

CRAWLING THE NETWORK: INFOBOTS

One aspect of the appeal of data mining lies in the joy of discovery, but is data mining the precursor to information warfare? If information is in a public space, can a website owner protect it from others' access or use? As the quantity of available information increases, automated tools to find information will become part of the network structure. Intelligent software agents will scour the web and report back. Can automated agents or "bots" collect anything they can turn up? Is the activity similar to a scavenger hunt for shells on the beach or is it snooping at the people under the next umbrella? In I, Robot Isaac Asimov proposed three ethical principles for robots:

1. A robot may not injure a human being, or, through inaction, allow a human to come to harm.
2. A robot must obey orders given to him by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

The legality of bots has made its way to the courts and is causing a controversy. In eBay, Inc. v. Bidder's Edge, Inc. (100 F. Supp. 2d. 1058 (Dist. N.Cal. 2000)) the tort of cybertrespass poses the theory that an uninvited bot is a parasite on the system capability of a website and the potential harm of a possible crash is grounds for an injunction against the bot and the entity controlling the bot. Other than engaging the server, no actual harm was shown. The robot query used up approximately 1% system resources by hitting the web site approximately 100,000 times a day in order to update the Bidder's Edge auction aggregation site.

In order to defend its systems from uninvited robots eBay used "robot exclusion headers," which prohibited the use of any robot, spider, other automatic device, or manual process to monitor or copy the web pages. The robot exclusion header is a message sent out to other computers programmed to detect and respond to headers and to notify the sender that eBay does not permit unauthorized robotic activity. A programmer can design a robot to read a data file (robot.txt) to understand the system message and respond to the "not welcome" command. eBay can identify robotic activity by monitoring the number of incoming requests from particular IP addresses. The robotic activity is not much different from a person using a search engine and the

search engine scouring the net to find the responsive link. The eBay system can identify multiple hits from the same IP address, an occurrence which may indicate robotic activity.

Bidder's Edge attempted to defeat the monitoring by using rotating proxy servers and disguising the origin of the hit. The number of hits did not in any way reach the level of a denial of service attack. The Bidder's Edge intent was to update its own auction site with available prices and merchandise from several auction sites. This type of data aggregation is comparative pricing at its best. Bidder's Edge users went to eBay to make an eBay purchase. eBay tried to negotiate with Bidders' Edge so that the robot would only be sent in response to a user inquiry to an Ebay item from the Bidder's Edge site and Bidder's Edge could then update its site in response to the query. Bidder's Edge refused and said it wanted to send constant robots so that its database was always up to date and it could immediately update its users to changes. eBay objected because it had a contract approach to dealing with other "auction aggregators." eBay charges those types of users a pro rata portion of costs of maintaining the system when the bot uses the website in this style for this purpose.

The court applied the standard for a preliminary injunction, requiring eBay to show the likelihood of success on the merits and the possibility of irreparable harm. eBay filed nine causes of action which indicated that the real gravity of the complaint was Bidder's Edge's use of the eBay information. The court agreed with the eBay claim that the automated agents resulted in lost capacity to its computer system. The court accepted the argument that if Bidder's Edge activity was left unchecked it would encourage other auction aggregators to engage in the same activity, which could have the cumulative effect of harming the eBay's systems. The court reasoned that a property owner had the right to exclude others and that the bot activity was a trespass to chattels.

Commentators have taken issue with this novel legal theory because the eBay website is publicly available. However, the servers are private property and the owner has the right to control the level of access it grants to the public. The court reasoned that a trespass claim exists when the defendant exceeds the scope of the consent even though the use is not a conversion. This type of robot activity is referred to as possessory intermeddling; courts do not agree on what amount of intermeddling necessary to establish the tort of cybertrespass to chattels.

In applying the Asimov's robot ethics to the eBay case, one could find that 1) no harm occurred to any humans; 2) the robot did not disobey orders from human beings; and 3) the robot's existence is subject to law.