SYRACUSE JOURNAL OF SCIENCE & TECHNOLOGY LAW

VOLUME 30 SPRING 2014 ARTICLE 2, PAGE 56

NOTICE AND MANIFESTATION OF ASSENT TO BROWSE-WRAP AGREEMENTS IN THE AGE OF EVOLVING CRAWLERS, BOTS, SPIDERS AND SCRAPERS: HOW COURTS ARE TETHERED TO THEIR APPLICATION OF *REGISTER* AND *CAIRO* AND WHY CONGRESS SHOULD MANDATE USE OF THE ROBOTS EXCLUSION STANDARD TO PREVENT CIRCUMVENTION OF RESPONSIBILITY

Michael H. Laven

TABLE OF CONTENTS

I. INTRODUCTION TO BROWSE-WRAP AGREEMENTS

A. What are Browse-Wrap Agreements?

In 2012, when Internet users browsed the World Wide Web looking for the best price on a new Apple product, Thanksgiving flight or car insurance, they inevitably encountered a brave new world of manifestation of assent to a contract: the world of "click-through" and "browse-wrap" agreements.[1] The click-through agreement probably garners the most awareness from the average Internet user, as satisfactory completion usually involves clicking "agree" or "yes" before the one is allowed to continue – a physical action from the user that is mandatory.[2] However, much more commonplace, as at appears on virtually every website, although much less conspicuous, is the browse-wrap agreement. This type of agreement is found on websites of all varieties, including commercial, educational and personal websites, and allows for acceptance of the website's "terms of use"[3] simply through the conduct of continued use of the website. The user therefore has notice of the terms, may read them if they desire and may discontinue their use of the website if dissatisfied with the terms offered. Bits and pieces of litigation have arisen involving both click-through and browse-wrap agreements, certain issues have been settled, but, with technology evolving so quickly, the current state of the law leaves many uncertainties for web users and designers alike.[4]

These types of agreements are really just the natural evolutionary extensions of judicial decisions of the recent past, which dealt with the validity of "pay-now-terms-later" contracts ("PNTL"). A classic example of a PNTL contract comes from the realm of shrink-wrapped software. If one were to recall physically purchasing software from a brick-and-mortar retailer in the 1990's, one might remember that the software was shrink-wrapped in plastic. Once the

---

[1] *See* Ian Rambarran & Robert Hunt, *Are Browse-Wrap Agreements All They Are Wrapped Up to Be?*, 9 TUL. J. TECH. & INTELL. PROP. 173 (2007) (contending that the two most common types of standardized electronic agreements, encountered by Internet users, are click-through and browse-wrap agreements).

[2] *Id.* at 174, 177-78. Rambarran & Hunt further explain: "A click-through agreement is usually conspicuously presented to an offeree and requires that person to click on an acceptance icon, which evidences a manifestation of assent to be bound to the terms of a contract."

[3] Essentially, the "terms of use" of a website are contractual terms offered by the website designer, and are found commonly after the user clicks a hyperlink at the bottom of the website's homepage. Once the user clicks the hyperlink, which might display the text "terms of use" or "terms of service," the user is directed to another webpage where the actual terms are displayed for the user to read.

[4] *Id.* at 174-75.

purchaser got home, unwrapped the product and inserted the software for installation onto their computer, the purchaser was presented with the End User License Agreement ("EULA"), which the purchaser was required to agree to in order to use the software. These agreements, and others with similar attributes, came to be known as PNTL because the purchaser's notice to the terms of use came after the exchange of money for goods.[5] From a policy rationale standpoint, it would not be conducive to retail shopping if a cashier was required to read all terms of use to a customer at the point of sale, and, because of this, courts have been willing to enforce the terms of PNTL agreements if the purchaser has the opportunity to reject the terms and return the product.[6]

A landmark PNTL decision came about in *ProCD v. Zeidenberg*[7], essentially a click-through case in the early to pre-internet era. Zeidenberg, the defendant, purchased the plaintiff's software and was then presented with the terms of use later, when he installed the software on his computer. The court held that the defendant, through his continued use of the software after presentation of the terms, amply manifested acceptance, particularly because the defendant had the option of discontinuing use of the software and returning it.[8] This decision, relying heavily on the Uniform Commercial Code[9], opened the door for click-through and browse-wrap future validity, stating "[a] vendor, as master of the offer, may invite acceptance by conduct, and may propose limitations on the kind of conduct that constitutes acceptance."[10]

B. The Notice Requirement

---

[5] *See* Florencia Marotta-Wurgler, *Some Realities of Online Contracting*, 19 SUP. CT. ECON. REV. 11, 12 (2011), (describing PNTL's as those in which "sellers do not make their contracts available to consumers until *after* they purchase the product.").

[6] *See* Hill v. Gateway 2000, Inc., 105 F.3d 1147, 1149 (7th Cir. 1997) (where the customers did not have notice of the full contractual terms until after the computer, purchased over the phone from a Gateway representative, arrived, stating: "[p]ractical considerations support allowing vendors to enclose the full legal terms with their products. Cashiers cannot be expected to read legal documents to customers before ringing up sales.").

[7] ProCD v. Zeidenberg, 86 F.3d 1447 (7th Cir. 1996) (holding that because the defendant used the "shrinkwrapped" software, after having a chance to reject the terms of use and return the product, the acceptance was binding).

[8] *Zeidenberg* is not so different from the web user confronted with a browse-wrap terms of use – that user can simply leave the page if not satisfied with the terms.

[9] Zeidenberg, 86 F.3d 1447; *see also* U.C.C.§ 2-204 ("A contract for sale of goods may be made in any manner sufficient to show agreement, including conduct by both parties which recognizes the existence of such a contract.").

[10] *See* Zeidenberg, 86 F.3d at 1452.

Usually, when a party is contemplating a physical world offer, the opportunity to review the terms before possibility of assent is nearly automatic, unlike browse-wrap agreements, where the conspicuous display of such offer to the party might be in question. Stated simply, contracts require that the offeree have notice of the offer being presented, and this aspect, coupled with formation of assent, present unique complications in the context of the web.[11] Turn again to the web user confronted with a browse-wrap agreement – the statement of terms, once found, is the offer, and the notice requirement is likely satisfied if the hyperlink leading to the textual terms can easily be found by the user. A look at Google's search engine homepage reveals a modern, common example: close to the lower right corner of the page is hyperlinked text that states "Privacy & Terms".[12] Clicking that link leads the user to a page has three prominent links, one of which is "Terms of Service"[13], and, when clicked, the user is led to a page that finally displays the actual text of the terms of service, accompanied, in the second paragraph, by the phrase, "By using our Services, you are agreeing to these terms."[14] This manner of browse-wrap agreement and the steps required by the user to find and read the terms are commonplace on the Internet. In the context of reasonable notice on the web, Google has an uncluttered home page, relative to many of its competitors, displaying the "Privacy & Terms" hyperlink fairly conspicuously.[15]

C. Browse-Wraps, the Law, and People

Over the last twelve years courts have attempted to tackle, with varying degrees of certainty, the validity of agreements created by the conduct of a webpage visitor. A proper chronology of cases would probably start with *Pollstar v. Gigmania, Ltd.*[16], where the term

[11] *See* Woodrow Hartzog, *Website Design As Contract*, 60 AM. U. L. REV. 1635, 1643 (2011) (arguing that when courts interpret contracts they "focus on whether the plaintiff had reasonable notice of and manifested assent to the online agreement.").

[12] https://www.google.com/intl/en/policies/?fg=1(last visited Nov. 4th, 2012 15:00 EST).

[13] The other two prominent hyperlinks are "Overview" and "Privacy".

[14] https://www.google.com/intl/en/policies/terms/ (last visited Nov. 4th, 2012 15:00 EST).

[15] For purposes of example, the author encourages you to visit http://www.yahoo.com/ and attempt to find the hyperlink for the terms of use; then, try the same thing on http://www.google.com and compare your experience. This analogy was relevant on November 4th, 2012 15:00 EST.

[16] Pollstar v. Gigmania, Ltd., 170 F. Supp. 2d 974, 982 (E.D. Cal. 2000) (possibly coining the browse-wrap term, stating "[t]he court hesitates to declare the invalidity and unenforceability of the browse wrap license agreement at

"browser wrap" or "browse wrap" may have very well made its debut in a judicial decision about the validity of the browser wrap agreement that was decided, factually, on whether the term's display was "sufficiently conspicuous". A series of *Ticketmaster* cases focused the spotlight on whether there was clear evidence that the defendant had assented to the terms and conditions of the plaintiff's website. In those cases, the court distinguished between the visibility of shrink-wrap terms and browse-wrap terms.[17] *Ticketmaster* came on the heels of *Specht v. Netscape Communications Corp.*[18], where the court, grappling with a browse-wrap agreement that accompanied the download of a software "plug-in"[19], dug deeply into the issue of visibility and notice, noting the subtle difference when a user must scroll down the page to successfully find the terms or a hyperlink that directs a user to the terms.[20] Although the courts were dealing with issues of contract law within a new manner of commerce – on the Internet – these cases, and a few subsequent which will be discussed shortly in this article, support the concept that the fundamental principles of contract have not been changed by these particular technological developments.[21]

C. Browse-Wraps and Agency in the Age of Bots

---

this time . . . people sometimes enter into a contract by using a service without first seeing the terms—the browser wrap license agreement may be arguably valid and enforceable.").

[17] Ticketmaster Corp. v. Tickets.Com, Inc., CV997654HLHVBKX, 2003 WL 21406289 (C.D. Cal. Mar. 7, 2003) (finding that "[t]he "shrinkwrap" cases find the printed conditions plainly wrapped around the cassette or CD enforceable. Even the back of your parking lot ticket may be enforceable.").

[18] Specht v. Netscape Commc'ns Corp., 306 F.3d 17, (2d Cir. 2002).

[19] A plug-in, generally, is a smaller software program that operates within, and enhances the function of, an existing software program. *Plug-in*, TECHTERMS.COM (2014), http://www.techterms.com/definition/plugin.

[20] Specht, 306 F.3d 17 (commenting on the "scroll-down" visibility issue, stating, "even though plaintiffs could not have learned of the existence of those terms unless . . . had scrolled down the webpage to a screen located below the download button . . . a reasonably prudent Internet user in circumstances such as these would not have known or learned of the existence of the license terms . . . and that defendants therefore did not provide reasonable notice of the license terms.").

[21] *See* Cairo, Inc. v. Crossmedia Services, Inc., No. C 04-04825 JW. 2005 WL 756610 (N.D. Cal. Apr. 1, 2005) (stating, "[w]hile new commerce on the Internet has exposed courts to many new situations, it has not fundamentally changed the principles of contract. It is standard contract doctrine that when a benefit is offered subject to stated conditions, and the offeree makes a decision to take the benefit with knowledge of the terms of the offer, the taking constitutes acceptance of the terms, which accordingly become binding on the offeree.").

As busy as folks are in the modern world, forget browsing the web yourself. You can simply dispatch your web robots ("bots") to do your bidding and have them return to you with data and information, the likes of which ye have never seen.[22] This article has already attempted to dissect browse-wrap agreements with regards to people, but humans are not the only visitors to websites on the Internet. They go by many names, but bots, spiders, crawlers, scrapers or whatever they might come to be called are constantly scouring the Internet, performing a variety of tasks for a variety of people and entities.[23] [24] These bots can perform many important tasks, such as helping a web user perform various "info-chores", but bots, which will most likely continue to play an important role in the future of the Internet, can be utilized for nefarious purposes as well.[25] In the year 2012, the average Internet user encounters bots more than some might realize. If you have ever unsuccessfully bid on an EBay auction item, chances are, you did not lose a legitimate fight - you lost to a bot.[26]

However, many e-commerce business models include bots.[27] The most prominent example of when bots (or data scrapers in this instance) are essential to a business model is with a search engine. All of the major search engines utilize bots to repeatedly scrape the Internet for websites and the content of those websites, and this data is returned to the search engine, analyzed and used to generate search query results for the engine's users. Even issues of copyright, in this context, have mostly been settled as fair use, because of the economic

---

[22] Read this sentence using a pirate voice/accent.

[23] *See* Andrew Leonard, *Bots Are Hot!* WIRED MAGAZINE, ISSUE 4.04, APR 1996, http://www.wired.com/wired/archive/4.04/netbots_pr.html (listing synonyms for autonomous web robots, such as "spiders, wanderers, and worms. Cancelbots, Lazarus, and Automoose. Chatterbots, softbots, userbots, taskbots, knowbots, and mailbots. MrBot and MrsBot. Warbots, clonebots, floodbots, annoybots, hackbots, and Vladbots. Gaybots, gossipbots, and gamebots. Skeleton bots, spybots, and sloth bots. Xbots and meta-bots. Eggdrop bots.").

[24] *See id.* (stating, "[s]trings of code written by everyone from teenage chat-room lurkers to top-flight computer scientists, bots are variously designed to carry on conversations, act as human surrogates, or achieve specific tasks - such as seeking out and retrieving information.").

[25] *See id.* (noting some of the potential evil purposes for bots: "Bots can be instructed to do whatever their creators want them to do, which means that along with their potential to do good they can do a whole lot of evil. Bots can steal information instead of simply retrieving it. A commercial bot - such as an online-shopping bot or a news-retrieval bot - could be designed to disable its competitors.").

[26] Google "autosniping" or "autobidding" to find a plethora of websites, such as www.auctionsniper.com, that offer bot services to the average Internet users, which will place a final bid in an EBay or similar web auction at the latest possible moment.

[27] Jeffrey M. Rosenfeld, *Spiders and Crawlers and Bots, Oh My: The Economic Efficiency and Public Policy of Online Contracts That Restrict Data Collection*, 2002 Stan. Tech. L. Rev. 3 (2002).

importance of search engines, a policy rationale proffered by the court in *Kelly v. Arriba Soft.*[28] Bots are certainly not going away in the foreseeable future – they are an essential part of the Internet and e-commerce.

Since 2004, courts have dealt with a variety of cases involving the validity and enforceability of browse-wrap agreements when the conduct of assent is "performed" by a bot and not a person. This raises questions of both contract and agency law. Can a bot manifest assent to a browse-wrap agreement like its human counterpart and do the person (the bot operator) and the bot share a principal-agent relationship under the law?

In the case of *Register.com v. Verio*[29], a data-scraping bot encountered a browse-wrap agreement while functioning in a manner clearly in violation of the terms presented, and the defendant claimed that these terms were rejected. The court determined that because the bot visited the site so very frequently, the defendant must have been aware of the terms and likened the situation to this: "Returning to the apple stand, the visitor, who sees apples offered for 50 cents apiece and takes an apple, owes 50 cents, regardless whether he did or did not say, 'I agree.'"[30] This decision piggybacked on some traditional common law legal doctrines of contract, including the concept that silence can function as assent.[31] In an obvious difference, this defendant did not visit the apple stand himself, albeit frequently. His bot, or tool really, took apples without paying. In *Register*, however, the court found that the defendant (through his bot) had notice of the terms offered and that such notice, coupled with an ample opportunity to reject the offer by discontinuing visitation of the plaintiff's website, constituted acceptance.[32]

---

[28] Kelly v. Arriba Soft Corp., 336 F.3d 811, (9th Cir. 2003) (holding that the search engine's bot copying of copyrighted images and the subsequent unlicensed display of thumbnail versions of those images constituted fair use "because use was transformative." Additionally, the court noted that the use "benefitted the public by enhancing Internet information gathering techniques.").

[29] Register.com, Inc. v. Verio, Inc., 356 F.3d 393 (2d Cir. 2004) (defendant Verio used a bot to obtain information, about Register's newly registered domain customers, for use in later in mass unsolicited emails. Verio claimed not to be bound by the terms of use attached to such a use of the data: "By submitting a WHOIS query, you agree that you will use this data only for lawful purposes and that under no circumstances will you use this data to ... support the transmission of mass unsolicited, commercial advertising or solicitation via email.").

[30] *See* Register.com, Inc., 356 F.3d at 403.

[31] RESTATEMENT (SECOND) OF CONTRACTS § 69(1)(A) (1981) ("[S]ilence and inaction operate as an acceptance ... [w]here an offeree takes the benefit of offered services with reasonable opportunity to reject them and reason to know that they were offered with the expectation of compensation.").

[32] *See* Register.com, Inc. 356 F.3d at 403.

About a year after *Register*, the courts dealt with a closely related issue in *Cairo v. Crossmedia Services*.[33] In *Cairo*, the defendant firmly denied having awareness of the terms of use, again, because only a bot, and not the defendant himself, had visited the plaintiff's site. The *Cairo* court cited to *Register* for the proposition discussed previously, which is crucial to the framework established in these cases: "[N]ew commerce on the Internet has exposed courts to many new situations, it has not fundamentally changed the principles of contract . . . when a benefit is offered subject to stated conditions, and the offeree makes a decision to take the benefit with knowledge of the terms of the offer, the taking constitutes acceptance of the terms, which accordingly become binding on the offeree."[34] Both of these decisions shined a spotlight on the bot's notice of the website's offered terms – imputing, vicariously, the bot's notice to the operator. The court seemed to be saying that notice to a bot would constitute constructive notice to the operator, giving the operator knowledge of the terms and knowledge of the fact that continued use of the website would manifest assent. But, as bots continue to evolve, and smart bot operators (and their legal counsel) become aware of the case law, it is likely that routes will be discovered that will effectively reduce the ability of plaintiffs to prove that a bot operator knew they had assented to a website's terms.[35] A possible battle in this regard might already be brewing between Microsoft and Google because of this very type of circumvention related activity.[36] Is there a solution that would be legally enforceable?

Within our current legal doctrine of enforceable contracts that are electronically signed and executed, the law of agency, in the realm of autonomous web bots, might not retain the traditional aspects legal scholars have come to expect.[37] Drafters of two proposed but failed

---

[33] C 04-04825 JW, 2005 WL 756610 (N.D. Cal. Apr. 1, 2005) (holding that "[s]imilar to the circumstance in *Register.com*, Cairo's visits to CMS's web sites with knowledge of CMS's Terms of Use constituted acceptance of the terms, which accordingly are binding on Cairo.").

[34] *Id.* at 5.

[35] What if the bot only visits the site a single time for scraping? What if the bot operator employs a new bot for each scrape? Would this make the terms unenforceable for lack of notice and lack of knowledge of assent under the standard used in *Register* and *Cairo?*

[36] *See* Amit Singhal, *Microsoft's Bing uses Google search results—and denies it*, GOOGLE OFFICIAL BLOG, http://googleblog.blogspot.com/2011/02/microsofts-bing-uses-google-search.html. (last visited Feb. 6, 2014).

[37] These "aspects" of agency law would include the traditional doctrines of actual and apparent authority, agent-third party liability, principal-agent liability and agent-principal liability.

acts[38] tried to establish that an "electronic agent, as defined, is in essence a tool of its user."[39] The problem is that, under a characteristic closely associated with the law of agency, if an agent goes on a "frolic of its own", the act is outside the scope of the agent's authority and can relieve the principal of *respondeat superior* liability.[40] As the drafters of the EUTA realized that bots could someday "act autonomously, and not just automatically",[41] their efforts seemed to predict this very conundrum of agency law principle. A bot is probably not an agent under the concepts of agency law, but rather simply a tool of its operator. Society generally does not condone relieving a criminal suspect of liability simply because the crime was committed with some manner of criminal tool rather than with the suspects bare hands. The operator-bot relationship seems to be not much different than the criminal-tool relationship in that the operator is using a tool - the bot - to visit a website, rather than to visit the site in person.[42] With bot operators potentially protected from liability using the shields of agency and contract law, where is the future of e-commerce and the Internet headed?

II. HOW TO FIGHT THE BOTS

A. Good Bots

Most website providers probably do not want to keep all bots from visiting every part of their site. Keeping a website totally hidden from all types of bots would be akin to a store owner having a hidden location that no customer could possibly find unless they were personally informed of the secret address. The store would have no random foot traffic either. Only invited

[38] The UNIFORM ELECTRONIC TRANSACTIONS ACT (UETA) and the UNIFORM COMPUTER INFORMATION TRANSACTIONS ACT (UCITA).

[39] *See* Stephen T. Middlebrook& John Muller, *Thoughts on Bots: The Emerging Law of Electronic Agents*, 56 Bus. Law. 341, 348 (2000).

[40] *See, e.g.* Makor Issues & Rights, Ltd. v. Tellabs Inc., 513 F.3d 702, 707-08 (7th Cir. 2008) (stating, "[a] routine invocation of respondeat superior, which would impute the mistake to the corporation provided only that it was committed in the course of the employee's job rather than being 'a frolic of his own.'").

[41] *See* note 21, § 2 cmt. 5, at 9, UNIFORM ELECTRONIC TRANSACTIONS ACT (UETA) (1999), *available at* http://www.law.upenn.edu/bll/ulc/ulc.htm.

[42] "In person" might be a strange concept in this context as it refers to a person sitting at a computer, entering a URL into a web browser and viewing a website's content. It is not "in person" in the sense that the human and the website server are likely many miles apart, physically.

guests could shop. Arguably, the most important bots on the Internet are those of major, reputable search engines. Search engines, for quite some time, have been the essential mechanism for how web users find content. Search engines employ bots to report back on the websites that exist in the World Wide Web – not only what the website is and where it is located in cyberspace, but the content of the site, page by page. Bottom line, if you run a business on the Internet, you need to allow the bots of search engines to list and catalog the pages of your site or you will not have any visitors. Of course, there are many reasons to have websites that are only for a select set of visitors, which one would want to keep private from all uninvited visitors, including search engines, but this aspect can easily be accomplished using login and password protection.[43] This is not the problem contemplated in this article.

B. The Robots Exclusion Standard

In 1994, a standard, albeit purely advisory and not legal in nature, was invented to give website providers a mechanism to deal with bots and, in essence, control the bots' access to information and function within their site.[44] Known commonly as the Robots Exclusion Standard ("RES") or Robots Exclusion Protocol, the system essentially involves placing a data file "in the top-level directory of your web server" and using the "the /robots.txt file to give instructions about their site"[45] which can include information for cooperating bots (the key word being "cooperating") to access and follow. This information would ordinarily be a list of which pages and sub-pages within the website are available for scraping and which pages are not. Under this protocol a web designer can have a publicly available website with certain portions that are kept private from search engines, web users and the like. The robots.txt file could certainly contain data other than file directory access lists – it could contain the text of the terms of use offered by the website designer. This may sound like a quick and easy solution, until one considers the amount of bots that may exist, now and in the future, which are likely to be

[43] Note how your bank account information does not appear in a Google search of your own name and information. Google your full legal name and Social Security number – you won't find anything. (Author is not responsible for any identity theft as a result of trying this.)

[44] Martijn Koster, *A Standard for Robot Exclusion*, ROBOTSTXT, http://www.robotstxt.org/orig.html#status (last visited Feb. 5, 2014). The invention of the RES is generally attributed to Martijn Koster.

[45] *See About /robots.txt : In a nutshell, available at* http://www.robotstxt.org/robotstxt.html.

operated by those who will not cooperate and follow the RES. It is an unfortunate situation, as the RES has become, quite organically, the "de facto standard on the Internet."[46] This only works to magnify the true flaw of this otherwise effective system: it relies on compliance instead of mandating it under force of law.[47] A similar, albeit less commonly used system also exists, known as Sitemaps[48], which is essentially the reverse of the RES, allowing for the *inclusion* of particular uniform resource locators ("URLs") and a limited amount of accompanying meta-data.[49] Under this system, web designers can submit, periodically, to each chosen search engine, a Sitemap file that details the content and location of each page and directory within the site. It is theoretically feasible that Sitemaps could perform the same function, with regard to containing the textual terms of use of the website, as a legally enforceable RES system, but, as RES is already a more commonly employed as a tool for this purpose, and as RES would apply to all visitors of a website (not just those who have received a Sitemap file), it would seem to jump out as the obvious choice.

III. THE EFFECTIVENESS OF THE CURRENT FRAMEWORK AND WHY IT WON'T LAST

A. What the courts can do right now

When a court addresses a dispute akin to *Cairo* or *Register*, it must engage in a factual search for some evidence demonstrating the bot's ample notice of the terms offered by the website. This notice translates to the bot operator's notice and the bot operator further has knowledge that they have assented to the terms - all because the repeated actions of the bot. What can courts do when the evidence of notice, and hence the bot operator's knowledge of assent, cannot be construed? Both of the browse-wrap cases previously examined in the context

---

[46] *Id.*

[47] *See* Lourenço, A. & Belo, O., *Applying Clickstream Data Mining to Real-Time Web Crawler Detection and Containment Using ClickTips Platform., available in* Reinhold Decker, Advances in Data Analysis:

Proceedings of the 30th Annual Conference of The Gesellschaft Für Klassifikation E.V., Freie Universität Berlin, March 8-10, 2006 (Google eBook) (discussing how web crawlers can be confused with regular users or even impersonate others – the programs do not have any strict regulation enforcing and limiting their actions).

[48] *See* http://www.sitemaps.org/ (defining Sitemaps as an XML file which allows webmasters to include additional URL information including when it was last updated, how often it changes, and how important it is in relation to other URLs in the site).

[49] Meta-data is data that is attached to or about other data: in this example, the URL itself is the data, and information about the content of the URL, how often updated, etc., would represent the meta-data.

of bot assent, *Register* and *Cairo*, basically determined that frequent bot scrapings of a site puts the bot operator on notice of the terms – after which, the continued visitation and scraping of the site constitutes assent to the terms within. The situation that might throw courts a curveball, and which is bound to present itself sooner rather than later, is one where the bot is making a single scraping, never to return, and the bot operator cannot be found to have actual notice, or the constructive notice essentially applied in *Register* and *Cairo*. The court could try and force the facts of a hypothetical situation like this into the limited framework of *Register* and *Cairo*, but it would be a difficult, if not futile, task. Those cases sought to find clear notice and, within this hypothetical, it will become much more difficult for the court to ascertain in the future. Difficulties under this framework might also arise under other, more complicated, methods of attempted circumvention. For example, what if a clever data scraper turned unwitting Internet users themselves into bots? This might not be merely a hypothetical situation for long – a dispute may already have begun between Microsoft and Google, brewing since at least 2011, which has the potential to explode into a landmark case-in-the-making regarding browse-wrap agreements.[50]

It would be naïve to expect that in a world of bot use proliferation, plenty of it for illegitimate and nefarious purposes, that there will not be bot operators, with qualified attorneys,[51] who will advise their clients how to avoid liability under this current framework. This article has already discussed examples of how that might happen, effectively giving the bot operator a loophole to avoid responsibility under the law.

B. Is a Bing-Google storm brewing?

When considering how a potential framework-defying case might come about, the question arises as to how a potential plaintiff would become aware that a bot had visited its site and in some way violated its terms of use. In *Register*, the plaintiff became aware of what was occurring, because the company began receiving complaints from customers (those who had recently registered new domain addresses with the plaintiff) complaining about the unsolicited

[50] *See* Amit Singhal, *Microsoft's Bing uses Google search results—and denies it*, GOOGLE OFFICIAL BLOG (Feb. 1, 2011), http://googleblog.blogspot.com/2011/02/microsofts-bing-uses-google-search.html.

[51] Some, not all, attorneys are qualified.

mass emails they had received from the defendant company. The defendant was first easily discovered because it explicitly referenced the plaintiff company in the solicitations made to the plaintiff's customers. The defendant quickly changed the emails to not include the reference to the connection to the plaintiff, but at this point, the game was up.[52] This case demonstrates not only how a plaintiff can become aware of malfeasance, but also how a defendant will act with intent to avoid detection and responsibility.

In *Cairo*, the plaintiff was put on notice of a violation of its terms of service when it reviewed its server logs and easily discovered that the defendant was copying its promotional and circular materials and reposting the data on the defendant's site, without permission from the plaintiff.[53] This represents another method that can be used by plaintiffs to discover a defendant's actions. Additionally, the defendants in *Cairo*, alleged that its "computer search programs cannot read the Terms of Use posted on a web site, and they do not report the presence of such Terms of Use."[54] Simply put, a bot operator openly displaying a purpose to avoid responsibility, while claiming that the bot shields him from liability.

Apparently, Bing, the search engine arm of Microsoft[55], and hopeful competitor of Google, has been violating Google's terms of service by means of a fairly clever "back door" attempt to reverse engineer Google's prized search algorithm.[56] In addition to using bots (spiders) to crawl the web compiling data on every possible website, Google uses a mathematical algorithm to determine which of those listings are most relevant as results to return to its users during search queries. The difference between Google's algorithm and that of its competitors is what distinguishes the quality of their respective search engine products. Google's terms of service clearly state two important things: "[b]y using our Services, you are agreeing to these terms" and "[y]ou may not use content from our Services unless you obtain permission from its owner or are otherwise permitted by law."[57] Any attempt by Bing to reverse engineer Google's

---

[52] *See* Register.com, 356 F.3d at 397.

[53] *See* Cairo, Inc., 2005 WL 756610, at * 3.

[54] *Id.*

[55] Bing is simply the name a Microsoft *product*. BING, *available at* http://www.bing.com/?publ=DBING (last visited Feb. 20, 2014).

[56] *See* Singhal, *supra* note 36.

[57] GOOGLE, *Terms of Service*, last modified March 1, 2012, *available at* http://www.google.com/intl/en/policies/terms/.

algorithm or other search content results would almost certainly violate the terms – but only if Bing is bound by those terms.

The accusation, from Google's own Official Blog, is that Bing is monitoring the search queries and results that are done through Microsoft's Internet Explorer 8 ("IE8") web browser[58]. When IE8 users enter a Google search using IE8, the resulting data is being reported to Microsoft (Google suspects), which is then being used to bolster Bing's performance. Google appears to have caught Bing red-handed, using several "synthetic queries" (essentially bogus queries and results) that were then found to exist within Bing with no other reasonable or possible explanation, a la *Feist*.[59] No lawsuit has been filed to date, but Google's response to the situation does not show approval: "We look forward to competing with genuinely new search algorithms out there—algorithms built on core innovation, and not on recycled search results from a competitor."[60] Imagine for a moment that Google sues Bing over this issue – what result under the current framework?

Allegedly, Bing is not technically visiting Google's site (human or bot) but, rather, is monitoring the actions of its own users, who are visiting Google during their day-to-day use of IE8. Although the IE8 users have likely consented to this practice in the fine print of the IE8 terms of use, they have still essentially been turned into bots by Microsoft. Microsoft seems to have found a clever way (intentionally or not) to avoid the *Cairo* framework – by letting their IE8 users perform the bot scraping (simply by using Google search) and subsequently collecting that data from the IE8 users, data that, most importantly, did not come directly from Google. Assuming the IE8 users are subject to Google's terms, would Bing also be subject to those terms because they were monitoring the search queries of its customers? Would IE8 users be pleased to learn that they have unwittingly violated Google terms of service? Would a court find that Microsoft was aware of Google's terms of service with ample opportunity to review and reject them under these specific circumstances? With so many unanswered questions from just one real-world example, only a simple, unified approach will suffice.

---

[58] On March 14th, 2011 a newer version of the browser, Internet Explorer 9, was released.

[59] *See* Feist Publications, Inc. v. Rural Tel. Serv. Co., Inc., 499 U.S. 340 (1991) (wherein the plaintiff placed fake telephone listings in its product to determine if the defendant was copying. The trick worked for proving the defendant's actions, but liability was not found for reasons of copyright law and originality).

[60] Singhal, *supra* note 36.

IV. A RECOMMENDED SOLUTION TO FUTURE PROBLEMS

A. Congressional Action

When a tool, designed for this very purpose, already exists but suffers from a fatal flaw that deprives it of legal enforceability, Congress should step forward and mandate the use of the RES for all bot operators. This law would not apply to website designers, who would have the *caveat emptor* option of utilizing this tool or not when making and publishing their website. The law should require that bot operators must program their bot to access a robots.txt file every time it visits a site and the bot must report all data in the robots.txt file back to the operator. Website designers could use the file to exclude certain portions of the site from scraping and could also include their terms of use in the RES file. This would take the framework, as it applies to humans in the line of cases discussed in Section I(C) and make it equally applicable to bots, whether the operator of such bot claimed to have notice of terms or not. In the same way that a physical property owner could post a warning at the perimeter of his property, a website designer can expect that terms of use will be binding on human visitors if they are displayed conspicuously and, now, a designer could expect their terms to be binding on a bot operator. Under this proposed law, a bot operator should be deemed, under the law, to have been provided ample opportunity to see and review the terms of service on the very first visit to the site, if those terms are made available in the robots.txt file. The law would leave little room for a bot operator to claim that they have not assented to the terms for lack of notice. The law could also create a mandatory principal-agent relationship between the operator and the software, regardless of whether the bot is automatic or autonomous. The operator would always bear responsibility for the actions of the bot, no matter how clever their programming might be. This law might be the easy part; the hard part would involve international enforcement in a Web that is, after all, World Wide.[61]

[61] The law would need to become a part of the TRIPS treaty (administered by the World Trade Organization) that requires signatories to have minimum standards of intellectual property regulation. Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, THE LEGAL TEXTS: THE RESULTS OF THE URUGUAY ROUND OF MULTILATERAL TRADE NEGOTIATIONS 320 (1999), 1869 U.N.T.S. 299, *available at* http://www.wto.org/english/tratop_e/trips_e/t_agm0_e.htm.

Incorporation of this law into TRIPS[62] or a similar enforceable treaty would prevent bot operators from avoiding the liability created under the new American law by simply taking their operations offshore. The operator could simply locate their servers and computers in a country that does not mandate use of the RES system. As website terms of service are often in place to help protect the intellectual property of the website, particularly copyright and trademark, fundamental arguments could be made that TRIPS is the proper vehicle to implement RES legal standards world-wide. Issues, however, may arise within the clauses, common to website terms of use, that deal with choice of forum and arbitration. Bot operators would not be pleased with being subject to the arbitration and choice of forum clauses for sites in all corners of the world.[63] Nevertheless, the notable strengths of arbitration clauses, an "important mechanism for defining rights and obligations resulting from new contractual forms" in the international commerce context, would easily outweigh the concerns of inconvenience.[64]

The use of arbitration clauses in international contracts provides many positive benefits including flexibility of procedure, lower costs and speed.[65] In the long run and in the interest of promoting free trade, "treaty-based, international arbitration offers a private court system that enforces contracts."[66] In fact, these clauses have been referred to, by the Supreme Court, as an "indispensable precondition to achievement of the orderliness and predictability essential to any international business transaction."[67]

B. Conclusion

This proposed solution of this paper solves the potential problems that could arise under the limited framework provided by *Register* and *Cairo*, yet still gives web designers the flexibility needed to work with the "good bots", so that the future of e-commerce, in this context, can proceed uninterrupted. Web designers can breathe a sigh of relief knowing that any

---

[62] Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994, Marrakesh Agreement Establishing the World Trade Organization, Annex 1C, Legal Instruments--Results of the Uruguay Round vol. 31, 33 I.L.M. 1197 (1994).

[63] TRIPS currently applies in all WTO states, *see http://www.iipa.com/trips.html*

[64] Thomas H. Oehmke, *Arbitrating International Claims—At Home and Abroad*, 81 AM. JUR. TRIALS 1 (2001).

[65] *Id.* at §3.

[66] *Id.*

[67] *See* Scherk v. Alberto-Culver Co., 94 S. Ct. 2449, 2455 (1974).

misappropriation of their site or its services or content will not go unpunished. Web designers can operate their site in a free and open manor, making it easily accessible to major search engines, individual users and other relevant scrapers, without having to implement costly and time consuming monitoring practices to become aware of and fend off potential wrongdoers. The future of e-commerce on the Internet is ever expanding and nobody can say for sure when the explosion in growth and importance will subside. Having a law in place to regulate the liability of bot operators - one with international enforcement, that removes the burden of responsibility from the shoulders of the website designer and places it squarely on that of the bot operator, where it belongs - would encourage and incentivize e-commerce investment while de-incentivizing unscrupulous bot operation.