the context of search may be best understood as a right to inform themselves freely by exploring the Web to its full potential, using available search technologies and services which enhance the findability of information, ideas and resources in the public networked information environment. The use of search engines has become a daily activity for end-users. End-users rely on search engines to find news and other resources, to inform themselves about products, political candidates, diseases, and to reach destinations and other online services. The user's freedom obviously implies a right to be able to choose which available navigational media to use. In addition, the user has a general interest in navigational

media of high quality. In Chapter 10 the question will be explored to what extent these interests of end-users could inform additional legal and regulatory involvement directed at the search engine market.