

# Intellectual Property Aspects of Web Publishing

Holger M. Kienle and Daniel German  
University of Victoria  
Victoria, Canada  
{kienle,dmgerman}@cs.uvic.ca

Scott Tilley  
Dept. of Computer Sciences  
Florida Institute of Technology  
stilley@cs.fit.edu

Hausi A. Müller  
University of Victoria  
Victoria, Canada  
hausi@cs.uvic.ca

## ABSTRACT

This paper addresses how intellectual property affects the Web in general, and content publishing on the Web in particular. Before its commercialization, the Web was perceived as being free and unregulated—this assumption is no longer true. Nowadays, content providers need to know which practices on the Web can result in potential legal problems. The vast majority of Web sites are developed by individual such as technical writers or graphic artists, and small organizations, which receive limited or no legal advice. As a result, these Web sites are developed with little or no regard to the legal constraints of intellectual property law. In order to help this group of people, the paper tries to answer the following question: What are the (typical) legal issues for Web content providers to watch out for? This paper gives an overview of these legal issues for intellectual property (i.e., copyrights, patents, and trademarks) and discusses relevant law cases. As a first step towards a more formal risk assessment of IP issues, we introduce an IP maturity model that captures a Web site's IP coverage with five different maturity levels.

## Categories and Subject Descriptors

K.5.m [Legal Aspects of Computing]: Miscellaneous;  
H.5.4 [Information Interfaces and Presentation]: Hypertext/Hypermedia; D.2.9 [Software Engineering]: Management—Copyrights

## General Terms

Legal Aspects

## Keywords

Copyright, deep linking, hypermedia, intellectual property, link law, trademarks, patents, open content, World Wide Web

## 1. INTRODUCTION

*“... if the promise of hypertext actually is to be realized and disseminated to users, then early-stage identification of legal issues and compliance with legal obligations will be necessary.”*

– Jones at Hypertext '87 [15]

The above quote is from the first conference on hypertext in 1987. At the time, hypertext systems existed mainly as research prototypes and had little commercial significance. Even though the question of the legal implications of hypertext systems was raised early on, there has been little research in this direction—a notable exception being Nelson's Xanadu system, which has introduced a mechanism for dealing with copyright issues [26] [32]. In the meanwhile, the Web has emerged as the most popular hypermedia system by far, achieving widespread hypertext authoring for the first time.

The Web has made a transition from an academic research network to a commercialized commodity. In the early days of the Web, the main content publishers were academics that wanted to make information freely accessible to interested readers, typically fellow researchers. Unrestricted linking between documents of unrelated authors was part of the publishing process. In fact, the simplicity of both publishing and interlinking of documents appealed to many people and was a major driver for the Web's success. Furthermore, the Web was perceived as being free and unregulated [22, 25].

At some point, business started to realize that the Web provided new opportunities for them. A business' Web presence is now seen as an important success factor that affects its reputation, sales, and marketing. Since Web sites are now perceived as assets, businesses increasingly want to protect their Web presence and control the way in which others utilize their Web sites. As a result of the increasing commercialization of the Web, legal issues play a more and more important role for both content providers and consumers. Until 1994, courts in the United States did not have

to deal with Web cases. In 2001, the state and federal courts have published more than 1,100 opinions involving the Web [18]. Whether one likes this development towards a more regulated Web or not, content providers and users need to know which practices on the Web can result in potential legal problems.

The vast majority of Web sites are developed by individuals and organizations that receive limited or no legal advice. As a result, these Web sites are developed with little or no regard to the legal constraints of law. Consequently, this paper assumes that the reader has an information technology background (e.g., from computer science or technical writing), but no (formal) background in law. In order to help this group, the paper tries to answer the following question:

What are the (typical) legal issues for Web site owners to watch out for?

Web site owners should know about legal issues from two perspectives: as content providers and as consumers. On the one hand, they need to know how to protect their own (intellectual) property, embodied in their Web sites. On the other hand, they need to know the legal constraints of utilizing someone else's Web site. Legal problems arise when Web site owners believe that their rights have been violated. It is hard to predict what a certain owner will view as a violation. For example, some content providers view a deep link into their Web site as a copyright infringement (cf. *Ticketmaster vs. Tickets.com*), while others welcome such links because they bring visitors to their site.

This paper focuses on legal issues raised by intellectual property (IP) law because most relevant cases fall in this area. IP law includes copyright, trademarks, and patents. It provides protection of, for instance, a Web site's content, designs, processes, and implementation. However, it should be noted that IP is not the only legal concern on the Web. For example, trespass law has been used to challenge unwanted emails and Web spiders [28]. In one case, eBay sued Bidder's Edge for trespass to chattel<sup>1</sup> because of a Web crawler that frequently visited eBay's site to scrape auction information [35].

This paper cannot give legal advice or rules that state what kinds of conduct violates IP law. Unfortunately, case law is still evolving and there is significant legal uncertainty. However, the paper serves to highlight the areas of IP that have caused legal problems for Web content providers in the past. This allows content providers to make a more informed risk assessment in the future.

This paper is organized as follows: Section 2 gives a brief survey of IP legal issues surrounding Web sites. The discussed issues are of interest to Web site publishers from

two perspectives: (1) opportunities to protect their own IP and (2) potential legal threats when using IP of others. Section 3 extends the document maturity model introduced by Huang and Tilley [16] to reflect IP issues. We believe that since legal issues are of more and more concern to Web site owners, they should be part of an evaluation of the quality of a Web site. Section 4 summarizes the paper and draws some conclusions.

## 2. INTELLECTUAL PROPERTY ASPECTS

A general definition of IP is that of useful information or knowledge. IP is a broad field, covering copyright, patents, and trademarks. IP law gives a creator of IP certain exclusive rights to exploit his or her creation, thus "preserving the motivation needed for people to undertake the development of new and useful things that could benefit society" [6]. In doing this, IP law has to strike a balance between IP holders and the public. Judge Kozinski puts it as follows [23, page 204]:

"Overprotecting intellectual property is as harmful as underprotecting it. Creativity is impossible without a rich public domain."

As further discussed below, this balance is important for both patents and copyright.

Table 1 gives an overview of the legal issues addressed in this paper. Because of space limitations, this paper can provide only a brief treatment. For this reason, important legal cases are given along with references that provide more in-depth discussion.

In the following, we assume U.S. law and explicitly note when discussing laws in other countries. There are several international treaties (e.g., for copyright there is the Berne Convention of the World Intellectual Property Organization—WIPO) that are the basis for national laws, guaranteeing a certain uniformity. However, there are still many more or less subtle differences among countries. For example, the term of copyright in the world is 50 years after the author's death, except in the United States where it is 70 years—this is the result of the *Sonny Bono Copyright Term Extension Act* passed by Congress in 1998.

### 2.1. Copyright

A copyright applies to work that is expressed in a "tangible medium for the purpose of communication" [4]. Examples of such works are literature, music, paintings, movies, and television broadcasts. In some countries, copyright also includes computer programs and information embedded in databases.

The copyright law has to balance the rights of the content creators to be compensated adequately for their efforts, and

<sup>1</sup>Chattel is a legal term for personal property such as books, cars, and computers [35].

IP Law	Infringement	Legal Cases
copyright	digital copy	<i>Playboy vs. Russ Hardenburgh</i> (1997) [27, p. 12B-12]
	intermediate RAM copy	<i>Intellectual Reserve vs. Utah Lighthouse Ministry</i> (1999) [14] [11], <i>Religious Technology Center vs. Netcom</i> (1995) [14], <i>MAI Systems vs. Peak Computers</i> (1993) [38]
	linking	<i>Ticketmaster vs. Tickets.com</i> (2003) [14], <i>Washington Post vs. TotalNews</i> (1997) [39] [20] [31], <i>Futuredontics vs. Applied Anagramics</i> (1998) [14] [31], <i>NVM vs. De Telegraaf</i> (2000) [17]
	thumbnails	<i>Kelly vs. Arriba</i> (2003) [41] [42] [31]
patent	business methods	<i>Amazon vs. Barnesandnoble.com</i> (2002) [40] [18] [45], <i>Priceline vs. Microsoft</i> (2001) [18]
	software	<i>Eolas vs. Microsoft</i> (2004) [1]
trademark	HTML meta tags	<i>Brookfield vs. West Coast Entertainment</i> (1999) [30] [21]
	HTML contents	<i>Playboy vs. Welles</i> (1999) [30] [21]

**Table 1. Summary of IP issues that affect Web content publishers**

the rights of citizens for access to information. Copyright grants several *pecuniary rights*, which are held exclusively by the owner. Among these are making and selling copies of the work, and making derivative works. A consumer right is *fair use*, which allows the consumer to use copyrighted material in certain ways regardless of the copyright owner's wishes [33, 34]. For example, limited quoting from a copyrighted book is permissible under fair use. Copyright protection currently endures for the life of the author plus an additional 70 years after the author's death.

Typically, a copyright is owned by the author of the work. However, it is also possible that the copyright is owned by the author's employer (i.e., *work for hire*) [8]. Depending on the business and employer relationship, it can be difficult to determine the copyright owner (e.g., for independent contracting). This means that a company might not have copyright of its own Web site. A development agreement can be used that explicitly states whether the site is considered work-for-hire or not, and whether the developer transfers the copyright to the company [18].

A copyright is subject to no formalities, which means that no action is required from an author to obtain a copyright. Hence, the content of a Web site such as text, images, and JavaScript code has a copyright even without an explicit copyright notice or registration. While a copyright notice is not required, it puts others on notice that the owner has expressed explicit interest in the copyright. The Copyright Act says that "the notice shall be affixed to the copies in such manner and location as to give reasonable notice." The Copyright Office has not issued guidelines for Web sites, but it is common to include the notice in the footer of every page.

As with other works, Web sites can be also registered by the Copyright Office [7]. Such a registration is a nec-

essary prerequisite to file an infringement suit. Since Web sites tend to evolve, repeated registration of the site might be necessary. In *IMS vs. Berkshire*, IMS alleged that Berkshire copied content from eight pages of its e-Basket Web site, which tracks magazine advertising. The obtained information was then used in a competing system. IMS published the e-Basket site on January 14, 2003 and registered it on March 7, 2003. Berkshire allegedly accessed the site around March of 2002. IMS claimed that the registered site is the same one that Berkshire infringed upon. However, IMS failed in its copyright registration to refer to its pre-January-2003 site as pre-existing work. Because of this, a district court ruled that the e-Basket site that Berkshire had accessed in 2002 was not protected by the registration secured in 2003, and therefore dismissed the claim to copyright infringement [2].

Not all works are protected by copyright. For example, Web sites and other publications of the Federal Government are in the public domain. Web sites can also use an explicit open content license to replace copyright. Open content licenses are similar in that they allow the (non-commercial) copying of content, which can be a more appropriate business model than copyright. Examples of open content are Creative Commons [9] and the GNU Free Documentation License (GFDL) [12]. The GFDL is primarily meant for software documentation, but not restricted to it. For example, the Wikipedia online encyclopedia uses the GPL for its content [44].

In the rest of this section, we discuss several practices of Web publishing that can pose legal risks because of copyright law.

**Digital Copies.** Since works published on the Web are protected by copyright even without a notice or registration, copying and reposting is likely to constitute a copyright in-

fringement. This is contrary to popular assumptions that content on the Web is “free.” To clarify, Pamela Samuelson, a law professor, explicitly states: “If some people think it’s OK to post other people’s work on a bbs [i.e., bulletin board service] or listserv because the author didn’t put copyright notices on work, they are relying on a mistaken assumption” [34]. There are several supporting cases, for instance, *Playboy vs. Russ Hardenburgh*. Hardenburgh encouraged subscribers of his BBS to submit copyrighted images of Playboy magazine and then made them available for download. The court decided that this infringed Playboy’s copyrights and that the defendant is liable for direct infringement [27]. As a general guideline, Isenberg, attorney and founder of GIGALAW.com, recommends to not “publish anything on your Web site unless you know who created it” [18]. While this might seem obvious, it is often ignored in practice.

**Intermediate RAM Copies.** Viewing of a Web page in a browser creates a copy of the page in the user’s computer. This raises the question whether this intermediate RAM copy is a reproduction of the work and hence constitutes a copyright infringement.

Several cases have found that intermediate copies are indeed covered by copyright. In *Intellectual Reserve vs. Utah Lighthouse Ministry*, the owners of a Web site hosting copyrighted material were held liable of contributory infringement because of intermediate RAM copies in the users’ browsers [11]. Addressing the users’ RAM copies, a district court stated [14]:

“When a person browses a website, and by so doing displays the [copyrighted material], a copy of the [copyrighted material] is made in the computer’s random access memory (RAM), to permit viewing of the material. And in making a copy, even a temporary one, the person who browsed infringes the copyright.”

In its decision, the court relied on *MAI Systems vs. Peak Computers*, in which an appellate court ruled that loading software into a computer creates a fixed copy of that software [38]. Even though this interpretation of intermediate copies has been widely criticized, it seems now firmly established in case law.

According to this view, a user activating a link may be an infringer and the author of the Web page containing the link may be a contributory infringer. However, this interpretation is limited by fair use; one can argue that by making a Web site available on the Web, content publishers expect others to browse (and download) their work (cf. *Religious Technology Center vs. Netcom*).

**Linking.** Spinello observes that “linking is the essence of the World Wide Web, and there is little doubt that legal or technological constraints on linking would have substantial negative ramifications for the web” [37]. Regardless, a

number of legal cases suggest that Web content providers cannot expect to freely link to other Web sites as a general rule.

Berners-Lee calls it a myth that a normal link could possibly infringe copyright. He makes the observation that “making [a] reference with a hypertext link [as opposed to giving the URL without explicitly linking to it] is more efficient but changes nothing else” [5]. A district court in *Ticketmaster vs. Tickets.com* came to the same conclusion [14]:

“hyperlinking does not itself involve a violation of the Copyright Act (whatever it may do for other claims) since no copying is involved.”

However, linking to a page that infringes on copyright can be problematic. In *Intellectual Reserve vs. Utah Lighthouse Ministry*, a Web site first illegally published copyrighted material for viewing. After being directed by a district court to remove the material, the site complied, but instead posted links to other sites that illegally contained the same material. The court ruled that users that viewed the material infringed on copyright because of the temporary RAM copies and that, hence, the links themselves constitute contributory infringement.

Another potential problem is the practice of some Web sites to frame content of other sites. Frames allow it to display content from different sources in one browser window, thus making it “easy to create the impression that the owner of the surrounding frames is in fact responsible for the defining document” [5]. From a copyright perspective, such a composition might constitute a derivative work [39]. In this case, the user of the Web browser might be the direct infringer, the frame provider a contributory infringer. The first case that involved framing was filed in 1997 by a group of six content providers (among them the Washington Post) against TotalNews, which operates a Web site that frames the Web sites of online newspapers. TotalNews packaged the news stories with their own advertisement in a separate frame. In the complaint, the plaintiffs claimed, among other things, copyright infringement. The case was settled out of court in December 1997. As part of the settlement TotalNews stopped framing the suing newspaper sites. In *Futuredontics vs. Applied Anagramics*, the latter framed Web pages of its competitor, Futuredontics, and argued that its “frame provides a ‘lens’ which enables Internet users to view the information that [Futuredontics] itself placed on the Internet.” An appellate court ruled in favor of Applied Anagramics finding that no derivative work was created by framing.

Some site owners have tried to prohibit *deep links* into their Web sites. A link is called deep if it bypasses the front page (or official entry points) of a Web site. In *Ticketmaster vs. Tickets.com*, Ticketmaster sued a competitor, Tick-

ets.com, in July 1999 for using deep links to their event pages [11]. Tickets.com advertises events and sells tickets. If Tickets.com does not sell tickets for an event, it provides a link to another ticket broker. If the exclusive broker for an event is Ticketmaster, customers can click on a hyperlink (“Buy this ticket from another on-line ticketing company”), which shows them Ticketmaster’s event page. Ticketmaster claimed that URLs to interior pages are subject to copyright protection. The district court disagreed, stating that “a URL is simply an address, open to the public, like the street address of a building, which, if known, can enable the user to reach the building. There is nothing sufficiently original to make the URL a copyrightable item.”

While deep linking is probably permissible under U.S. copyright law, the European Union’s database protection directive can affect deep linking [24].<sup>2</sup> The directive defines a database as “a collection of works, data or other independent materials arranged in a systematic or methodical way and individually accessible by electronic or other means” [17]. This broad definition can cover Web pages that provide a collection of information, such as job postings or news stories. For example, in *NVM vs. De Telegraaf*, De Telegraaf operated a real estate search agent that extracted listings from the Web site of a broker [17]. A lower court held that even extraction of small pieces of information can be infringing if that data is of great value to end users. Subsequently, the Dutch Supreme Court ruled in the case that owners of online databases can prohibit deep linking to the contents of their database.

Wood analyzes the legal risk surrounding linking and gives the following recommendations [46]:

- Don’t deep link without permission.
- Don’t frame other entities’ pages without permission.
- Do not use links to sites that encourage others to engage in copyright infringement.

Because of the legal uncertainty that is still prevailing, these guidelines are conservative and hence quite restrictive.

**Thumbnails.** In *Kelly vs. Arriba Soft*, the courts had to decide whether the use of copyrighted pictures as thumbnails is permissible. Kelly is a professional photographer who offers his photographs for sale through his Web site. Arriba Soft is an image search engine, which indexed Kelly’s photographs. As part of the indexing, thumbnail pictures are generated. Kelly alleged that the thumbnails constitute a copyright infringement. Kelly argued that thumbnails contain most of the creative elements of the corresponding full-size picture (such as subject, composition, and lighting). In this case, the court decided that Arriba’s thumbnails are fair use (based on the four fair use factors

given in the Copyright Act). However, one should not generalize from this ruling that thumbnails are always permissible. In fact, the court cautioned that “we do not suggest that the inferior display quality of a reproduction . . . will always assist an alleged infringer in demonstrating fair use.” Still, according to Sableman, the ruling indicates that “the fair use doctrine is likely to protect normal indexing and summarizing use of Internet content as hyperlinks” [31].

## 2.2. Patents

Similar to copyright, patents can foster innovation because they give an innovator an incentive to come up with new ideas; otherwise, the idea could simply be taken (without paying the inventor’s cost). To obtain the patent, the inventor has to fully disclose it to the public. In return, the patent grants the inventor the right to exclusively use it for 20 years.

In contrast to copyrights and trademarks, an inventor must file a patent application with the Patent and Trademark Office (PTO). To be patentable, an invention must meet certain tests, among them novelty and nonobviousness. An invention is novel if it is not part of *prior art* (i.e., the existing body of technology that could be reasonably known to someone working in the field). An invention is nonobvious if it is not an obvious variation of prior art. Once a patent is granted, the invention should be marked with the word “patent” followed by the patent number. “Patent pending” is often used to inform people that a patent has been applied for [6].

Traditionally, patents have been granted for industrial innovations such as machines and manufacturing processes. However, patent law has increasingly broadened its scope to accommodate technological changes. For example, business methods and software are now patentable as well. Amazon’s 1-Click patent (discussed below) is an example of a business method.

In a lawsuit, the patent holder has to show that at least one claim of the patent covers its product or process. The alleged infringer can try to establish that the patent is invalid (e.g., due to prior art). The patent holder can ask for damages, typically profits that have been lost as a result of the infringement. Compared to copyright registration, the cost of obtaining and enforcing a patent is significant. Patent attorneys, (professional) prior art searches, and several fees are part of the application and maintenance process. Furthermore, patent lawsuits are expensive and maybe, as a consequence, many of them are settled. Lessig states that it takes on average \$1.5 million for both sides to take a patent dispute to trial [23, page 213].

Kirsch, a patent attorney, characterizes the increasing importance of patents as follows [18, page 113]:

“Technology companies that just a few year

<sup>2</sup>The U.S. Congress is currently considering a similar law [13].

ago wouldn't have ever considered the impact of patent protection—either offensively or defensively—are now devoting many resources to ensuring they are protected, and minimizing the possibility of infringing third-party patent rights. Those companies that don't consider these issues do so at their peril.”

In the following, we discuss two patent cases to exemplify the impact that they can have on Web sites and Web technology.

**Business Methods Patents.** Patents about Web technology or functionality held by one organization can impact other Web sites. The 1-Click patent<sup>3</sup> that Amazon obtained describes an arguably obvious method of how to do e-commerce on the Web. The basic idea is to purchase an item with a single click, using the user's shipping address and credit-card information already on file. Briefly after the patent was granted, Amazon sued Barnesandnoble.com and won an injunction in December 1999 (just in time right before the Christmas shopping season) [45]. It seems that the court's decision had an impact on other Web content providers—Apple put out a press release in September 2000 that it had licensed the 1-Click patent from Amazon for its Apple Online Store [3]. Barnesandnoble.com appealed and a higher court overturned the preliminary injunction in February 2001, also raising doubts about the validity of the patent [40]. In that decision, the case was remanded to another appeals court for a full trial. However, the parties decided to settle in March 2002. The terms of the settlement were not disclosed.

*Priceline vs. Microsoft* is a similar case, in which Priceline claimed that Expedia's *Price Matcher* service infringed on its reverse auction patent [18]. Expedia agreed to pay royalties and the case was settled.

**Software Patents.** An example of a recent software patent case is a claim from Eolas that it holds a patent for the automatic downloading of embedded content (i.e., plug-ins).<sup>4</sup> The patent potentially affects content that is referred to by `<embed>`, `<object>`, and `<applet>` HTML tags. In fact, the patent contains the following example of an HTML tag to describe an embedded object:

```
<EMBED TYPE = "type" HREF = "href"  
      WIDTH = width HEIGHT = height>
```

Eolas sued Microsoft, claiming that Internet Explorer infringes on the patent. (Other browser vendors such as

Netscape and Opera are potential targets as well.) A jury granted Eolas \$521 million in damages in August 2003, which was affirmed by court [1]. In the meanwhile, the PTO decided November 12, 2003 to re-assess the validity of the patent. There are several proposals to work around the patent. One of them is to insert a dialog box that explicitly prompts the user to launch the embedded application to get around the patent's definition of an “automatically invoked” application. Another one proposes to use DHTML to launch applications instead of statically embedded HTML tags. Regardless of the workaround, it will have a significant impact on Web site developers and users.

## 2.3. Trademarks

According to the Lanham Act, a trademark includes “any word, name, symbol, or device ... used by a person ... to identify and distinguish his or her goods.” Similar to a copyrighted work, trademarks need not be registered; they are acquired through use. Conversely, a trademark has to be actively in use, otherwise it becomes invalid. Federal trademarks, which are registered with the PTO, are identified with the “®” symbol. Unregistered trademarks can be identified with the “™” symbol to put others on notice that a trademark is claimed.

Trademark law allows several companies to hold the same trademark, as long as there is no consumer confusion (e.g., because of different markets or products). A problem arises for domain names, because only one trademark can hold the corresponding domain name. This has resulted in many domain name disputes [36].<sup>5</sup>

Trademark owners can sue for infringement if someone uses their trademark in a way that is “likely to cause confusion” by consumers. If the trademark is famous, owners may prevent others from using the same trademark under the theory that other uses would *dilute* the value of the famous mark even if there is no likelihood of confusion. Deciding what constitutes infringement or dilution is difficult to determine and as a result rulings are hard to predict. Furthermore, Isenberg observes that “often, it seems that, especially in Internet cases, some courts have gone out of their way to find a trademark famous (and therefore protected by the anti-dilution law) if the defendant using the trademark appears to be acting in bad faith” [18].

In a panel about link law, one lawyer summarized the legal situation as follows [29]:

“Basically if you use third party trademarks in links or in meta-tags without permission you're taking a risk. There is the possibility you will get sued for trademark infringement or dilution.”

<sup>3</sup>The patent was granted in the United States on September 28, 1999 under the title “Method and system for placing a purchase order via a communications network” (U.S. patent 5,960,411).

<sup>4</sup>The patent was filed in 1994 and granted in 1998. It is entitled “Distributed hypermedia method for automatically invoking external application providing interaction and display of embedded objects within a hypermedia document” (U.S. patent 5,838,906).

<sup>5</sup>We do not further discuss this issue here because these disputes center around a Web site's URL and not its content.

In the following, we discuss how trademarks affect (commercial) contents and the use of meta tags.

**Contents and Disclaimers.** *Playboy vs. Welles* is a case that shows that judicious use of trademarks can be fair use [21]. Terry Welles is a photo model and was Playboy's Playmate of the Year 1981. She started her own Web site in June 1997 and used terms such as "Playmate of the Year" and "Playboy" in her site's pages. She placed a disclaimer at the bottom of her pages that stated that the site is not affiliated with Playboy. Playboy sued Welles for trademark infringement and dilution. A lower court denied Playboy's motion for a preliminary injunction. The court found that Welles made fair use of the trademark because they were used in good faith and accurately described her Web site. The court viewed the use of a disclaimer in her favor.<sup>6</sup> An appellate court confirmed the lower court's decision. Kuester and Nieves recommend that

"one preventative measure against a trademark or copyright claim is a common disclaimer. The disclaimer should state that the Web site and owner of the Web site are not sponsored or affiliated with any owners of the specific trademarks used on the Web site" [20].

The protection of a trademark can be limited by the First Amendment (free speech). In general, contents of a non-commercial nature (e.g., news coverage) has a stronger free speech protection than commercial contents such as the one in *Playboy vs. Welles*. However, courts have disagreed what constitutes noncommercial speech [21].

**Meta Tags.** HTML documents allow the author to specify information about the document in addition to the actual content. Such information, called meta data, is not shown by the Web browser and consists of key-value pairs.<sup>7</sup> The `keywords` attribute can be used to quickly identify and summarize the contents of a page. For example, a Web page that sells Mikimoto jewelry could use the following keywords:

```
<META name="keywords"
content="Mikimoto, jewelry">
```

Search engines can use this information to efficiently index a page without scanning the page's contents. However, keywords can be used to mislead a search engine in the sense that the given keywords might not correspond to the actual content. However, it should also be mentioned that this issue might become moot, because sophisticated search engines simply ignore meta tags. Meta tags can infringe on trademark law if the `contents` attribute contains

<sup>6</sup>Ironically, the disclaimer might have helped to give the site a higher ranking for searches that contained the trademarks used in the disclaimer.

<sup>7</sup>The HTML standard itself does not specify the keys. Instead, profiles are used to specify valid keys and their meaning.

terms that are trademarked. In the above example, the use of Mikimoto can be a trademark violation.

In *Brookfield vs. West Coast Entertainment*, the plaintiff held the trademark "MovieBuff," which was used by the defendant in a domain name (`moviebuff.com`) and in the Web site's meta tags. The court concluded that the defendant used the trademark to attract customers to his Web site: "Using another's trademark in one's metatags is much like posting a sign with another's trademark in front of one's store." The court saw this as a trademark infringement under the doctrine of initial interest confusion [30]. Interestingly, the court said that the defendant could use the term "movie buff" in meta tags, because it is a legitimate word in the English language (dictionary doctrine [21]).

### 3. AN IP MATURITY MODEL

Huang and Tilley propose a documentation maturity model (DMM) akin to the Capability Maturity Model (CMM) for software [16]. While the DMM is specifically targeted towards technical documentation for program understanding, its general framework can be used for Web site contents as well. The DMM assesses the quality of documentation using five maturity levels. The maturity levels are used to characterize product quality as well as process maturity.

Since the IP aspect of the contents published in Web sites is increasingly important to many site owners, we propose to extend the DMM to capture the maturity of a Web site with respect to IP. In this paper, we discuss the product quality part of the DMM. This allows an assessment of the current state of the Web site's IP maturity, but ignores the process by which the maturity level has been achieved.<sup>8</sup>

#### 3.1. Web Site Product Quality Levels

To assess product quality, the DMM defines several product attributes along with the required characteristics for each maturity level. For example, graphics formatting ranges from (1) static and informal (e.g., GIF images) to (5) editable (i.e., live documents [43]). The levels are inclusive, meaning that a certain maturity level satisfies the characteristics of all lower levels.

We propose the following levels to assess a Web site's maturity regarding IP:

**Level 1 – None:** The Web site has no copyright notices, relying on the implicit protection that copyright provides. Similarly, trademarks (ones own and other's) are not acknowledged.

<sup>8</sup>The details of the DMM's process maturity levels have not been defined yet.

**Level 2 – Informal:** Copyright notices are present at some Web pages or specific works accessible via the Web page, but not applied consistently. (Instead of copyright notices, licenses such as open content can be used as well.) Similarly, Trademarks are only partially acknowledged.

**Level 3 – Copyright Notice:** Every Web page has a valid copyright or license notice (copyright symbol, followed by year of first publication and the copyright owner's name). The notice is typically placed at the footer of the page. However, if the page contains content of different copyright owners, individual notices are necessary. In order to achieve consistence and complete of the notices, they can be generated (semi-)automatically. Optionally, the Web site is registered with the Copyright Office. Trademarks that the Web site publishers are aware of are acknowledged. Every trademark's occurrence can be separately acknowledged or the first or prominent occurrence of the trademark. There can be an additional notice along the lines that "other company, product and service names may be trademarks of others."

**Level 4 – IP Agreements:** The Web site contains an agreement that identifies its IP and possibly the IP of others. It also governs acceptable behavior by users of the site regarding the site's IP. Typically, IP agreements are contained in a dedicated Terms of Use page that is linked to at the footer of every page. Such a page constitutes a *browse-wrap* agreement between the owner and user of the site [18]. Since the user of the site is typically not required to read the Terms of Use page and also does not explicitly accept it (e.g., by clicking on an "I agree" button), commentators have questioned the enforceability of such a contract [10]. However, courts have increasingly broadened their view of what constitutes an enforceable contract [19]. Part of the IP notice can be other statements, such as link policies and provisions against reverse engineering.

**Level 5 – Comprehensive Agreement:** This level is beyond the scope of the paper because a full agreement between the owner and user of the site addresses issues unrelated to IP ranging from the site's jurisdiction [36] and disclaimer of liability to privacy statement and acceptable user behavior. In general, the agreement has to be comprehensive in the sense that all potential legal issues are addressed. In order to use the full functionality of a Web site (e.g., activities that go beyond basic browsing such as ordering of products), the user typically has to register and explicitly accept a click-wrap contract.

The introduced maturity levels are a simple tool for Web

site owners to assess their IP maturity and are a first step towards a more formal risk analysis.

## 4. SUMMARY AND DISCUSSION

This paper has explored intellectual property issues on the Web by discussing case law in copyrights, trademarks, and patents. Web site owners and content providers need to know about legal aspects that potentially affect their activities for two reasons. First, they have to know how to manage and protect their own IP. Second, they must respect the IP of others in order to minimize legal risks.

Large commercial Web sites are aware that IP is an important issue for their business and hence have taken steps to protect it. Such sites typically have achieved Level 5 in our quality model. For example, the Amazon Web site ([www.amazon.com](http://www.amazon.com)) has

- the copyright notice "©1996-2004, Amazon.com, Inc. or its affiliates" at the very bottom of every page.
- a Conditions of Use page that contains the following IP agreements:
  - an explicit notice regarding copyright that "all content included on this site, such as text, graphics, . . . , and software, is the property of Amazon.com . . . and protected by United States and international copyright laws."
  - a list of its trademarks and trade dress.
  - a statement of its patents, among them the infamous 1-click patent: "One or more patents apply to this Site and to the features and services accessible via the Site, including without limitation: U.S. Patent Nos. 5,715,399; 5,960,411; . . ."
- a comprehensive coverage of legal issues, including a Privacy Notice page.

The above evidence from the Amazon site shows that great care has been taken to protect the site's IP and other legal issues. Note that we could assess Amazon's quality level from the outside by observing their site; in contrast, assessing the process maturity would require access to the company's internals.

While it is difficult to give clear guidelines for content publishers due to evolving case law and diverging jurisdictions, we have highlighted the areas of IP that have raised legal issues in the past. Even though few Web sites have access to professional legal advice, they can still assess and manage IP issues themselves. As a first step to a more formal risk assessment, we have introduced Web site IP maturity levels, which allow to assess a Web site's coverage of IP issues. Web site owners can use the maturity levels as a guidance to improve their own site's quality with respect to IP.



## References

- [1] Court upholds \$521 million jury verdict against microsoft for patent infringement. *Packets Cyberlaw Newsletter*, 1(8), 2004. [http://cyberlaw.stanford.edu/packets/vol\\_1\\_no\\_8/001918.shtml](http://cyberlaw.stanford.edu/packets/vol_1_no_8/001918.shtml).
- [2] Misuse of password found not to violate digital millennium copyright act. *Packets Cyberlaw Newsletter*, 1(13), 2004. [http://cyberlaw.stanford.edu/packets/vol\\_1\\_no\\_13/002120.shtml](http://cyberlaw.stanford.edu/packets/vol_1_no_13/002120.shtml).
- [3] Apple. Apple licenses amazon.com 1-click patent and trademark. Sept. 2000. <http://www.apple.com/pr/library/2000/sep/18amazon.html>.
- [4] R. August. *International Business Law: Text Cases, and Readings*. Prentice Hall, 4th edition, Feb. 2003.
- [5] T. Berners-Lee. Links and law. *World Wide Web Consortium*, 1997. <http://www.w3.org/DesignIssues/LinkLaw.html>.
- [6] K. W. Bowyer. *Ethics and Computing: Living Responsibility in a Computerized World*. IEEE Computer Society Press, 1996.
- [7] Copyright Office. Circular 66: Copyright registration for online works. <http://www.loc.gov.org/copyright/circs/circ66.html>.
- [8] J. P. Costello. Copyright and work made for hire. *IEEE Software*, 11(3):93–94, May 1994.
- [9] Creative Commons. <http://creativecommons.org/>.
- [10] S. J. Davidson, S. J. Bergs, and M. Kapsner. Open, click, download, send...what have you agreed to? The possibilities seem endless. *Computer Associations Law Conference (CLA 2001)*, 2001. <http://www.cla.org/click%20contract.pdf>.
- [11] I. J. Garrote. Linking, framing, and copyright: A comparative law approach. *SSRN Electronic Paper Collection*, 2001. <http://ssrn.com/abstract=280596>.
- [12] GNU Free Documentation License. <http://www.gnu.org/copyleft/fdl.html>.
- [13] J. Grove. Wanted: Public policies that foster creation of knowledge. *Communications of the ACM*, 47(5):23–25, May 2004.
- [14] D. L. Hayes. Advanced copyright issues on the internet. Feb. 2004. [http://www.fenwick.com/docstore/355/Advanced\\_Copyright\\_02-29-04.pdf](http://www.fenwick.com/docstore/355/Advanced_Copyright_02-29-04.pdf).
- [15] E. Henry W. (Hank) Jones, III. Developing and distributing hypertext tools: Legal inputs and parameters. *Hypertext '87*, pages 367–374, Nov. 1987.
- [16] S. Huang and S. Tilley. Towards a documentation maturity model. *21st Annual International Conference on Design of Communication (SIGDOC '03)*, pages 93–99, Oct. 2003.
- [17] P. B. Hugenholtz. The new database right: Early case law from europe. *Ninth Annual Conference on International IP Law & Policy*, Apr. 2001.
- [18] D. Isenberg. *GigaLaw Guide to Internet Law*. Random House, Oct. 2002.
- [19] H. M. Kienle, D. German, and H. A. Müller. Legal concerns of web site reverse engineering. *6th International Workshop on Web Site Evolution (WSE 2004)*, Sept. 2004.
- [20] J. R. Kuester and P. A. Nieves. Hyperlinks, frames and meta-tags: An intellectual property analysis. *IDEA: The Journal of Law and Technology*, 38:243ff, 1998.
- [21] F. G. Lastowka. Search engines, html, and trademarks: What's the meta for? *Virginia Law Review*, 86:835–884, 2000.
- [22] L. Lessig. *Code, and Other Laws of Cyberspace*. Basic Books, 1999.
- [23] L. Lessig. *The Future of Ideas: The Fate of the Commons in a Connected World*. Random House, 2001.
- [24] S. M. Maurer, P. B. Hugenholtz, and H. J. Onsrud. Europe's database experiment. *Science*, 294:789–790, Oct. 2001.
- [25] M. D. Medley, R. Rutherford, R. W. Roth, G. E. Anderson, and S. A. Varden. Ethical issues related to internet development and research. *ITiCSE Working Group Papers*, 30(4):61b–76b, Dec. 1998.
- [26] T. H. Nelson. Xanalogical structure, needed now more than ever: Parallel documents, deep links to content, deep versioning, and deep re-use. *ACM Computing Surveys*, 31(4es), Dec. 2000.
- [27] M. B. Nimmer and D. Nimmer. *Nimmer on Copyright*. Matthew Bender & Company, 2002.
- [28] M. O'Rourke. Is virtual trespass an apt analogy? *Communications of the ACM*, 44(2):98–103, Feb. 2001.
- [29] Panel. Panel discussion: Link law on the internet: A panel discussion. *IDEA: The Journal of Law and Technology*, 38:197ff, 1998.
- [30] S. U. Paylago. Search engine manipulation: Creative use of metatags or trademark infringement? *IDEA: The Journal of Law and Technology*, 40:451ff, 2000.
- [31] M. Sableman. Link law revisited: Internet linking law at five years. *Berkeley Technology Law Journal*, 16(3), 2001.
- [32] P. Samuelson. Intellectual property rights for digital library and hypertext publishing systems: An analysis of Xanadu. *Third annual ACM conference on Hypertext (Hypertext '91)*, pages 39–50, Dec. 1991.
- [33] P. Samuelson. Computer programs and copyright's fair use doctrine. *Communications of the ACM*, 36(9):19–25, Sept. 1993.
- [34] P. Samuelson. Copyright's fair use doctrine and digital data. *Communications of the ACM*, 37(1):21–27, Jan. 1994.
- [35] P. Samuelson. Unsolicited communications as trespass? *Communications of the ACM*, 46(10):15–20, Oct. 2003.
- [36] K. E. Smith. Trademark disputes in cyberspace. *Mixed Media*, pages 33–38, Feb./Mar. 1998.
- [37] R. A. Spinello. An ethical evaluation of web site linking. *Computers and Society*, pages 25–32, Dec. 2000.
- [38] R. H. Stern. Glitches left in software copyright system. *IEEE Micro*, 13(3):75–77, June 1993.
- [39] R. H. Stern. Content providers: "I was framed". *IEEE Micro*, 17(3):7–9, 76–79, May/June 1997.
- [40] R. H. Stern. Amazon's one-click patent loses its teeth. *IEEE Micro*, 21(1):7–10, Jan./Feb. 2001.
- [41] R. H. Stern. Challenging search engines under copyright law: Part 1. *IEEE Micro*, 22(3):6–7, May/June 2002.
- [42] R. H. Stern. Challenging search engines and pop-ups under copyright law: Part 2. *IEEE Micro*, 24(1):7, 71–73, Jan./Feb. 2004.

- [43] A. Weber, H. M. Kienle, and H. A. Müller. Live documents with contextual, data-driven information components. *20st Annual International Conference on Documentation (SIG-DOC 2002)*, pages 236–247, Oct. 2002.
- [44] Wikipedia: The Free Encyclopedia. <http://en.wikipedia.org/>.
- [45] T. Wolverton. Amazon, Barnes&Noble settle patent suit. *CNET News.Com*, Mar. 2002. <http://news.com.com/2100-1017-854105.html>.
- [46] D. J. Wood. Best practices for avoiding linking and framing legal liability. *GigaLaw.com*, June 2002. <http://www.gigalaw.com/articles/2002-all/wood-2002-06-all.html>.