

Search Engines: Seek and Ye Shall Find? The Position of Search Engines in Law

by Nico van Eijk

EDITORIAL

In the electronic world, the ways of receiving and imparting information have changed beyond all recognition because of the ease in sharing information over the Internet. We can scarcely begin to conceive the gigantic amount of information put there at our disposal by private and public suppliers. It is even more impossible to monitor where all this information comes from or how it is being sorted and selected by those who make it available to us.

Luckily, we are not left alone to dig out the particular piece of news that interests us nor do we lack a system that lists items of potential interest – as library catalogues once were able to do.

Search Engines are the librarians of the Internet. They are the magic little helpers of the electronic information supply. Once the computer is switched on, using a search engine may even be faster than consulting any catalogue, encyclopedia or dictionary, even if they sat on the shelf next to us. Additionally, the use of the search engine is likely to yield many more results, simply because the storage place on the web outdoes many times that of any shelf.

What are search engines really? How, if at all, are they regulated? Why would they come under regulation? How do they compare to other technological means that we use in the process of asking for and receiving information electronically? These are questions tackled by this IRIS *plus* article, which I am sure you will enjoy reading.

Strasbourg, January 2006

Susanne Nikoltchev

IRIS Coordinator

Head of the Legal Information Department

European Audiovisual Observatory

IRIS *plus* is a supplement to IRIS, *Legal Observations of the European Audiovisual Observatory*, Issue 2006-02



OBSERVATOIRE EUROPÉEN DE L'AUDIOVISUEL
EUROPEAN AUDIOVISUAL OBSERVATORY
EUROPÄISCHE AUDIOVISUELLE INFORMATIONSTELLE

76 ALLEE DE LA ROBERTSAU • F-67000 STRASBOURG
TEL. +33 (0)3 88 14 44 00 • FAX +33 (0)3 88 14 44 19
<http://www.obs.coe.int>
e-mail: obs@obs.coe.int



Search Engines: Seek and Ye Shall Find? The Position of Search Engines in Law

Nico van Eijk¹

Institute for Information Law (iVIR), University of Amsterdam

Search engines play a central role in knowledge acquisition and knowledge transfer in today's information society. They are to a large extent responsible for making information on the Internet easily accessible. However, search results are often manipulated. The highest position on the list of search results is for sale and information providers use advanced methods to mislead search engines. Search engines hardly get a mention in law. For example, definitions that are used in media and telecommunications law exclude search engines from sector-specific regulations. New legal attention is needed to counter excesses and to make search engines more transparent. At the same time the provision and use of more independent search engines should be encouraged. Consumers should be made more aware of how search engines work.

Search engines have become an essential part of the way in which digital information is made easily accessible. They are used by virtually all Internet users, who moreover believe that searching using search engines is the best way of finding web sites.² "Googling" has become an autonomous concept and an independent form of leisure activity, similar to "zapping" through television channels. Anybody who cannot be found via Google does not exist: "To exist is to be indexed by a search engine."³

This article is an initial move to position the phenomenon of the search engine within the field of law, with an emphasis on sector-specific regulation such as media and telecommunications law.⁴ Before moving on to this legal analysis, in order to provide a better understanding of the problem, I shall look briefly at the way in which a search engine works, the manipulation of search results and the underlying business model. There will be a lot of references to Google. This is done many times as a metaphor or as *pars pro toto*. After all, in a lot of respects "Googling" is synonymous with searching on the Internet, and the position of other search engines is the same as, or similar to, that of Google.

How A Search Engine Works

A search engine's main function is that of enabling access; it is a *gateway* to possibly relevant information on the Internet. However, it is a two-directional gateway: from the information provider to the user and from the user to the information provider. A search engine determines what information provided by an information provider can be found by the end-user as well as what information the end-user will ultimately find. The search facility provided and the underlying search algorithm thus control supply and demand. Or to put it more simply: it is a bottleneck with two bottles attached to it.

How does a search engine work?⁵ Most search engines use more or less the same method to achieve search results. The

process starts with searching the Internet for information. This automated process uses intelligent "sleuths" called *spiders*, *bots* or *crawlers*. These sleuths surf the Internet using criteria set previously by the search-engine provider. The information found is thus made uniform and structured, which lays the basis for its traceability. Then the information is indexed. This indexing determines what counts as relevant words or combinations of words; and non-relevant information, such as fillers and punctuation marks, is deleted. At this stage the information is also streamlined in such a way that, for example, differences between singular and plural forms or variations that occur as a result of declensions produce identical search results. Certain recognisable words, such as people's names and basic concepts, are possibly identified. The rest of the information is then "weighted", based on the frequency of words in a text and the contextual relevance or significance. This enriched information forms the ultimate basic material for the search engine.

By no means all the information that is present on the Internet is found and indexed by search engines. According to Lawrence & Giles, individual search engines index only 16% of all the information present on the Internet and all the search engines together cover no more than 42% of all the available information.⁶ Other estimations contradict this low number, but this does not detract from the observation that only a limited amount of the information present is, or can be, indexed. There are various reasons for this. Some of the information is hidden in files that cannot be indexed, such as word-processor files or text in graphics files. However, search engines are becoming increasingly intelligent and are more and more capable of analysing a Word file or a PDF file. There is also information that the providers do not want to have included in search engines. News information that is rapidly refreshed, for example, is not suitable for inclusion in search engines as the information quickly becomes obsolete (sometimes months pass before a spider attempts to re-index the site). There is also information that is accessible via the Internet but that is not itself present on the Internet, such as information that is stored in external databases. Moreover, the Internet is still constantly growing and changing.

When a search engine is consulted, a process is largely used that is the opposite of the indexing process. The end-user formulates a search question that is broken down and analysed by the search engine. In this process, non-relevant elements (such as fillers) are again deleted, the relationships between the search terms are looked at (this can be indicated in the search query using Boolean operators (such as AND, OR, NOT)), and the relative importance of the search terms entered is charted. This results in several search results, which are displayed on the end-user's screen.

This model of collecting, ordering and making available information is only one reflection of reality. What actually



happens before a search result is made available is very complex and is characterised in an important way by the many subjective elements woven into the process. I shall group these subjective elements here under the term “manipulation”.

Manipulation of Search Results

The manipulation of search results takes at least three forms: manipulation by the search engine, manipulation by the information provider and manipulation by “hackers”.

Search Engines

The first form of manipulation is that carried out by search-engine providers. They draw up the criteria on the basis of which the information present on the Internet is collected, ordered and made available. Information that is not searched for is not found. If a spider is instructed to ignore certain information, this information will never appear as the result of a search action. The analysis of a search query and the answer to be given are determined by the algorithm that the search engine uses. This algorithm is the true secret to the way the process works and the ultimate manipulation tool. It resembles to some extent the secret recipe for Coca-Cola. Here are a few examples from practice to illustrate the manipulation by search engines:

Some search engines offer the opportunity of “buying” a high position on the list of search results.⁷ There are different variations of this. The simplest method involves literally selling the position. Other search engines priority-index the pages of paying parties so that they rank higher in the list of search results.

Some search engines deliberately do not reproduce certain results. For example, it is claimed that Google does not make certain search results available in the case of search queries from China.⁸ It would do this to protect its business in China and to prevent being put on a blacklist by the Chinese government. Part of Google’s defence is that these are sites that are inaccessible from China anyway. It is also claimed that Yahoo takes the wishes of the Chinese government into account.⁹ In addition, users in Germany are also supposedly not given the same information as other people when consulting Google. In particular Nazi propaganda is not available to users of Google.de.¹⁰ Research shows that the results of search requests differ, not only depending on the search engine used, but also depending on whether Google.com, Google.de or Google.fr is used.¹¹ Some of these cases concern information, the dissemination of which is not permitted in certain countries on the basis of legislation or jurisprudence, such as racist or Nazi information, or about the provision of goods that are prohibited in this context (the “Yahoo”-case).¹² References to this information are also often prohibited. A similar situation occurs regarding violations of intellectual property rights or unfair competition. Although these are interesting issues, the emphasis here is on filtering as a method used independently by a search engine to influence the accessibility of information. The fact that filtering is possibly done on the basis of

legal restrictions or a judicial decision is therefore not explored in any further detail.

To look at another example: Google, and other search engines as well, determine the search results in part using a system of *page-ranking*.¹³ Internet pages which are referred to by other pages are given a higher position on the list of search results. Moreover, user behaviour is monitored – Google does this, according to its own statements, in a fully anonymous way – from which more information is distilled that is used to determine the search result.¹⁴

Last but not least, there are search engines that, in addition to automated systems, also use a human factor: search results are manually adjusted by their own employees on the basis of more detailed criteria that have been formulated, both subjectively and otherwise.

Information Providers

The second form of manipulation is that done by information providers. They can do this by paying for a higher ranking in some cases or by exercising direct influence on the search-engine provider, but more often it is by cleverly designing their own web information to create a profile in such a way that the information is placed high up on the list of search results by the search engines. In doing this, they attempt to anticipate the search engine’s algorithm (to the extent that this is actually known). A classic example is the manipulation of one’s own *metatags*¹⁵ by adding attractive search words that have nothing to do with one’s own service provision (such as football, pornography or the brand names of competitors). However, search engines are becoming increasingly clever and are often capable of “neutralising” the effects of manipulated *metatagging*. Now, more advanced methods are therefore being used to attract greater attention. Fake sites are being set up, for example, that contain a lot of references to one’s own site in order to influence page-ranking systems. Or popular sites are being copied and included invisibly in one’s own site so that unsuspecting users end up somewhere other than at the site they intended to access. These and other forms of manipulation or deception are known as *spamdexing*, *cloaking*, *link-farming*, *webring*, *redirects*, *doorway pages*, *page-jacking*, etc.¹⁶ All such methods aim to improve the ranking in the search results. These manipulation techniques are combated by the search engines but by no means always successfully. At Google the ultimate sanction is the exclusion of the offender, whose pages are then no longer indexed.¹⁷ The party concerned can then no longer be found via the search engine. The offenders are not just shady characters, but also governments and reputable companies, which use agencies to optimise the search results. An entire industry has grown up around this optimisation of search results. Under the name *search engine marketing* companies offer services aimed at improving rankings. They are also called SEOs: *search engine optimisers*, a nice euphemism. Serious international conferences are held by the sector about issues such as “how to build links that generate traffic to your web site, and how to avoid the penalties of “spamming” the search engine”.¹⁸ Several handbooks have been written about this topic as well.

Hackers

A third form of manipulation can be regarded as a form of “hacking”. Often this is not about directly commercial interests but about deliberately causing damage, making a social statement or doing it just for the fun of it. One example is so-called “Google-bombing”. There follow a few classic examples to illustrate this sport. Some time ago, when the search term “who is more evil than the devil?” would be entered in Google as the search request, the first reference to appear was “www.microsoft.com”. The search request “miserable failure” still produces George Bush’s curriculum vitae as the first result.¹⁹ A Google bomb uses the same manipulation methods as those described above to achieve a higher ranking.

So the results of search requests are constantly under pressure from manipulation by search-engine providers, information providers and “recreational users”. These effects are further reinforced by the business model of the search engines.

Business Model

Search engines generate income mainly from one source: advertising. If we take Google as an example, this search engine generates 96% of its income from advertising.²⁰ This income is generated mainly by “Google AdWords”. AdWords enables advertisers to put their advertisements together themselves and state how much money they are willing to spend. They are then charged on the basis of the number of times that the advertisement is clicked on. The advertisements appear on the Google web site next to the results of a search request. Google decides which advertisement appears when and does this mainly in relation to the search request.

The second source of income consists of placing the advertisements on third parties’ web sites. This is done via the “AdSense” program, which has two variations: “AdSense for search” and “AdSense for content”. With “AdSense for search”, advertisements are placed in relation to search requests on third parties’ web sites. With “AdSense for content”, advertisements are linked to the content of web sites. For AdSense, Google has a revenue-sharing model, with some of the advertising income generated going to the information providers. These are thus in a position to take this into account when putting together the content of their web site and to “optimise” the content (which puts the traditional distinction between commercial interests and the editorial process under pressure: there can be a great temptation to adapt the editorial information in such a way as to generate greater advertising income via Google).

Google prides itself on the fact that in its business model there is no direct link between the search result and the advertisements shown. However, this is – to some extent if not to a great extent – just a gloss. After all, there is at least an indirect link. In order to be able to place relevant advertisements the results of search requests have to be linked to them. So it is therefore likely that the algorithm for the search result takes

this implicitly or explicitly into account or is influenced by this interaction.

This is also one of the arguments used by Introna & Nissenbaum²¹ when they discuss the question of whether search engines can be left to normal market forces. They state that in the case of search engines, the conditions for a normally functioning market are not present in sufficient number. To support this, they refer among other things to the fact that on the demand side, users are not well enough informed, either about possible alternatives or about the search result provided, which may be subject to the manipulation described above. Introna & Nissenbaum assume that search results are determined mainly by the highest common factor,²² not by what is actually relevant information for the person who submits the search request. As regards the supply side, they contend that there is simply no level playing field. Strong market players will be better able to influence the search results. Equally, there will be little interest in including in the search results web sites that are of interest only to a small group of users. They assume that the Pareto effect²³ also operates in the market for search engines and in such a way that a high percentage of the search requests result in only a small percentage of the information present. Introna & Nissenbaum consider this an undesirable situation as they emphasise “the value of comprehensive, thorough, and wide-ranging access to the Web”. They state: “We base our case against leaving it to the market on the particular function that we see search engines serving and on the substantive vision of the Web that we think search engines (and search-and-retrieval mechanisms more generally) ought to sustain.” It will therefore come as no surprise that Introna & Nissenbaum bring search engines into the concept of the “public good theory” (in ethical terms) and see this as a topic worthy of public attention. It cannot be the case that access to the web is hampered by search engines that systematically give preference to popular and commercial information. The Internet would in that case become a library containing books without any covers or tables of contents: the information would be present but it would be impossible to find it.

The Search Engine in Law

This plea for intervention prompts the question of how the law should deal with search engines. It seems natural to apply the classical analytical framework to search engines, whereby a distinction is drawn between infrastructure, transport services and information services.²⁴ Search engines are a typical example of convergence, more particularly of convergence between transport services and information services. After all, the search engine has – among other things because of the underlying algorithm and the advanced technology – aspects of a routing service, which puts these activities, in the classical school of thought, into the category of transport services. The search engine is in this respect comparable to the technical aspects of an electronic programme guide (EPG) and it also has certain characteristics in common with an application program interface (API). However, the search engine is more. Just as is the case with an electronic programme guide, the search engine also has functionalities that make it a marketing instrument and a content-related service. In her thesis Helberger

states quite rightly that these aspects cannot be regarded as transport but rather are characteristic of information services.²⁵

The core functionality of a search engine consists in making information sources easily accessible but in part in a lot of cases it also consists in providing access to the information itself.²⁶ Just as is the case with the electronic programme guide, the search engine thus breaks through the classical distinction made in media law and telecommunications law, for which areas the question of whether we are dealing with infrastructure or transport services or information services is a crucial one. Not only can this dividing line be found in various national laws; it is also characteristic of the European communications framework, which explicitly determines that it does not relate to the content. This means that the main information-service aspects of the search engine fall between two stools. I shall come back to this.

A Contiguous Right?

In my view, the search engine is thus mainly an information service. Can this service be further qualified? More particularly, how does the search engine relate to freedom of expression, the basic right when it comes to content? Is the freedom to receive and impart information central to search engines or does the search engine possibly come under a right of access to information? I believe that in the case of search engines, the freedom to hold and impart information, or the freedom to access information is not the main aim, as such, because the underlying information is already present on the Internet and is thus public/disclosed/accessible. Search engines *facilitate* access to information, but do not offer access by themselves. The immediate subject is not the information itself, but rather the *making accessible* of the information. The functioning of a search engine therefore entails activities that are of crucial importance for making the actual perusal of information possible. Points of law relating to search engines have implications for freedom of expression, including the right to receive and impart information and the right to access information as a “contiguous” right. Once again, *making information accessible is not the same as disclosing or disseminating it*. However, it is my opinion that making information accessible should be able to claim similar status. In the case of Article 10 of the European Convention on Human Rights (ECHR), this concerns a “contiguous” right, as regards both the freedom to receive and to impart information, comparable to the right to access information. The fact that this is a contiguous right does not mean that this right is a subordinate one. I agree with opinions in the literature about the status of contiguous rights, which reject a possible hierarchical relationship.²⁷ None of this detracts from the fact that neither Article 10, ECHR, nor Article 19 of the International Covenant on Civil and Political Rights, includes an independent right to making information accessible.

The “Television without Frontiers” Directive

But at a European level can the search engine be found in positive law such as set out in the “Television without Fron-

tiers” Directive (TWF)²⁸ or the new EU regulatory framework for the communications sector?²⁹ It is clear that the TWF Directive, including the recently proposed amendments to it,³⁰ has too little to offer for the regulation of search engines. The Directive concentrates on television broadcasting services and explicitly excludes “communication services providing items of information or other messages on individual demand such as telecopying, electronic data banks and other similar services”. The proposed revision, although modernising the Directive by including – amongst other things – Internet-related broadcasting activities, does not change the basic scope of the Directive.

The Regulatory Framework for the Communications Sector

Do search engines come within the scope of the new European communications framework? Is there perhaps an electronic communication service in question? According to the Framework Directive (Article 2c), an electronic communication service is “a service normally provided for remuneration which consists wholly or mainly in the conveyance of signals on electronic communications networks, including telecommunications services and transmission services in networks used for broadcasting, but exclude services providing, or exercising editorial control over, content transmitted using electronic communications networks and services”. Search engines are strongly related to content and in some cases there are editorial checks, as a result of which they fall outside the definition of electronic communication services. In addition, Recital 10 of the Preamble to the Framework Directive again states explicitly that the provision of Internet content falls outside the definition of electronic communication services.

However, the framework has introduced a new fact with the concept of “associated facilities” (Article 2e of the Framework Directive). This is understood to mean “those facilities associated with an electronic communications network and/or an electronic communications service which enable and/or support the provision of services via that network and/or service”. These also include systems of conditional access and electronic programme guides. Electronic programme guides are to a certain extent comparable to search engines. Does this mean that search engines can possibly come under the concept of “associated facilities”, for example, if my proposition that this is primarily an information service does not hold? The drafting history, the Preambles and the text of the Directives provide little justification for this. Articles 5 and 6 of the Access Directive state that only facilities that are related to the provision of digital radio and television broadcasting services are eligible for regulation. Furthermore, the concept “facilities” is used in the context of the provision of universal service, which has just as little relevance for search engines.³¹ In short, search engines do not fit into the new communications framework.³²

E-commerce Directive

Does the E-commerce Directive³³ potentially offer solace and is the search engine possibly an “information society



service”? An information society service means “any service, normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services”. We can assume that search engines come within the scope of this definition. However, the relevance of the E-commerce Directive to the problem of search engines is extremely limited.³⁴ The Directive does not, for example, supervise the way in which the algorithm used works and the way in which rankings are decided. However, Article 6 of the E-commerce Directive seems relevant, for it sets out the requirement that commercial communication constituting (part of) an information society service has to be clearly identifiable as such.

Competition Law

In view of the increased interest in competition law, I feel obliged to approach search engines from this angle too, even if it is only indicatively. Competition law as a generic law is definitely relevant. There is certainly a market for search engines, and classical questions arise about dominance and possible abuse of market power. It might go too far (yet) to state that Google can be compared to Microsoft, but a problem similar to that of the Microsoft cases exists. Google’s market share on separate national markets and on the European market may still be a little smaller than that of Microsoft on the market for operating systems but it still definitely holds a dominant position: in different national markets Google has a share of over 80%.³⁵ This does not make it easy to assume that effective market forces are at work, which is considered by some authors to be important for compensating for or mitigating some of the adverse effects resulting from the business model described above.³⁶

As regards the applicability of competition law, in particular from a European perspective, I would like to refer to the section on conditional access and competition law in the recently published thesis by Natali Helberger and in particular to the considerations about the electronic programme guide.³⁷ As I stated earlier: the electronic programme guide can in several respects be compared with a search engine. Excesses in the case of electronic programme guides, for example as regards access to the same, can be the subject of competition law. It is also imaginable that there is a role for competition law if certain information providers are systematically excluded from indexing by dominant search engines. However, an autonomous right to inclusion in an index derived from competition law is difficult to imagine. There is, for example, no *essential facility* at stake: there are alternatives available in the form of access to other search engines, and the setting up of a new search engine is in itself not a problem. Moreover, we saw above that the business model of a search engine is based on manipulation. Intervention based on competition law could therefore have too negative an effect on the business proposition of search engines. Also, intervention based on competition law raises questions about the legitimacy of this in relation to freedom of expression.³⁸ Nor is it easy to realise possible non-commercial aims within competition law. Some authors are more optimistic on this point than others.³⁹

Renewed Attention for Search Engines

Are there reasons to strive for a more active role by governments and legislators regarding search engines? As already stated, the ethics experts Introna & Nissenbaum, advocated a more active approach back in 2000. Authors in other disciplines, the economy, mathematics and information technology also devoted attention at that time to the phenomenon of the search engine.⁴⁰ However, search engines were then still at the beginning of their development and there seemed to be little reason to opt for a *merit/public good* approach for search engines. The importance of search engines was indeed referred to, but without leading to concrete consequences regarding policy or regulation.

A lot has happened since then. There are still many search engines, but there is a major dominance by just a few (or just one?). Alternative sites that are, for example, based on non-commercial objectives, are scarce; do not belong to those that are frequently visited, or concentrate on niche interests.⁴¹ Another relevant development is the fact that Google’s flotation on the stock exchange and its stock performance indisputably established the commercial objectives of search engines and the fact that the operating model is actually based on the generation of income via the forms of manipulation described above. Added to this is the increased pressure from the *search engine marketing* world.

Yet the most crucial development is the increased importance of search engines for making information easily accessible; information that is increasingly the basis for today’s knowledge acquisition. At issue is knowledge acquisition that is dependent on search engines, which are often outside national or even European jurisdiction, meaning that their (creation of) search results cannot be checked.

The development of search engines is therefore increasingly monitored with Argus’ eyes. The web site www.googlewatch.com is an example of such monitoring. In Germany, a non-profit association has been set up with the lovely name *Gemeinnütziger Verein zur Förderung der Suchmaschinen-Technologie und des freien Wissenszugang*⁴² (SuMa-eV). The association strives for greater awareness and for free, versatile, non-monopolistic search engines. Action points include the development of independent search engines and the disclosure of search-engine algorithms. In other ways too Germany plays a pioneering role in relation to search-engine problems. In 2003 a comprehensive study was published by the Bertelsmann Stiftung about search engines,⁴³ followed by a more recent study about search engines and the applicable German regulations.⁴⁴ The topic has already been expressly discussed in the German Parliament.⁴⁵ Political parties adopt standpoints⁴⁶ and under increasing social pressure, search-engine providers collaborate on the drawing-up of a code of conduct.⁴⁷ In other countries, on the other hand, it has long been *comme il faut* to state that the Internet should be left alone (or to overreact due to particular events).⁴⁸ That there is a certain change taking place can for example be seen from the fact that Germany and France have announced a joint initiative to produce a European search engine. This so-called “Quero Project” should result in a Euro-



pean search engine capable of competing with Google and other non-European search engines.⁴⁹ However, Quero does not seem to have taken into account the issues mentioned in this contribution.

Analysis and Conclusions

In today's information society, information takes centre stage as a source of further social and economic development. This is not only the case *de facto*; it is also expressed in many political declarations.⁵⁰ Increasing amounts of information are made available via the Internet, which gives the Internet a pivotal role in knowledge acquisition. Search engines increasingly determine the accessibility of all the available information and in this regard hold a position of power that cannot be ignored. Search engines occupy the place of qualitative access methods based on objective, verifiable criteria, such as those used *inter alia* by libraries and in the educational or academic world. The social relevance of search engines is therefore major and is only increasing as more and more information (including government information) is made available via the Internet.

The legal aspects of search engines do not feature widely in current research agendas. This also applies incidentally to making information easily accessible in a more general sense, where I have observed that there is often a misunderstanding between the question of access to information (the availability of information) and *making information easily accessible* (finding/tracing information). Search engines attract a great deal of attention and are (or have been) to a greater or lesser degree a subject of research in other scientific disciplines, but there are substantial gaps – such as concerning the legal aspects – and there is an utter lack of a more coherent multi-disciplinary approach. Partly in view of the major social dimension of search engines, it is desirable that legal and multi-disciplinary research be launched on national and international scales.

A possible legal framework cannot ignore the problem of convergence. The limited legal attention devoted to search engines is, I believe, partly the result of the fact that the search engine is neither one thing nor the other: it concerns

issues that are considered to fall within telecommunications law and partly – if not very much so – issues to do with content. Partly because of this, there is a legal vacuum: *the search engine does not have a place in law*. There are several ways of legally embedding search engines in the future.

I believe that there is a lower limit when it comes to making the procedures used by search engines more transparent. This could be linked to the existing regulations regarding information obligations, such as Article 6 of the E-commerce Directive or the new Directive concerning unfair commercial practices.⁵¹ We should be aware here that the most popular search engines are not of European origin, so there is a problem of jurisdiction.⁵²

More drastic legal intervention requires a careful consideration of interests. Trying to make search engines provide only “objective” search results is not realistic given that the operating model of search engines is determined precisely by manipulation. Excesses of this market failure should nevertheless be examined more closely and be considered for regulation. I am thinking of more specific rules in the field of consumer protection and of forms of access regulation for information providers.

Next to the question of possible regulatory intervention, a clearer view on the legal conceptual position of search engines, in particular with respect to the constitutional environment, is needed.

To achieve both the provision of and the use of more “neutral” search engines some form of governmental intervention – to be derived from a duty of care as yet to be fleshed out within the framework of the information society – will be unavoidable. This could lead to the support of initiatives that aim to provide independent search engines. However, it is also important that there be accompanying media education that not only promotes the use of alternative search engines but also supervises the use of existing search engines. Recent research in the USA indicates that a lot of users have a naive picture of search engines.⁵³ In view of the fact that these forms of government intervention are within the domain of information law and concern freedom of expression, caution is advised in outlining possible government policies.

1) Prof. Dr. N.A.N.M. van Eijk is a professor at the Institute for Information Law (IViR, University of Amsterdam, www.ivir.nl).

2) K. Deelstra, *Handboek Search enginemarketing, over search enginepositionering en adverteren*, Culemborg: Van Duuren Media, 2005, p. 5-7; L. Rainie & J. Shermak, “Search Engine use November 2005”, memo Pew Internet & American Life Project/Comscore Communications, 2005, http://www.pewinternet.org/pdfs/PIP_SearchData_1105.pdf

3) L. Introna & H. Nissenbaum, “Shaping the Web: Why the Politics of Search Engines Matters”, in: *The Information Society*, 1 July 2000, vol. 16, no. 3, p. 169-185, p. 171.

4) In this contribution, the focus is on the media- and telecommunications law perspective, alongside which some general issues relating to freedom of expression and competition law are dealt with as well. I recognise the fact that other fields of law such as intellectual property law and privacy law are also relevant.

5) E.D. Liddy, “How a Search Engine Works”, in: A.P. Mintz (ed.), *Web of deception: misinformation on the internet*, Medford: Cyber Age Books 2002, p. 197-208 (197-201); Deelstra 2005, *op. cit.*, p. 10-11.

6) S. Lawrence & C.L. Giles, “Accessibility of information on the web”, in: *Nature*, no. 400, 1999, p. 107-109.

7) See, *inter alia*, S. Nicholson, “How Much of It is Real? Analysis of Paid Placement in Web Search Engine Results”, in: *Journal of the American Society for Information Science and Technology*, 2005.

8) <http://battellemedia.com/archives/000919.php> ; <http://www.google-watch.org/china.html> ; http://www.theregisters.co.uk/2006/01/25/google_censors_chinese_results/

9) Refer to reports on Yahoo (and Google) on the web site of reporters sans frontières: www.rsrf.org. (*inter alia* “Google - Yahoo market battle threatens freedom of expression”, 26 July 2006.

- 10) Report "Bundestag sorgt sich um Vielfalt im Suchmaschinenmarkt", <http://www.heise.de/newsticker/meldung/print/48398>, <http://www.heise.de/newsticker/meldung/print/57734>
- 11) J. Zittrain & B. Edelman, "Documentation of internet Filtering Worldwide", in: C. Hardy & C. Möller (ed.), *Spreading the Word on the internet*, OSCE: Vienna 2003, p. 137-148 (138-139). See also: A. Mowshowitz & A. Kawaguchi, "Bias on the Web", in: *Communications of the ACM*, September 2002/Vol. 45, no. 9, p. 56-60 and B. Friedman & H. Nissenbaum, "Bias in Computer Systems", in: *ACM Transactions on Information Systems*, Vol. 14, No. 3, July 1996, p. 330-347.
- 12) J. Reidenberg, "Yahoo and democracy on the internet" in *Jurimetrics* 2002, p. 261-280; P. de Meij, "Grensoverschrijdende uitingen op the Internet", in: Van ontvanger naar zender, opstellen aangeboden aan prof. mr. J.M. de Meij, Amsterdam: Cramwinckel 2003, p. 217-241.
- 13) <http://www.google.com/technology/>. Page-ranking is named after Larry Page, one of the founders of Google.
- 14) Such as Google's e-mail service, Gmail: <http://gmail.google.com/gmail/help/privacy.html>
- 15) The information hidden in the web site in the form of key words that are read by the search engine for indexing purposes.
- 16) Deelstra, *op. cit.*, p. 169-184.
- 17) Google has a complaints procedure for allegations of misuse: <http://www.google.com/intl/nl/webmasters/spamreport.html>
- 18) See *inter alia*: www.jupiterevents.com/sew/winter06/index.html
- 19) Incidentally followed by the CVs of Jimmy Carter and Michael Moore.
- 20) Google, IPO prospectus, 2004, p. 38.
- 21) Introna & Nussenbaum, *op. cit.*
- 22) The *page-ranking* system described above is set up on the basis of this.
- 23) The so-called Pareto principle (also known as the 80-20 rule, the law of the vital few and the principle of factor sparsity) states that for many phenomena 80% of consequences stem from 20% of the causes (www.wikipedia.org).
- 24) N.A.N.M. van Eijk, L. Asscher, N. Helberger, J.J.C. Kabel, *De regulering van media in internationaal perspectief*, background study to: WRR, *Focus op functies, uitdagingen voor een toekomstig mediabeleid*, The Hague: WRR 2005, p. 66-67.
- 25) N. Helberger, *Controlling Access to Content - Regulating Conditional Access in Digital Broadcasting*, Kluwer Law International, Information Law Series: London/The Hague/Boston 2005, p. 10-19.
- 26) As the search result is the result of an indexing of the content of a book, text, file, etc.. In some cases, too, some of the content is part of the search result displayed.
- 27) J.M. de Meij, A.W. Hins, A.J. Nieuwenhuis & G.A.I. Schuijt, *Uitingsvrijheid, de vrije informatieboom in grondwettelijk perspectief*, Amsterdam: Cramwinckel 2000, pp. 114 and 129. I believe that an equal status should also be obvious in view of the great importance of search engines for freedom of expression, which in any case justifies an equal status rather than a subordinate one.
- 28) EC Council Directive 89/552/EEC on the co-ordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the pursuit of television broadcasting activities, adopted on 3 October 1989, OJ L 298, 17 October 1989, p.23, as amended by Directive 97/36/EC of the European Parliament and of the Council of 30 June 1997 amending Council Directive 89/552/EEC on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the pursuit of television broadcasting activities, adopted on 30 June 1997, OJ L 202, 30 July 1997, p. 60.
- 29) In particular, Directive 2002/21/EC of the European Parliament and of the Council of 7 March 2002 on a common regulatory framework for electronic communications networks and services (Framework Directive), OJ L 108/33 (24 April 2002) and Directive 2002/19/EC of the European Parliament and of the Council of 7 March 2002 on access to, and interconnection of, electronic communications networks and associated facilities (Access Directive), OJ L 108/7 (24 April 2002).
- 30) See: http://europa.eu.int/comm/avpolicy/regul/regul_en.htm
- 31) Article 10 of the Universal Service Directive and Appendix I thereto: Directive 2002/22/EC of the European Parliament and of the Council of 7 March 2002 on universal service and users' rights relating to electronic communications networks and services (Universal Service Directive), OJ L 108/51 (24 April 2002).
- 32) Which still leaves unanswered the question of how electronic programme guides come within the framework. It seems highly justifiable to me that on the basis of the same exclusions as described for search engines, it can be argued that the communications framework can only cover the electronic programme guide as a transport service and not the electronic programme guide as an information service.
- 33) Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (Directive on electronic commerce), OJ L 178/1 (17 July 2000).
- 34) See also: S. Feliu, "Intelligent Agents and Consumer Protection", in: *International Journal of Law and Information Technology*, Vol. 9 no 3, p. 235-248 (238-239).
- 35) See: <http://www.checkit.nl> (data at beginning of 2005) and <http://www.webhits.de/deutsch/index.shtml?webstats.html> (data at 12 May 2005) respectively.
- 36) N. Elkin-Koren, "Let the Crawlers Crawl: on virtual gatekeepers and the right to exclude indexing", *University of Dayton Law Review*, 26 U Dayton L Rev. 179, winter 2001; Introna & Nissenbaum, *op. cit.*, p. 58-59.
- 37) Helberger, *op. cit.*, p. 119-195.
- 38) Concerning the relationship between competition law and freedom of expression, see *inter alia*, Larouche & van der Haar, *Een schets van het Europese Mediabeleid*, Amsterdam:Pallas Publications 2004, p. 15-22; C.f.: Search King, Inc. v. Google Technology, Inc Case No. CIV-02-1457-M.
- 39) Incidentally, I strongly support the idea that competition law is only a means and not an end in itself, but in practice, matters are a lot more complex.
- 40) Elkin-Koren, at *op. cit.*, Lawrence & Giles, at *op. cit.* and Mowshowitz & Kawaguchi, at *op. cit.*, among others.
- 41) This is as yet – apart from the dominance by Google – my own empirical finding on the basis of the fact that the average student working on his or her thesis is guided increasingly by what is found via Google, does not use any alternative search engines and hardly ever visits a library.
- 42) Editor's note: this might be translated as "non-profit organisation in support of search engine technology and free access to knowledge".
- 43) M. Machil & C. Welp (ed.), *Wegweiser im Netz, Qualität und Nutzung von Suchmaschinen*, Berlin: Bertelsmann Stiftung 2003.
- 44) W. Schulz, T. Held & A. Laudien, Search Engines as Gatekeepers of Public Communication: Analysis of the German framework applicable to internet search engines including media law and anti trust law, in: *German Law Journal*, Vol. 06, No. 10, 2005, p. 1419-1433.
- 45) See further: <http://www.heise.de/newsticker/meldung/48398>
- 46) Bundesfraktion Bündnis 90/Die Grünen, *Suchmaschinen: Das Tor zum Netz*, Berlin 2005, www.gruene-fraktion.de
- 47) "Verhaltenssubkodex für Suchmaschinenanbieter der FSM": http://www.fsm.de/inhalt.doc/Verhaltenssubkodex_SuMa.pdf
- 48) See: M.D. Birnhack & N. Elkin-Koren, "The Invisible Handshake: The Reemergence of the State in the Digital Environment", in: *Virginia Journal of Law and Technology*, 6 (2003), p. 1-57.
- 49) Refer to: <http://www.gtfa-2005.com/> (under: "Traitement automatique des continus multimédia (project quaero)").
- 50) Such as the European "Lisbon strategy", in which the development of a high-quality information society is central: http://europa.eu.int/information_society/index_en.htm
- 51) Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No. 2006/2004 of the European Parliament and of the Council ("Unfair Commercial Practices Directive"), OJ L 149/22 (11 June 2005).
- 52) In the USA, comparable recommendations by the Federal Trade Commission (FTC) apply in order to make commercial interests transparent but compliance with these varies from search engine to search engine. Moreover, there are search engines that do everything possible to avoid making commercial relationships transparent: D. Fallows/Pew Internet & American Life Project, "Search Engine Users", Washington 2005, p. 17-18.
- 53) Fallows, *op. cit.*, p. 16-21.