



**“Halt, who goes there?”  
On agents and conditional access to websites**

Martine L. Boonk

Computer/Law Institute, Vrije Universiteit Amsterdam, [m.boonk@rechten.vu.nl](mailto:m.boonk@rechten.vu.nl)

Arno R. Lodder

Centre for Electronic Dispute Resolution (CEDIRE.org), [lodder@cedire.org](mailto:lodder@cedire.org)

**Abstract**

*This paper deals with legal issues surrounding website access for software agents, notably the question how terms and conditions on a website can be presented in such a way that software agents and other automated means can adhere to them. We discuss the technology behind requesting website content, and indicate why website owners want to regulate the access to their sites anyway. The core of the paper is the analysis of the legal grounds for applying terms and conditions to visitors of websites: property and contract. We argue that under circumstances the mere visiting of a website may constitute a contract. In our discussion we take into account the theory of browse-wrap licenses and case law from both Common Law and Civil Law countries. A point that is given special attention is whether what holds for human users also holds for automated means, and we introduce ways to more effectively regulate website access for automated means.*

**1. Introduction**

Google is our best mate, but even with such powerful and convenient retrieval technology, finding the information one is looking for becomes more difficult every day. This is true for both academics and practitioners in the legal field. Most of us know where to find a particular Act, the text of a European Union Directive, or relevant case law. Retrieving all relevant (inter)national literature that appeared over the last couple of years in a specific legal domain is a harder and

more time consuming exercise. And how often is not much of your valuable time spent on finding relevant address information, or discovering who outside your core business is specialized in a particular topic?

It has been argued that given the enormous amount of information that is disseminated, we have no choice but to rely on others to sort out for us what is relevant and reliable (Rabinovich-Einy 2003). Others, however, cost money and their time is limited too. A better option is to let the technology do the job. Google does this quite well, but still requires the entering of queries by a human user. We believe that the amount of time it takes users to find relevant information could significantly decrease through the use of agent technology. For those who are not familiar with the concept, put simply, agents are pieces of software that perform tasks for their users, such as searching on the internet for relevant information, without the need for the human user to constantly monitor and instruct the agent. Basically, a general command or series of commands suffice. Comparable to a human agent acting on your behalf, which is also where it got its name from. For these computer agents to perform their tasks, they need access to websites. This paper deals with legal issues surrounding website access for software agents.

When it comes to the number of hits on a website, website owners have a somewhat paradoxical approach. At first sight, most of them want as many visitors as possible, but some guests are not really that welcome. As a matter of fact, website owners are keen on controlling access to protect their web site content against all kinds of unwanted access. In this respect, they often try and prevent spiders, crawlers, search agents and other automated browsing programs from accessing certain web pages or even the entire web site. Technological measures such as requiring a username and password can be helpful here, but they are not a panacea. After all, website owners do not want to discourage the right kind of users to easily access their web pages.

We will discuss whether website owners can successfully communicate terms and conditions concerning access to their website to both human users and automated means such as software agents. Since agent technology is currently being developed to offer users assistance in conquering the problem of over-abundance of information, this article focuses on website access

for software agents. Note though that this issue is not only important for the development of intelligent agent software, but also for search engines and other search tools.<sup>1</sup>

This paper is structured as follows. We first discuss the technology behind requesting website content, and why providers want to regulate access. The core of the paper is the analysis of the legal grounds for regulating behaviour on websites: property (section 4) and in particular a possible contractual basis (section 5) are discussed. These legal issues are addressed with a focus on both European legislation and the US Common Law doctrine. We end the paper with a proposal as to how more effectively regulate website access, and some concluding remarks.

## **2. Technology behind requesting website content**

In order to understand the legal analysis in this paper, we need first to explain what happens technically when a web page is requested by a user. First, a URL is entered into a web browser. A web browser can communicate with web servers through the use of the HTTP-protocol. It sends a HTTP-request to the web server of the website concerned. There are not less than eight methods to request data on a web server. A so-called “GET-request”, with which a copy of a particular web page can be requested, is the most common request.<sup>2</sup> The web server then returns the requested information to the computer on which the requesting program is located. Information on a website generally consists of documents written in HTML (Hyper Text Markup-Language). Additionally, a HTML document can refer to files that contain text, images, audio and video images. Such files are also passed on from the hosting web server to the requesting program.

When it comes to requesting web pages, automated browsing programs, including intelligent search agents, use the same HTTP-protocol as browsers do, sending get-requests for web pages themselves. For that reason, from a single request for a web page, a website owner is unable to distinguish between visits to her website by automated browsing programs and humans with browsers.

---

<sup>1</sup> For a survey of the current state of affairs of software agent technology in general, see the Agent Link III roadmap at <http://www.agentlink.org/roadmap/index.html>.

<sup>2</sup> For more information, see <http://www.w3.org/Protocols/rfc2616/rfc2616.html>.

In fact, the real difference between humans using a web browser and automated browsing programs lies in what happens to the requested information once it has been downloaded. Human users generally request a web page to view it on their computer screen, whereas automated browsing programs may process the downloaded data for other purposes, e.g. to compose a search index. More advanced search programs may process the data in a more sophisticated way, e.g. by searching for a specific phrase on a web page. Intelligent agents for example, tend to search for specific information, rather than crawling the web and harvesting every web page they come across (Boonk *et al.* 2005).

Content on a server can also be disclosed via other means than the WWW. When the content owner provides a different means for accessing content, this content can be viewed without a web browser. Google Web APIs service for example, is an Application Programming Interface (API) which enables users to directly send queries to Google's index of web pages in the same way as they would normally enter terms into a search form on the Google website, but without having to use a web browser.<sup>3</sup> This could be particularly useful when using a computer program to query Google's index of web pages.

From a technological point of view, every web page which is not protected by technological measures can be accessed by anyone. Many website owners who, for various reasons, may not want (certain types of) automated means accessing their website, do not (want to) take technological measures to protect the data on their websites. They can notify visitors by adding regulations to their website in either machine readable language or terms and conditions in natural language. For example, a web site owner can add a so-called robots.txt file to her website.<sup>4</sup> This file sums up the locations that website owners do not want automated browsing programs to visit. As a consequence a program "knows" it should not request those web pages.<sup>5</sup> Although the majority of all automated browsing programs are programmed to "read" a robots.txt file before requesting other pages of a web site, there is no legal obligation to adhere to them (Boonk *et al.* 2005). The same holds for metatags, which can be added to particular web pages, indicating automated browsing programs that they should not index that page or follow the links on that page (Subirana & Bain 2005, p. 132).

---

<sup>3</sup> See <http://www.google.nl/apis/> for a more detailed account.

<sup>4</sup> <http://www.robotstxt.org/wc/faq.html#what> ("What is a WWW robot").

<sup>5</sup> See for examples: <http://www.ariadne.ac.uk/issue15/robots/>.

In an effort to regulate access by automated browsing programs, website owners often add terms and conditions to a website that prohibit or restrict the use of (certain types of) automated means.<sup>6</sup> Alternatively, many terms of use oblige users to have their programs adhere to machine-readable exclusion clauses.<sup>7</sup>

### **3. Why providers want to regulate access**

Although web site owners may well have good reasons to regulate access for automated means on their websites, simply refusing all automated means does not resolve any problems. First of all, not all automated browsing programs can be lumped together. A search bot which has been designed for indexing purposes may simply harvest all data on a website, continuously following every available link on each page. An intelligent search agent on the contrary, is more likely to request one page; then process it and stop once it has found the required information.

Secondly, from a user's point of view, the prospect of websites that can be searched automatically is very welcome, as it saves time. These "quality differences" among search programs makes it virtually impossible to devise a universal solution for the problem of programs being unable to comprehend natural language.

There are various reasons for regulating website access for automated means. For instance, not all web pages are suitable for access by automated means, e.g. when they contain (dynamic) content, such as pages with forms, which usually cannot be processed by automated means. To add a web page which contains "special offers" that frequently change, to a search index is highly unpractical, because the information on that page changes so quickly that the copy stored in the search engine's index only during a short period corresponds with the actual information. Hence, a website owner may not want automated means to index such a web page either. Additionally, some automated browsing programs occasionally cause inconvenience. A website may be visited by bots that do too many get-requests in too little time, which may cause the web server to malfunction. Further, visiting bots can generate too much traffic as they simply request every

---

<sup>6</sup> Examples of terms that prohibit the use of automated means:

<http://www.findgift.com/Services/Terms-and-Conditions/> and, <http://info.lycos.com/legal/legal.asp>.

<sup>7</sup> Examples of terms that require adherence to robots.txt: <http://pages.ebay.co.uk/help/community/png-user.html> and <http://www.unibook.co.uk/terms.php>.

available web page and follow every available link on a web page instead of searching for specific information. Finally, many website owners want to protect the content on their website against copying of the content of the entire web site. A recent Dutch case, *NVM v. Zoekallehuizen.nl* deals with this issue.<sup>8</sup> NVM, an association of estate agents, has a website [www.funda.nl](http://www.funda.nl) on which the total supply of houses for sale of all its estate agents is published. Individual estate agents also publish their supply of houses for sale on their own website. In this case, *Zoekallehuizen.nl* is accused of copying unprotected data from all web pages of the individual estate agents to its own website, which offers a collection of the current supply of houses for sale as well, and deep-linking to protected data on the websites of the individual estate agents.

#### **4. Property as a legal basis**

On what legal basis a web site owner can successfully regulate website access for software agents and other automated means, is as yet unclear. At times, the web site owner may be able to appeal to copyright protection of the content. Website content can roughly be divided into two types, viz. content which is protected by intellectual property rights on the one hand and “unprotected” content on the other. Examples of the latter are pricelists, or web pages on which houses are advertised. Much of the information on the internet is not copyright protected (and not even protected by rights derived from the European Union Database Directive). In this paper, we will not discuss protection by intellectual property rights any further. Instead, we focus on the regulation of websites that mainly contain unprotected data. With regard to unprotected data, access regulation is even more important to a website owner. After all, in case a contract between parties is declared invalid, there is no intellectual property claim to fall back on.

Some authors argue that apart from intellectual property rights (notably copyright and database rights), it may also be possible to claim property rights of a different nature.<sup>9</sup> In several cases in the United States, the doctrine of trespass to chattels has been invoked successfully by web site

---

<sup>8</sup> At the moment of writing this article, the court has not yet ruled in this case. As of 16 March 2006, the verdict will be published at <http://www.rechtspraak.nl>.

<sup>9</sup> In Dutch law, the notion of a property right specifically refers to an all embracing right that can be exercised over physical objects. Anglo-American law does not have a universal notion of property. In Anglo-American law, the notion refers to a more general right to both tangible and intangible property.

owners trying to ban specific users from their websites without applying technological means. This doctrine assumes that the owner of a computer system can exercise certain rights over her system. Whether such a computer system can also comprise intangible goods, is subject to discussion (Bellia 2004, McGowan 2004, Burk 2000, Elkin-Koren 2001, O'Rourke 2001). In the eBay-Bidder's Edge case, the court argued that by continuing to request eBay's web pages while eBay had not only added a robots.txt to its pages, but had also explicitly stated that it did not want Bidder's Edge to crawl the eBay website and thus burdened eBay's systems, the conduct of Bidder's Edge constituted trespass to eBay's chattels.<sup>10</sup>

Under Dutch law, there have been two cases in which the plaintiff claimed that the defendant had violated the plaintiff's subjective right to (parts of) a computer system.<sup>11</sup> Both in the XS4ALL/Abfab case<sup>12</sup> and in the Netwise/NTS case a subjective right on (parts of a) computer system was claimed.<sup>13</sup> In the case between internet provider XS4ALL and spammer Abfab, the highest Dutch court of justice, Hoge Raad, ruled that the company Abfab was not free to send unsolicited email to XS4ALL clients, because XS4ALL had summoned Abfab to stop sending the emails and a continuation of sending unsolicited emails was regarded a breach of XS4ALL's exclusive rights to her computer system.<sup>14</sup> However, to what measure such a property right can be recognised under Dutch law and what it should encompass, remains questionable.

## **5. Contract as a legal basis**

The presence of Terms and Conditions on a web site implies an underlying contract. The question, however, is whether it is possible to conclude a contract under these circumstances. Although the question whether consent to an offer to form a contract has been manifested is a

---

<sup>10</sup> eBay, Inc. v. Bidder's Edge, Inc., 100 F. Supp. 2d 1058 (N.D. Cal. 2000).

<sup>11</sup> There has been at least a third Dutch case, Court of Justice The Hague 21 juli 2004 (nationalevacaturebank.nl vs. CVBank), where property on intangible goods was granted, but the reason we do not discuss it here that it has been generally considered a misjudgement.

<sup>12</sup> The Abfab/Xs4all case differs from the other cases in this paper, in that it was not about the request for information (pull), but about the unsolicited sending of data (push).

<sup>13</sup> Voorzieningenrechter Rotterdam 5 december 2002 (Netwise vs NTS) LJN AF2059, m.nt. T. de Graaf ([www.javisite.nl](http://www.javisite.nl)), m.nt. A.R. Lodder, Computerrecht 2003-2.

<sup>14</sup> HR 12 maart 2004 (Ab.Fab vs XS4all), LJN AN8483, m.nt. A.R. Lodder, vd Linden/Lodder, Annotaties internetrecht. Kluwer, 2006. See also [www.internetrechtspraak.nl](http://www.internetrechtspraak.nl).

fact-specific inquiry, some general considerations can be given to decide whether and how such a contract could come into being.

### **5.1 Does mere visiting a website constitute a contract?**

The first question in this respect is how the request for a web page is to be qualified. A related question is whether advertising a URL constitutes an offer and typing the URL an acceptance. In other words, can visiting a website be compared with opening the door to enter a shop?

As discussed above, a web site can only be viewed by a user when she first sends a request to the web server for a particular page. Upon the request, the web server will return information. Basically, this can either be the requested page (or the notification that the requested page cannot be found on that server), or the request for additional information, such as credentials required for viewing that page. Since all of this goes automatically, it is difficult to construct the corresponding consent of both the person responsible for the web server and the person requesting or the person behind the program which requests a website to make such an offer or accept the offer (Kerr 2004). Additionally, even if an offer and acceptance could be identified in such transactions, not all rights and duties are agreements under, e.g. Dutch law. For example, the agreement to attend a dinner party is no more than a rule of decency rather than a legal agreement. At least one of the parties must have an obligation to the other which is enforceable in a court of law (cf. Dutch law, Asser-Hartkamp 2004, p. 6-11). In our opinion, neither the request for a web page, nor the return of information can easily be regarded as output which can be enforced in a court of law.

The issue becomes clearer when we compare this situation with opening a door to enter a shop in real life. Is a contract concluded at the moment a door is opened to enter a shop? Under Dutch law, a visitor to a website cannot be said to have consented in anything legally enforceable yet. Opening a door in this case is more likely to be a physical action than an act that constitutes the conclusion of a contract.

When terms and conditions are presented in a shop, they will at least be part of the contract that is concluded when an item from the shop is being purchased. Nonetheless, it has been accepted that

that a simple exoneration clause presented clearly to customers in a shop or business accommodation is valid because customers can be considered to have taken notice of the clause.

Returning to the issue of website access, article 10 of the E-commerce Directive requires a service provider to provide information on “(...) (a) *the different technical steps to conclude the contract.*”<sup>15</sup> According to Lodder (2002, p. 83), the Directive provides for such information in order to prevent that people are contractually bound before knowing it. The Directive does not deal with contracts that regulate the service of offering information on a website. Although the Directive only discusses the conclusion of a contract for the deliverance of services and goods for which an order has been placed, it is questionable whether this would imply that a contract for the service of offering information on a website could be concluded by merely returning a web page upon request.

If knocking on the door of a web site is considered a mere fact, then when and how do agreements with respect to website access come into existence? This question is particularly important when discussing the applicability of terms and conditions regarding website use. Such terms can serve several purposes, for example regulating access to websites, behaviour on that web site, the use of the content of a website, privacy issues, etc. How can these terms become part of an agreement between the website owner and the website user?

## **5.2 Browse-wrap agreements**

When a website is not protected by technological measures, terms of use are generally presented either as click-wrap agreements or as browse-wrap agreements. Click-wrap agreements typically require a visitor to actively declare consent to terms before continuing to view the web site. Browse-wrap agreements are based on the assumption that one’s assent to such a contract is achieved by the mere act of browsing a web page.<sup>16</sup> In practice, many web pages can be viewed without first having read those terms of use.

---

<sup>15</sup> 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).

<sup>16</sup> See e.g. [http://www.zvulony.com/browse\\_wrap.html](http://www.zvulony.com/browse_wrap.html). Kunz *et al.* (2001).

In our opinion, the applicability of terms to a contract on a website depends on the nature of the contract that is to be concluded. To our knowledge, almost all website owners with terms and conditions on their websites regard their website as a service or a part of their service, which they only offer under their terms of use. In the EU context best-known as Information Society Services (cf. Directive 2000/31/EC). At least two types of websites can be identified:

1. Websites on which actual goods or services are sold, and;
2. Websites which only offer information, which can be considered (as) a service.

On both types of websites, the service of offering information can be offered under terms and conditions. With regard to the conditions for providing the service of offering information, browse-wrap agreements often appear. A possible agreement between parties is then meant to cover the service of providing a website with particular information to the user under the terms of use of the website.

When it comes to the purchase of goods or services though, users are more likely to expect additional terms and conditions governing their contract, so they can be expected to inquire the possibility of additional terms. Moreover, the moment a contract covering the actual purchase of goods or services has definitely come into being, can be stated more clearly, because it coincides with the moment of purchase. As outlined in article 9, 10 and 11 of the E-commerce Directive, the website owner must have fulfilled certain information requirements as well at this point. It is thus highly unlikely for such contracts to be concluded via a browse-wrap agreement.

Under Dutch law, for a tacit acceptance of an offer it is sufficient that a user has been notified of the other party's practice of applying terms at the moment of conclusion of the contract. Knowledge of the existence of terms that may be applicable to a contract at the moment of conclusion of the contract is required; otherwise terms cannot be part of the contract.<sup>17</sup> When pages can be viewed without even coming across a notification that terms are applicable, it would be a justifiable point of view to state that terms can certainly not apply until after a notification that specific terms are applicable.<sup>18</sup> The question remains whether mentioning the existence of

---

<sup>17</sup> In Dutch law, this is stated in article 6:217 BW.

<sup>18</sup> In this respect, knowledge of the existence of terms governing a contract must be distinguished from the knowledge of the content of those terms.

terms by adding a link to one or more web pages is sufficient for these terms to be enforceable. Apart from notifying the user that terms exist, the website owner must also give the visitor to a website a reasonable possibility to take notice of the contents of terms, under penalty of voidable terms.<sup>19</sup> It should be noted though that the average internet user is in general aware of the existence of terms of use regarding the browsing on a website. So, it could be argued that for a “take notice” the existing general practice on the internet suffices. For instance, we all know that cookies can be placed on our computers in order to facilitate the browsing of a site. The EU Directive 2002/58/EC on privacy in the sector telecommunication explicitly addresses this topic. Users may, of course, agree explicitly to each cookie being installed, but this means practically that browsing becomes almost impossible. Therefore users generally indicated in their browser, often without knowing, that they accept certain cookies being placed. Another example, if one copies a CD to the computer while being connected to the internet, Windows Media Player places additional information such as album covers, song titles, etc. This is also a standard configuration. In sum, both in practice (standard configurations) as well as in case of implied consent based on existing practice on the internet, many contracts are concluded with terms applicable without the users knowing it.

### **5.3 Case law**

In recent cases, some further guidelines have been given to determine how a web site user can be made aware of both the existence and the conditions to a contract. In the United States, there have been several cases with regard to the applicability of such terms.<sup>20</sup> A well-known case in this respect is the *Specht v Netscape* case, in which a U.S. district court ruled over the enforceability of a browse-wrap agreement that had been put on a Netscape’s website. The court ruled the terms to be unenforceable because the visitors on that website had not been given the possibility to agree to those terms.<sup>21</sup> From the *Register v Verio* case, it follows that a professional party who repeatedly visits a website and thus gathers knowledge of terms on that website, particularly when the party engages in activities that are usually restricted such as harvesting email addresses on a web site, has accepted the terms.<sup>22</sup> In this case, the question whether there was a browse-wrap agreement was left aside. In a more recent US case between *Hubbert and Dell*, the terms

---

<sup>19</sup> Article 6:233 BW.

<sup>20</sup> [http://gsulaw.gsu.edu/lawand/papers/su03/darden\\_thorpe/](http://gsulaw.gsu.edu/lawand/papers/su03/darden_thorpe/)

<sup>21</sup> See e.g. [http://www.cio.com/research/legal/edit/010402\\_agree.html?action=print](http://www.cio.com/research/legal/edit/010402_agree.html?action=print)

<sup>22</sup> *Register.com, Inc. v. Verio, Inc.*, 356 F.3d 393 (2d Cir. 2004).

and conditions containing the arbitration clause were declared part of the contract between the customer and the provider.<sup>23</sup> The Dutch case *Netwise/NTS* had an outcome similar to *Register/Verio*.<sup>24</sup>

The above cases may serve as a starting point for determining the circumstances and the moment that human users can be regarded to have consented to terms and conditions on a web site. Whether a user can be regarded to have consented to terms is largely a factual question. Besides the considerations regarding standard practice in the previous section, under different circumstances a user cannot have consented to terms until she has been informed at some time that terms exist by, e.g. a link on one or more of the web pages she has visited and access to the terms has been provided. It can be concluded that consent to terms on a web site can be assumed more easily depending on the number of pages of a website visited after the user has come across the notification of the existence of terms on a web site. Additionally, it can be stated that a visitor who does not agree with the terms on a website should stop visiting that site. If she does continue to request web pages, such behaviour may be explained as a tacit consent to the terms of use.

With regard to the case law discussed above, we do remark that although all of the above cases involved automated browsing programs requesting web pages, in none of the decisions the possibility was expressed that the user of such programs may not have noticed terms because her program queried the website or, more precisely, the web server. A reason for this could be that the browsing programs in question were built to access one or a few particular web sites, which were pre-programmed to some extent by their user. The fact that users are considered responsible for the behaviour of current browsing programs because they are regarded as instruments could be an underlying reason as well, even though all mentioned cases lack argumentation of that kind.

In the near future, in case agents would become more autonomous, it will be more difficult to relate an agent's actions to its human principal. With regard to autonomous software agents, which distinguish oneself from the other automated browsing programs by their autonomy, the degree to which one can be held responsible and liable for actions committed by such agents is already subject to discussion for quite some time (Lerouge 1999, Weitzenboeck 2001).

---

<sup>23</sup> *Hubbert v. Dell Corp.*, 2005 Ill. App. LEXIS 808 (Aug. 12, 2005):

<http://www.state.il.us/court/Opinions/AppellateCourt/2005/5thDistrict/August/Html/5030643.htm>

<sup>24</sup> Voorzieningenrechter Rechtbank Rotterdam, 5 december 2002 (*Netwise/NTS*), LJN: AF2059

Interestingly, a small number of websites, for example the Veron-website, explicitly hold users of automated means responsible for the behaviour of those programs on the web site.<sup>25</sup>

#### **5.4 From humans to agents**

It can be concluded that a user's tacit consent to terms of use on a website may be possible, depending on the actual circumstances. However, in our opinion, a user cannot tacitly consent to conditions that go beyond provisions which are necessary for a proper functioning of the service of offering information on that website. Examples of such provisions are access regulations, provisions with regard to a user's behaviour on a website (e.g. the prohibition of the use of bad language on a forum), and provisions about what may be done with the content (privacy policy, intellectual property notices). When consent to terms can be assumed, a contract between the website owner and a visitor to the website with regard to the service of offering information on that website may come into being.

The same holds when a web site is visited by agents or other automated means, even though in reality agents are still unable to process terms of use on a website. As automated means are still regarded as instruments rather than actors by some (For a discussion in Dutch on this topic, see Voulon 2004 and Lodder 2004), their human principles in any case responsible for each of their actions initiated on a web site.

From the above, it follows that regulating agent access through the use of terms and conditions is not effective due to the inability for agents to understand that there are terms and conditions on a web site. Unless website owners call the owners of agents and other automated means to account for not adhering to terms on a website, they fail to make clear through terms of use that some automated means are not welcome. The more autonomous agents are become, the less practical it will become to require agent users to read terms on a website first before sending their program to process the website. While natural language processing for computer programs such as agents is still in its infancy, it is necessary to look at other options. So how can agents and other automated means be made aware of access regulations on websites?

---

<sup>25</sup> <http://www.veron.nl/copyright.htm>

## **6. More effective regulation of website access**

Several suggestions can be made to regulate website access for agents more effectively, without having to take technological measures.

First of all, one can develop a web site which is completely inaccessible for agents. For example, websites which are made with the Macromedia Flash program require a Flash player to be viewed. There is usually some text on the home page and users are required to press “enter” to access the other pages. Web pages that consist only of pictures, as is often the case with Flash-pages, can only be viewed by human users, since pictures cannot not be processed by automated means. This method is a rather rigorous one, because it simply excludes all automated means.

Secondly, we suggest the development for agents (as more sophisticated automated browsing programs) a special agent.txt file, that allows to give specific access instructions to agents. It is even possible to devise special protocols reserved to agents, through which an agent can identify itself. That way, a web site owner could for example decide only to return web pages to agents with particular properties or whose users have explicitly consented in the web site’s terms and conditions. As far as autonomous agents are concerned, a further possibility concerns the development of protocols for online negotiation between agents and website owners (cf. Brazier *et al.* 2003). The conditions on which a particular agent is allowed to access a specific web site can thus be negotiated on an individual basis. Although agent negotiation may appear to be a promising solution, adherence to such measures would still happen on a voluntarily basis. When agents do not identify themselves as such, they can still access all web pages without knowledge of the web site owner. The success of such methods depends on the willingness of both parties. Another objection is that terms normally are not negotiable. During a meeting of the Dutch association of IT & Law on agents in 2000 Jac. Rinkes made this point in reply to the suggestion that agents could negotiate over contract terms. He told that when he negotiated to exclude a specific term applicable to the buying of a washing machine, he did manage to speak to the shop owner, but was ‘thrown out’ without a washing machine within five minutes anyway.

A third option to more effectively regulate access to websites for automated means could be to legally impose adherence by automated means to any machine readable no-robots clause (robots.txt, no-robots metatags, agent.txt etc.). Unfortunately, measures of that kind are unlikely

to resolve the issue of regulating agent access to websites. After all, without the application of technological measures, web site owners cannot see whether it is an agent that does an individual get-request. Moreover, such a solution disregards the reason for which robots.txt were originally devised, namely to resolve the nuisance caused by automated browsing programs visiting web pages that were unsuitable for them. Imposing adherence to such clauses fail to distinguish between the various reasons for banning automated means. It makes rather a difference when a website owner bans a program to prevent nuisance or to protect content. For that reason, any solution that does not take into account the various reasons website owners have when they inform users that they do not want automated means accessing their website, is insufficient.

More importantly, it should be noted that *how* a user views a particular web page on her own computer is usually not visible to website owners. From the website owner's viewpoint, both a computer program and a human user (who uses a computer program –browser- to request pages as well) request a web page in exactly the same way. Basically, their manner of requesting pages only differs with regard to the number of requests or the interval with which the requests are sent. In the light of the above explanation of how websites are requested, this means that website owners try and dictate what users do with the website content once it has been transferred to a user's computer. We think the power of a website owner cannot possibly go as far as to dictate what users do with unprotected data which had been sent to their computer upon request. It is highly questionable why automated means representing a person should not allowed access when the person they represent is free to request a web page.

## **7. Concluding remarks**

The internet is a strange phenomenon. It is important to be present, it is more important to be visited, and it is most important to be visited often and be ranked among the first ten hits of search engines. A necessary condition to be ranked high on a query result is that bots, automated programs, do visit your web page. It appears that some bots (e.g., from Google) are more welcome than others (e.g., from Bidder's edge). In general, providers of web site services do not always like bots, e.g. because they take resources of their servers or they index often changing information. The mind-boggling question here is: how to regulate website access to content that is in principle made available to all. We discussed what legal

grounds exists for regulation of website visits, with an emphasis of you might say more sympathetic automated means: intelligent agents.

We analysed briefly whether property of servers could form the legal basis. There is no final answer yet. The contractual basis appears to be more natural from a legal point of view. We distinguished two reasons why one might assume a contract has been concluded:

- on the basis of the standard configuration of browsers that allow the provider of a website to perform specific acts (such as storing cookies);
- on the basis of standard practice – we all know providers apply policies and give directions regarding website use. If we do not investigate the existence and content of such terms, we can in principle be bounded.

This all applies to human users, what about their automated counterparts? This issue is more difficult to solve. One problem is that agents or bots cannot read natural text, so they can only be bound to directions that are presented in a recognizable format. An option is to address their human users, but that is not a real option since there can be many users and addressing them all personally is quite difficult (if they can be traced at all).

Over time, however, the use of agents is expected to become more common. At that moment it becomes necessary for website owners to present their terms of use in an appropriate format (Lodder & Voulon 2002). After all, agents can generate access by human users or increase the dissemination of the content offered. And basically, that is what all website owners want.

## Acknowledgements

The first author has written this article within the framework of the Virtual Bridges project (see [http://www.iids.org/research/legal\\_aspects](http://www.iids.org/research/legal_aspects)).

## References

- Asser, C. and Hartkamp, A.S. (2004), 4: *Verbintenissenrecht, Dl. 1, De verbintenis in het algemeen*, Deventer 2004
- Bellia, P.L., (2004), Defending Cyberproperty, 79 *N.Y.U. L. Rev.* 2164 (2004).
- Boonk, M.L., Groot, D.R.A. de, Brazier, F.M.T. and Oskamp, A. (2005), Agent Exclusion on Websites, In: *Proceedings of The 4th Workshop on the Law of Electronic Agents (LEA 2005)*, pp. 13-20, <http://pubs.cli.vu/pub238.php>.
- Brazier, F.M.T., Oskamp, A., Schellekens, M.H.M. and Wijngaards, N.J.E. (2003), Are mobile agents outlawed processes?, In: Oskamp, A. and Weitzenboeck, E. (eds.), *Proceedings of LEA 2003: The Law and Electronic Agents*, pp. 127-139,
- Burk, D.L. (2000), The Trouble With Trespass, 4 *J. Small & Emerging Bus. Law* 27 (2000)
- Elkin-Koren (2001), N., Let the Crawlers Crawl: On Virtual Gatekeepers and the Right to Exclude Indexing, 49 *J. copyright Soc'y* 165 (2001)
- Kerr, I.R. (2004), Bots, Babes and the Californication of Commerce, *University of Ottawa Law and Technology Journal*, Volume 1, issue 1-2, 20004. <http://www.uoltj.ca/articles/vol1.1-2/2003-2004.1.1-2.uoltj.Kerr.285-324.pdf>.
- Kunz, Christina L. *et al* (2001), Click-Through Agreements: Strategies for Avoiding Disputes on Validity of Assent, 57 *Bus. Law.* 401 (2001).
- Lerouge, J.F. (1999), The use of electronic agents questioned under contractual law: suggested solutions on a European and American level, *John Marshall Journal of Computer and Information Law* 1999.
- Lodder, A.R. (2002), 'Directive 2000/31/EC', in Lodder, A.R. & Kaspersen, H.W.K. (eds.), *eDirectives: guide to European Union Law on E-commerce*, Kluwer Law International 2002.
- Lodder, A.R. (2004), 'Man and machine: what's the difference.' *JAVI*, 3(1):20-21, 2004, <http://www.javisite.nl/Artikel/0401-020AR.pdf>
- Lodder, A.R. & M.B. Voulon (2002), 'Intelligent agents and the information requirements of the Directives on distance selling and e-commerce', *International Review of Law, Computers & Technology*, Volume 16, no. 3, p. 277-287.
- McGowan, D. (2004), "The Trespass Trouble and the Metaphor Muddle" . *Minnesota Legal Studies Research Paper* No. 04-5. Available at SSRN: <http://ssrn.com/abstract=521982> or DOI: 10.2139/ssrn.521982
- McGowan, D. (forthcoming), "Website Access: The Case for Consent" . *Loyola-Chicago Law Journal*, Forthcoming <http://ssrn.com/abstract=420620!>;

**O'Rourke** (2001), M., Property Rights and Competition on the Internet: In Search of An Appropriate Analogy, 16 *Berkeley Tech. L. J.* 561 (2001)

**Rabinovich-Einy**, O. (2003), 'The Ford-Firestones of the Future: Resolving Offline Disputes in an Online Society', in: A.R. Lodder e.a. (eds.), *Proceedings of the ODRworkshop.org, Edinburgh 2003*, p. 37-50.

**Subirana**, B. & Bain, M., *Legal Programming: Designing Legally Compliant RFID and Software Agent Architectures for Retail Processes and Beyond*, Springer 2005

**Voulon**, M.B. (2004), 'Intelligent agent: instrument in plaats van actor. Het gebruik van de term intelligent agent in juridische beschouwingen omtrent het tot stand brengen van overeenkomsten', *JAVI* 3(1): p. 16-19, <http://www.javisite.nl/Artikel/0401-016AR.pdf>

**Weitzenboeck**, E.M. (2001), *Electronic agents and the formation of contracts*, Eclip-report EP 27028 2001.