Electronic Agents as Search Engines

Cruquenaire, Alexandre

Published in:
International Journal of Law and Information Technology

Publication date:
2001

Document Version
Publisher's PDF, also known as Version of record

Electronic Agents as Search Agents: Copyright related aspects

ALEXANDRE CRUQUENAIRE[1]

Abstract

In the Internet environment, search engine play an important role. Indeed, it would be utopian trying to find precise information on the web without this kind of services. These activities however raise copyright related questions. The mode of operation of search engines, which relies broadly on hyperlinks, could lead to copyright infringement, particularly concerning the moral prerogatives of the author. Furthermore, the practical difficulties to identify and sue copyright infringers bring the copyright owners to sue ISPs which indirectly contribute to illegal activities. Search engine could therefore be subject to indirect liability since they refer to illegal contents. Considering the importance of search engines, a specific legal regime should be designed, as in the US. However, nor the e-

[1] Researcher, Centre de Recherche Informatique et Droit (CRID), University of Namur.

commerce directive, neither the copyright in the information society directive address the question of the liability of search engines.

1 Introduction

An electronic agent acts with a certain degree of autonomy in relation to its user. Some electronic agents can be described as "intelligent" since they learn from experience in order to improve the performance of their tasks in compliance with their user's requests. Electronic agents are designed to accomplish tasks that human beings could have done.

Many of the electronic agents available on the web today are search engines or similar software[2]. In this context, this article will focus on the legal issues related to search engine activities.

There are three main types of search tools available on the web: (1) the classical search engines; (2) directories and portals, and (3) the specialised search engines.

The "classical" search engines are traditionally devoted to two major tasks: firstly, they gather information on the web and classify it following criteria generally determined by their user; secondly, they display lists of links to web sites related to an end-user request. These types of search tools use robots to gather the information on the web and to classify it. Considering the abundance of information, the human role is limited to the definition of the criterion used to classify the information gathered (generally the information is classified by keywords). The keywords relating to a specific content are collected in different ways[3]: fromthe words contained in the URL address of the concerned web page; from the words contained in the title and the first paragraph of the web page; fromall the words of the document concerned (scoring method); from the keywords contained in the meta-tags section of the HTML code of the page concerned. This latter method is particular since the keywords are defined by the author of the indexed content.

Some search tools work in a different way: they contain a database of links to web sites that have entered into an agreement with the service provider concerned. This means that the service provider has prior control on the content of the site towards which he/she links. The aim of the service is not to

---

[2] See e.g. *http://www.egghead.com*

[3] Cf. Verbiest, " Entre bonnes et mauvaises références. A propos des outils de recherche sur Internet ", online at *http://www.grolier.fr/cyberlexnet/COM/A990225.htm*.

provide the broadest information to the user but rather the information related to partner web sites that are of interest to the user. We will call these types of tools "directories".

Specialised search engines are dedicated to very specific topics (ticket selling, gathering of press articles, etc.). Generally, the information is only gathered on web sites chosen by the service provider. The area covered is more limited and the service provider controls the extent of this area.

One can make a distinction between two types of copyright related issues in the context of search engine activity: the possible copyright infringements committed by the search engine (discussed in section 2 below) and the copyright protection of the database created by the search engine (discussed in section 3).

2 Liability issues

The question of liability for copyright infringement is even more strongly linked to the "general" liability issues such that copyright owners use tort law more and more frequently to protect their copyright. Indeed, since it is sometimes very difficult to identify and sue the direct infringer, the copyright owners try to protect their rights by suing the online intermediaries (namely, the link providers and the hosting providers).

It is therefore important not to examine the possible copyright infringement raised by the linking activity solely from the copyright point of view, but also from the tort law point of view. We will first consider the questions related to direct copyright infringement (section 2.1) and afterwards the general tort liability questions (section 2.2).

2.1 How can a search engine infringe copyright?

Copyright covers different prerogatives which are, mainly, reproduction (in any form and on any medium), public performance and adaptation. The author has the exclusive right to authorise the exercise of these prerogatives. The author has also a moral right which allows him (at least)[4] to refuse any modification of his work which could be deemed prejudicial to his honour or reputation.

[4] See the minimal protection required by article 6bis of the Berne Convention for the protection of literary and artistic works (available on the WIPO web site, online at *http://www.wipo.org*).

Can the activity of collecting information on the web and creating links between this information and keywords be viewed as a copyright infringement? The fact of gathering information from the web does not raise any particular copyright problem – rather, it concerns the *sui generis* right[5].

The difficulties arise mainly from the creation of links between different contents. First the question of the legality of the linking practice is considered (section 2.1.1). Afterwards, the situation of the search tools is addressed (section 2.1.2) and finally the setting up of a specific regime for the search tools is considered (section 2.1.3).

2.2.1 The legality of hyperlinks with regard to copyright

The existence of hyperlinks is inherent to the HTML language. Indeed, HTML documents contain only plain text. To insert an image in a HTML document, a link to the image concerned is necessary. The link is a part of the HTML code of a page which identifies the location (internal or external[6]) of the image and instructs the Internet user's browser to pick up the referred element and to reproduce it.

A hyperlink is composed of two elements: the HTML code, which is invisible for the user, and the visible part which is generally made of words, pictures and/or logos.

For the purpose of this study, one should consider different types of links, depending on their mode of operation. One could make a distinction between invoke-to-load links and auto-load links[7]: the first type need an action from the user before displaying the targeted element (clicking on the visible part of the link) while the second is invisible for the user and runs automatically without any user intervention. One could also make a distinction depending on the method of linking: HREF link, inlining, framing. The HREF link redirects the browser of the user to another web site in a completely new navigation frame (and the sole URL displayed in the address frame of the browser user interface is the URL of the linked page)[8]. The inlining method consists in picking up an element and integrating it into the linking page, without any particular sign that allows the user to notice this process which therefore remains

---

[5] This is discussed in Grosse Ruse, "Electronic Agents and the Legal Protection of Non-creative Databases", also appearing in this Special Issue.
[6] according to the location of the image (inside or outside the server where the linking document is located).
[7] Cavazos & Miles, "Copyright on the WWW: linking and liability", *Richmond Journal of Law and Technology*, Volume IV, Issue 2, Winter 1997, online at http://www.richmond.edu/~jolt/v4i2/cavazos.html, pp.2-3.
[8] Strowel, "Liaisons dangereuses et bonnes relations sur l'Internet. A propos des hyperliens", (1998/4) *Auteurs & Média* p. 297.

invisible. Of course, the inlining technique uses auto-load links. Version 2.0 of Netscape Navigator has introduced a new technique – called "framing" – consisting in dividing the user's browser interface into several parts each of which are independent from the other. This technique allows one to display simultaneously contents from various sources while maintaining the "location" of the user at the linking site URL. One should also mention the deep linking technique consisting in bypassing the homepage of a web site and referring directly to specific pages contained in the site concerned.

The legality of the hyperlink is a very sensitive and discussed question. The linking technique has a decisive influence on the answer. Indeed, the different aspects of copyright could be differently affected by the abovementioned linking techniques.

Does a hyperlink imply a reproduction of the linked work? Generally, doctrine considers that the insertion of a hyperlink does not constitute a reproduction of the linked work[9]. Indeed, the hyperlink is only an instruction in HTML code which redirects a software (the browser of the Internet user) to another web site. The sole reproduction is made by the user's browser on the random access memory (RAM) of the user's computer. The insertion of an hyperlink only constitutes a reference to an external content which could be compared to a footnote in a book since the user needs to go to another web site to view the content[10]. However, one could consider that the insertion of a hyperlink could constitute a reproduction in the case of inlining. Indeed, this technique implies that the designer of the linking page gives instruction (in HTML language) to the user's browser to pick up an object on another site and to reproduce it *within* the linking page. The user has therefore the feeling that this external content is part of the linking page. In this case, it is submitted that there is a reproduction of the linked object in the linking page[11]. It is generally objected that the reproduction takes place only on the user's computer. However, the designer of the linking page planned this reproduction within his own page at a determined place by providing HTML instructions. One could make a comparison with the digitization process. It is not discussed that the digitization of a work constitutes a form of reproduction[12] (and adaptation, see infra) and the digitization consists in translating the work into binary language which is understandable by the computer. While reading the instructions, the computer understands that it has to reproduce the digitized work. While reading the HTML

---

[9] Cavazos & Miles, op. cit., p. 3; Carrière, "Hypertextes et hyperliens au regard du droit d'auteur: quelques éléments de réflexion", *Cahiers de propriété intellectuelle*, (1997) p.477; Visser, Dutch report in *Copyright in cyberspace*, ALAI Study days, 4-8 June 1996, (Otto Cramwinckel: Amsterdam 1997) p. 127.

[10] Carrière, op. cit., p.474; Lucas, *Droit d'auteur et numérique,* (Litec: Paris 1998) nr. 592.

[11] See for instance Strowel (op. cit., p. 302), who considers that there is at least a temporary reproduction of the linked work.

[12] Cf. e.g. Strowel & Trialle, *Le droit d'auteur, du logiciel au multimédia*, Cahier du CRID, n°11, (Bruylant: Bruxelles 1997), nr. 253; Sirinelli, French report in *Copyright in cyberspace*, ALAI Study days, 4-8 June 1996, (Otto Cramwinckel: Amsterdam 1997), p.108.

instructions, the computer also reproduces the digitized work. A hyperlink is also a form of machine code corresponding to the action of reproducing a designated work. The difference is probably in the fact that the linked work itself is not reproduced on the linking page, except when displayed on the computer of the user. However, one should make a distinction according to the type of link. Indeed, a form of reproduction exists in some circumstances. In the case of the HREF links, the HTML instructions given to the browser is to "leave" the linking site and to "go to" another web site (the linked site). In the case of inlining, the HTML instructions is to "reproduce" the linked content within the linking page when displayed to the user. It is held that, even if, strictly speaking, the HTML instruction does not contain the linked work itself, the inlining technique should be viewed as a form of temporary reproduction or public performance of the linked content. Moreover, it could also be argued that since the linking page is built by the insertion of various elements incorporated by the use of hyperlinks, it constitutes a 'virtual'[13] derivative work from the linked work, which implies an adaptation of the linked work[14].

The argument that a hyperlink could be considered a communication to the public (public performance) is not more sustainable – except concerning inlining (see supra) – for the above mentioned reasons[15].

The hyperlink technique also raises problems in the matter of moral rights. For instance, case law provides that the reproduction of press articles on a web site containing a general database of press articles infringes the moral rights of the author, since the journalist allows the reproduction of his/her work in a particular newspaper after considering the political and/or ideological trend of the concerned newspaper [16].

Even if the existence of a reproduction or communication to the public of the linked work in the linking page is rather uncertain, one could consider that the establishment of connections between different contents could lead to moral rights infringement in the case of connections that could be considered harmful to the honour or the reputation of the author.

---

[13] 'virtual' in the sense that the derivative work is only visible to the user who visits the web site where the linking page is located.
[14] Dusollier, "Les outils de référence: les cartes au trésor de l'Internet", in *Droit des technologies de l'information, regards prospectifs*, Cahier du CRID, n°16, (Bruylant: Bruxelles 1999), pp. 38-40.
[15] Contra, see Spacensky ("Promotion d'un site web et risques encourus: quelle responsabilité pour les outils de recherche et les créateurs de liens hypertexte?", online at *http://www.grolier.fr/cyberlexnet/COM/A980621.htm*), who considers that the hyperlink reference constitutes a communication to the public since the link provider takes the initiative to communicate the linked work to the Internet users.
[16] Cour d'appel de Bruxelles (cessation), 28 October 1997, *Auteurs & Média*, 1997, p. 383 (*Central Station* case).

The inlining technique increases the infringement risks. Indeed, the removal of a work from its original web site and its incorporation in another web page without the prior consent of the author could be considered as an infringement to his/her moral right to the respect of the integrity of his/her work. Concerning the minimal protection of the moral right guaranteed by the Berne Convention, the new environment within which the work is incorporated has to be harmful for the honour or reputation of the author.

Furthermore, in addition to the problems of the respect of the integrity of the work, a particular moral right problem could occur in the case of inlining if the name of the author of the linked work does not appear on the linking page[17]. This problem is more acute in the case of deep linking since sometimes the name of the author only appears on the homepage of the linked site.

Finally, the insertion of hyperlinks could be subject to liability for breach of copyright if the visible part of the hyperlink is composed in all or in part of elements protected by copyright[18].

2.1.2 The search engines case

Search engine activities present some particular risks with regard to copyright. These risks have to be assessed by considerating the type of activity of the service provider concerned (cf. supra): classical search engine (section 2.1.2.1), directory (section 2.1.2.2) or specialised search engine (section 2.1.2.3).

2.1.2.1 The classical search engine activity

Except in the case of inlining, the patrimonial prerogatives of the author are not very much involved (cf. supra on the hyperlink techniques). A search engine generally does not use inlining techniques. However, in October 1998, Altavista launched a new type of search engine called "AV Photo Finder" which displays to the user a mosaic of images corresponding to a keyword request. From the beginning of this new service, artists and particularly photographers reacted very negatively, invoking a violation of their copyright[19]. As explained before, in the case of inlining, it could be argued that a

---

[17] Spacensky, op. cit.
[18] Cf. Carrière, op. cit., pp. 477-478.
[19] See e.g. the Search Engine Report Newsletter (04/11/98): "AltaVista Photo Finder has artists concerned", online at *http://www.searchenginewatch.com/sereport/98/11-photofinder.html.*

(at least temporary) reproduction and/or a public performance occurs on the linking web site[20]. The central District Court of California ruled that this type of search engine activity could be considered fair use of the works concerned[21]. However, the transposition of this ruling into European law is quite uncertain.

Concerning moral rights, one has to be more qualified since search engine activities are more likely to infringe these prerogatives of the author. Indeed, in some circumstances, the author could invoke a violation of his moral rights where there is:

- a connection between his/her name or work and keywords that could be considered harmful to his/her honour and/or reputation;

- a connection between his/her name or work and other works that could be considered harmful to his/her honour and/or reputation (for instance: if the search engine displays an art picture of a nude in the middle of pornographic pictures[22]; if the references of an article against revisionism is displayed in the middle of racist materials) ;

- reproduction of a work within an environment different from its original environment (cases of inlining and deep linking). For instance, if the search engine reproduces an art picture of a nude in the middle of pornographic pictures (again).

These possible infringements to moral rights raise problems that are particularly difficult to manage since search engines do not have the ability or the resources to control all the web contents and connections resulting from the searches of their robots.

This situation is not too problematic concerning tort law since liability requires the demonstration of fault (negligence) that a normally attentive person would not have committed (see infra). However, copyright-based liability is particular: the sole demonstration of the unauthorised use of an author's prerogative (or the attempt to use such a prerogative) is sufficient to lead to the liability of the infringer, even if he/she has adopted a normally careful conduct.

Concerning the possible reproductions or public performances of the linked works, some technical elements have to be taken into account. Exclusion protocols allow creators to prevent others from referring to particular page(s) of their web sites:

---

[20] Verbiest, op. cit., p. 4.
[21] *Kelly et al. vs. Arriba Soft Corp.,*: Central District of California Court, case no. SA CV 99-560 GLT [JW], online at *http://legal.web.aol.com/decisions/dlip/kellyorder.pdf*
[22] Example quoted from Verbiest, op. cit., p. 4.

- robot.txt files : these files are placed at the root of the structure of a web site and are first read by the search engines' robots visiting the web site. It is then possible to put instructions in these files in order to prevent the automatic search robots to refer to certain pages. The weakness of this system is that the robot.txt file is directly connected to the domain name insofar that only the domain administrator can insert the required instructions. Therefore, it is sometimes practically difficult to use this exclusion mechanism when the concerned content is hosted under a sub-domain[23];

- meta-tags instructions: it is also possible to insert exclusion protocols in the meta-tags section of the HTML code of each web page. This exclusion protocol can be used for each web page whatever its location.

The non-use of these exclusion protocols raises a problem if the author of the referred content invokes later his copyright in order to sue the search engine provider. It is obvious that the existence of copyright cannot depend on the use of these mechanisms. However, one could wonder what the consequences of giving public access to a content without any restriction, are.

In the context of the creation of a direct hyperlink to a protected work hosted on a web site without any particular restriction from the author or host, reference is sometimes made to an implicit licence to hyperlink[24]. It is felt that this argument is derived from the bona fide principle and, more particularly, from the so called "théorie de l'apparence" – the appearance theory – which is a specific application of this general principle. Pursuant to this theory, third parties can invoke the behaviour of a rightowner if this behaviour has misled them. The consequence of this is that the appearance created, even without any fault, by the rightowner can be invoked as constitutive of a right in favour of the misled third parties. The application of this theory, as it has been established by Belgian case law[25], requires the fulfilment of three conditions: there must be an appearance which does not conform to the real situation; this appearance must result from the rightowner behaviour (even without any fault); this appearance has misled third parties.

The argument of the implicit license granted by the copyright owner for the creation of hyperlinks to his/her work is derived from similar considerations: the lack of restrictions by the copyright owner who puts his/her work on the web is constitutive of an implicit license since this behaviour is, without a doubt, likely to mislead potential linkers on their ability to link (or could reasonably be viewed by third parties as an implicit consent). However, it is better to address this problem from the angle of the

---

[23] Verbiest, op. cit., p. 2.
[24] Strowel, op. cit., p. 301; Carrière, op. cit., pp. 480-482 ; Julia-Barcelo, "On-line intermediary liability issues: comparing E.U. and U.S. legal frameworks", (2000) *EIPR*, pp. 116-117.
[25] See Kruithof, "La théorie de l'apparence dans une nouvelle phase", note on Cass. 20 June 1988, (1991) *R.C.J.B.* p. 68.

bona fide principle (théorie de l'apparence) than from the implicit license angle since copyright law generally does not recognise unwritten licenses. Copyright doctrine is generally reluctant to admit the interference of other law disciplines in its scope. However, it is submitted that bona fide is a general principle of law which should therefore apply also to copyright owners[26]. Copyright has certainly a particular role to play as a special matter of law but one cannot ignore all the general principles of law. It does not mean that, in general, every behaviour of the copyright owner could have as consequence a limitation of the copyright prerogatives. Particular behaviour of copyright owners should however be taken into account.

In the case of search engines, different elements could lead to an application of the "théorie de l'apparence", such as:

- the copyright owner decides to publicly disseminate his/her work on the web;

- taking account of the huge amount of information available on the web, the visiting rate of a web site depends highly on its reference in the search engine's databases consulted by Internet users;

- there exist technical means to exclude one or several pages from reference in that these database.

In this particular context, one could reasonably think that an author who does not use exclusion protocols gives the search engine providers the impression of wishing to be referred. Indeed, to be referred in search engines is a "natural" wish in the Internet environment and the fact of not using exclusion protocols could mislead search engine providers.

The situation is more difficult in the case where the author has stated on his/her web site that he/she does not want third parties to link his/her site. It is submitted that the provisions excluding hyperlinking cannot have been designed for search engines since the designer or builder of the web site knew that search engines usually act with searching robots and are not able to control the content of the referred sites. The "théorie de l'apparence" could therefore apply also in this context.

The fact that the "théorie de l'apparence" could allow reference to protected works available on web sites cannot howeverprevent the author from invoking his/her moral rights if the references made are prejudicial to his/her honour and/or reputation.

---

[26] Cf. Caron who considers that copyright cannot simply ignore the general principles of law in "France: les limites externes au droit d'auteur (abus de droit et droits du public)", in *The copyright boundaries*, ALAI Study Days, September 1998, (Oxford Press 1999), p. 237.

The problem raised by the "théorie de l'apparence" approach in the scope of copyright is that it could lead to a situation where the author would be obliged to use all the protection available in order to keep all his/her rights. One could not accept this solution. However, we think that the search engines case is particular in the Internet context considering the importance of their role in information localisation.

In order to preserve copyright integrity and the activity of search engines in the Internet context, one could wonder whether the European Commission should not be inspired by the American system which provides a specific regime balancing these two important and legitimate interests (see infra, section 2.1.3), though concerning only the indirect liability (see infra, section 2.2).

2.1.2.2 The directories' activities

The situation of directories is very different since the directory providers choose to refer to some site (contractual agreement). They are therefore able to control the content of the linked sites. As the inclusion in a directory results from a voluntary act of the copyright owner, it would be difficult to imagine action by a copyright owner against reference to his/her work. However, there could be problems if the database contains wrong information or makes connections harmful to the honour or reputation of the author (cf. supra). In this particular context, these problems have to be considered from the contractual point of view (what have the parties provided?).

2.1.2.3 The specialised search engines

The situation of specialised search engines is also particular since the human role is much more important in the choice of reference: human beings decide to which sites a link is to be made, and the choice of the categories of the database (a more precise criterion than in the 'classical' search engine case). This difference in the roles does not lead to any difference of legal solution in the matter of copyright since copyright-based liability does not require the demonstration of fault (see supra). However, this difference could lead to different solutions in the matter of tort law (see infra).

2.1.3 A specific regime concerning search engines?

Considering the important role played by search engines in the Internet context and the possible copyright infringement risks raised by this activity, one could wonder whether it would not be appropriate to enact a specific regime. This question is even more relevant since the identification of copyright infringements is quite impossible even in the case where human beings participate in the indexing process. The U.S. have provided a specific regime in their Digital Millenium Copyright Act (DMCA)[27] concerning indirect liability. This regime could inspire the European legislator in the area of direct liability.

The DMCA (§ 512, d) provides that the information location tool provider cannot be held liable for reference to web sites containing illegal material in certain conditions[28]. The service provider will not be held liable for referring to sites containing illegal materials or hosting illegal activities if the service provider:

(1) does not have actual knowledge that the material or activity is infringing; in the absence of such knowledge, is not aware of facts or circumstances from which infringing activity is apparent; upon obtaining such a knowledge or awareness, acts expeditiously to remove or disable access to the material;

(2) does not receive a financial benefit directly attributable to the infringing activity in a case in which the service provider has the right and ability to control such activity;

(3) on proper notification of claimed infringement, acts expeditiously to remove or disable access to the material;

The service provider has also to adopt a policy providing for the termination of subsciptions in the case of repeated infringements, and not to interfere with technical measures used by copyright owners to identify or protect copyrighted works (§512, i). The DMCA provides for specific procedures of "notice and take down" ($512, c, 3) and "counternotice and put back" (§512, g) which allow the service provider to assess clearly and easily the relevant character of a claim and to decide to take it (or not) into account without any risk of possible liability.

---

[27] Cf. Title II "Online copyright infringement liability limitation" (pp. 19-29), online at *http://lcweb.loc.gov/copyright/legislation/hr2281.pdf*

[28] See Julia-Barcelo, op. cit., p. 114.

However, this specific regime only applies concerning indirect liability issues (links towards illegal contents) and does not relate to the infringement risks resulting from the search engine activities themselves. Certain types of linking remain then subject to liability if they could be considered as direct infringement (see supra).

Considering the importance of the role of search engines in the Internet context, one should adopt a specific regime to balance the interests of the copyright holders and the interests of the search engine providers. Two solutions could be considered:

- the provision of a sort of compulsory license such as for reprography[29]: considering the fact that reprography is an economically and socially important activity, several European legislators have decided to provide a general legal license authorising reprography. In order to preserve the interests of the copyright holders, the Belgian legislator has provided a fair compensation. The money is collected upon the importation of photocopiers and upon the photocopiers' use according to the number of copies done. The amounts collected are shared between copyright holders. One could imagine a similar regime concerning search engine activity. However, there are at least four major problems in this context: first, the question of the determination of the concerned acts (the mere provision of hyperlinks?); secondly, the discussed character of the legality of the mere provision of hyperlinks; thirdly, the particular problem raised by moral rights: indeed, it seems difficult to admit a general limitation since the moral prerogatives are the most concerned; finally, it would not be very realistic to impose a fair compensation upon each hyperlink provision (if the provision of hyperlink is concerned) since it would lead to the death of the Internet system which is based on the hypertext architecture.

- The provision of a regime of exemption of liability for search engine activities. A regime inspired from the DMCA provisions concerning the connections established to illegal content could be an interesting solution. This could take into consideration the importance of the role of the search engine providers, the impossibility to monitor the content of the linked web sites (including the consideration of the possible linking restrictions provided by the copyright holders on the web pages)[30] and the copyright holders' interests. A notice and take down regime could be created in order to exclude any liability for copyright infringement resulting

---

[29] Concerning the Belgian approach, see Young & Roosen, "Le droit d'auteur et la reprographie enfin réconciliés", (1998) *Auteurs & Média*, pp. 93-105.

[30] Even in the case of specialised search engine, the information location tool provider cannot monitor the content of the linked web pages since it can change at each moment. Furthermore, the assessment of the legality of a content and/or of a connection established between different contents is sometimes impossible.

from hyperlinking: the link provider could not be held liable if, after receipt of a proper notice of claim from a copyright holder, he decides to remove the disputed hyperlink within a determined time period. One could also imagine a specific regime for the sole classical search engine activities and not for directories (in order to make a distinction between conscious linking and robot-made linking).

The European directive on the harmonisation of copyright in the information society does not address this question[31]. Recital 16 states that "liability for the activities in the network environment concerns not only copyright and related rights but also other areas, such as defamation, misleading advertising, or infringement of trade marks, and is addressed horizontally in 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), which clarifies and harmonises various legal issues relating to information societies services including electronic commerce (…)"

However, the European directive on e-commerce[32] does not address the liability issues related to search engine activities (a re-examination is foreseen on this topic – see article 21 of the Directive) either. The 50th recital of the e-commerce directive though insists on the fact that 'it is important that the proposed Directive on the harmonisation of certain aspects of copyright and related rights in the Information Society and this Directive come into force within a similar time scale with a view to establishing **a clear framework of rules relevant to the issue of liability of intermediaries** for copyright and related rights infringements at Community level'.

---

[31] Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, *O. J.*, L 167 , 22/06/2001, p. 0010 - 0019

[32] 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'), *O.J.*, L 178 , 17/07/2000 p. 0001 - 0016

2.2 Tort liability issues

As explained before, copyright owners frequently use tort law against online intermediaries in order to protect their rights against infringements committed by third parties.

Belgian case law has recently given two interesting examples of this new trend. Indeed, the International Federation of the Phonographic Industry (IFPI), has won one proceeding and lost another related to unlawful MP3 files hosting web sites on a tort basis.

In the first case, IFPI sued a hosting provider (Skynet) for having refused to delete from its server a web site containing hyperlinks to another web site containing downloadable unlawful MP3 files. The action was based on the fact that the refusal from Skynet constituted an act contrary to fair commercial practices [33]. The first court (Koophandelsrechtbank van Brussel) followed the opinion of the IFPI and condemned Skynet[34]. This decision has been reversed by the Brussels court of appeal[35]. This latter ruled that IFPI cannot require from an ISP to remove hyperlinks from its server unless IFPI does not comply with the following conditions, directly inspired from the e-commerce directive :

- IFPI has to notify the presence of alleged illegal contents on the ISP server ;
- IFPI has to expressly accept to take upon itself all possible liability resulting from an abusive removal of the contents concerned.

The court also reminds that one could not require from an ISP to assess the legality of a content, so that IFPI has to give the ISP the needed elements to establish the illegal character of the contents concerned. Since IFPI did not fulfil these requirements, the court reversed the first degree court ruling.

In another case involving the dissemination, by the way of an Internet web site, of a collection of hyperlinks (25,000) to unlawful MP3 files hosting web sites, the Antwerp civil court (Rechtbank van eerste aanleg) ruled, in a summary injunction[36], that the fact of making available for the public a collection of such hyperlinks constitutes an offence since this facilitates the access to the illegal (copyright infringing) contents. In this decision, the court ruled that setting up a hyperlink to illegal

---

[33] Article 93 of the act on fair trade practices and consumer protection (loi du 14 juillet 1991 sur les pratiques du commerce et sur l'information et la protection du consommateur, *M.B.*, 29 August 1991)
[34] Koophandelsrechtbank te Brussel, 2 November 1999, (1999/4) *Auteurs &Média*, p. 474.
[35] Hof van Beroep te Brussel, 13 February 2001, available at http://www.droit-technologie.org.
[36] Rechtbank van eerste aanleg van Antewerpen (kortgeding), 21 December 1999, ARK nr. 99/594/C, unpublished.

MP3 files hosting web sites can be compared to the fact of giving the user the key he needs to download illegal music. The behaviour of the linker makes an infringement possible and is therefore subject to liability[37]. Concerning the scope of the decision and its impact on further cases, one has to be very cautious since the factual background was very particular (the links provider had already been warned by the IFPI and his site was closed by a first hosting provider; however, he opened another site with the same content) and certainly had a very important impact on the final decision of the court.

Due to the practical difficulties to sue the people who illegally put reproductions of copyrighted works on the Internet, this trend will probably grow in the near future. This notably means that the links providers will be in the line of fire of the copyright owners.

The specific issue here is the existence (or not) of fault of the search engine provider. To address it, a distinction should be made according tothe type of search tool provided (see supra).

2.2.1 Classical search engines

The major problem is the control of the linked content. As the links result from information gathered by robots, one could not ask the search engine providers to monitor each linked content, otherwise, it would lead to the closing down of the search engine activities. Considering the amount of information gathered and indexed by the search robots, it would be difficult to demonstrate fault. The courts generally refuse to hold the intermediaries liable in circumstances where they were not duly warned by copyright owners[38]. If it is generally admitted that the online intermediaries cannot monitor the hosted or linked content, the courts are facing difficult problems when they have to assess the relevant character of a claim notification and/or of the response given by the concerned intermediary. This situation raises many problems for the intermediaries who receive a claim: should they comply with the complainant request or should they maintain the disputed content? The answer given by the intermediary could constitute a fault harmful for the copyright owner or an unjustified breach of agreement with regard to the author of the linked content (for instance, in the case of agreement relating to the reference in the search engine database)[39]. A better approach would be to adopt a specific regime inspired from the DMCA which seems very well balanced: the search engine providers cannot be held liable for providing the link except if they receive proper notification of copyright

---

[37] Based on common tort law (article 1382 of the Civil Code).
[38] Except in particular circumstances - see for instance the anonymous hosting in the *Estelle Halliday case*: Cour d'appel de Paris, 10 February 1999, online at *http://www.legalis.net/jnet/decisions/illicite_divers/ca_100299.htm*
[39] Cf. for instance the Skynet case concerning a hosting agreement, supra.

infringement and do not properly react in order to disable access to the alleged infringing content. The major advantage of a regime similar to the DMCA would be to set up a clear framework allowing each of the interested parties to assess the legality of the behaviour of the other party and to determine the appropriate answer to give (What is a proper notification of a claim? What should the search engine provider do to avoid any liability?).

2.2.2 The directories

At a first look, the situation seems rather different since the referred sites have been chosen by the link provider. However, the situation is more or less similar since the link provider cannot control the content of the linked sites which can change each second. The problems and solutions are therefore very similar: except in the case of negligence in the initial reference (lack of control in the case of manifestly infringing content), it seems difficult to demonstrate a fault of the service provider. However, the fact that the service provider is in contractual relations with the author of the alleged infringing content could impose on him broader obligations in term of collaboration with the complainant (to be able to communicate rapidly the contact references of the alleged infringer, for instance).

2.2.3 The specialised search engines

In this case, the service provider has no control on the linked content. The situation is a little bit different than the situation of the directories since there is no contractual relation between the linked content's author and the service provider.

2.3 Conclusion

The legal situation of the link providers and, in particular, of the search engine providers is rather unclear. This uncertainty raises many problems and needs to be removed as soon as possible in order to clearly balance the interests of the copyright owners and the interests of the search engine providers who play a very important role in the Internet architecture. In this context, the DMCA solutions should inspire the EC legislator.

3 The protection of the database generated by electronic agents

A database is, pursuant to the database directive[40], protected by copyright and by a *sui generis* right[41]. Considering search engines, there are practically no specific relevant copyright issues.

The directive provides that `databases which, by reason of the selection or arrangement of their contents, constitutes the author's own intellectual creation shall be protected as such by copyright' (article 3.1). The directive maintains the traditional criteria of originality. To be protected by copyright, a database has to be original (the expression of the personality of its author). In the case of databases generated by search engines, two questions arise: is the structure (selection or arrangement of the content) of the database original? Whose personality shows through this structure? The answer to the second question determines the copyright owner.

Concerning the originality of the database, very few of the search engines generated by databases will meet the originality requirements since they are as exhaustive as possible and/or classified following strictly logical criterion (price, length, etc).

As regards the few original databases, one has to wonder who has a decisive influence on their form. Since the current electronic agents available on the web are mostly search engines, it is obvious that there is no database structure resulting from the sole search robot activity. The structure of the database will therefore result from the service provider instructions which have determined the categories in which the data are classified and the type of search (the possible selection of the data).

The scope of copyright protection granted by the Directive is usual (articles 5 – restricted acts – and article 6 – exceptions): the author has the exclusive right to reproduce, adapt, translate, distribute and publicly perform. The sole mandatory exception regards the normal use needs (article 6.1.). A list of optional additional exceptions is provided in article 6.2: (a) reproduction for private purposes, (b) teaching illustration or scientific research, (c) public security or judicial procedure. Finally, article 6.2 provides an "open" clause for the traditional national exceptions.

---

[40] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, online at *http://europa.eu.int/eur-lex/en/lif/dat/1996/en_396L0009.html*

[41] See Grosse Ruse, "Electronic Agents and the Legal Protection of Non-creative Databases", also appearing in this Special Issue.