“Halt, who goes there?”
On agents and conditional access to websites
Martine L. Boonk, Computer/Law Institute, Vrije Universiteit Amsterdam
Arno R. Lodder, Centre for Electronic Dispute Resolution (CEDIRE.org)

Abstract
Google is our best mate, but even with such powerful and convenient retrieval technology, to get out the right information becomes tougher every day. The amount of time it takes users to find relevant information could significantly decrease through the use of agent technology, if, of course, designed properly. This paper is not about software design, on the contrary, we analyse legal issues surrounding website access for software agents.

A paradox exists with regard to the number of hits. Providers want as many visitors as possible, but some guests are not really that welcome. Website owners are keen on controlling access to protect their web site content against all kinds of unwanted access. In this respect, they generally try and prevent spiders, search bots and other automated browsing programs from accessing certain web pages or even the entire web site. Technical measures can be helpful here, but from a legal point of view the question is how a website owner should regulate the internet traffic, how he can stop agents and the like from accessing information that in principle is free to all.

One of the ways to do this, is by presenting General Terms and Conditions on a web site that prohibit automated browsing programs from accessing (parts of) that web site. Just like their human counterparts, agents probably never read those terms, if they are capable of understanding them at all. Nonetheless, pointing agents or their human users to those terms might help in banning the unwanted guests, in particular if terms oblige a user to have his spiders, bots etc. adhere to machine-readable exclusion clauses, such as a robots.txt-file or "no-robots" meta tags.¹

On which legal basis a web site owner can successfully ban such computer programs from accessing his site, is as yet unclear. At times, the web site owner may be able to appeal to copyright protection. But many of the information on the internet is not copyright protected (and not even protected by rights derived from the European Union Database Directive). In American literature, it has been suggested that a party has an exclusive right to a computer system on which a web site is hosted. An intensive discussion is going on as to whether unwanted access to websites may under circumstances be regarded as the aged Common Law tort trespass to chattels. To our knowledge, this issue has not been discussed as elaborately in European courts yet. In Europe, there have mainly been cases concerning unsolicited email. And although the ownership of the servers may give providers the right to restrict the use of their (virtual) property, so far these concern the permission to send unsolicited packets, rather than the status of requesting web pages without prior consent.²

¹ See e.g. http://store.naturalimpact.com/terms.html
Whether terms and conditions or torts are used as the legal ground, the central question is how to provide the other party, an agent in our case, with the necessary information regarding his rights and duties concerning website access. Related to that question is under which circumstances a website owner has done enough to hold the other party liable for the damages caused by unwanted access. Finally, since General Terms and Conditions indicate the existence of a contract between parties, the question arises when such a contract comes into being. Can the domain name be considered as an offer to visit, and the get-request (e.g., in case of human users typing the URL) as the acceptance so that each visit to a website is based on an information society service contract? All these questions are addressed in our paper, with a focus on both European legislation and the US Common Law doctrine.