

# EXHIBIT 1

**Disney Enterprises, Inc. et al v. Hotfile Corp. et al,**  
**1:11-cv-20427-KMW (S.D. Fl.)**  
**Rebuttal Report of Professor James Boyle**

1. I am currently the William Neal Reynolds Professor of Law at Duke University, and have been retained by Farella, Braun + Martel LLP on behalf of the Defendants in this action as an expert witness.

**Background and Qualifications**

2. I received an LL.B. (Hons) from Glasgow University (1980), and an LL.M. (1981) and S.J.D. (1986) from Harvard Law School. I have been a law professor since 1982, teaching at American University, and at the Universities of Pennsylvania, Harvard and Yale as a Visiting Professor. In 2000 I joined the law faculty at Duke. My other qualifications, awards and publications were listed in my initial expert report.

3. I have not previously testified as an expert. I am being remunerated for my work as an expert in these proceedings at the rate of \$750 per hour.

4. The Documents that were used in support of my opinions are listed below under the heading "Documents reviewed".

**Scope of Expert Assignment**

5. I have been asked by Farella, Braun + Martel LLP on behalf of the Defendants to provide an expert rebuttal report to a statistical report prepared by Dr. Richard Waterman,<sup>1</sup> (The Waterman Report) on the uses of Hotfile.com. The Waterman report also includes a section (Exhibit C) by Mr. Scott Zebrak. In that section Mr. Zebrak details the methods by which he assessed the copyright status of the 1750 files in Dr. Waterman's sample. He also lists those files, together with his assessment of their legal status. I have studied both Dr. Waterman's and Mr. Zebrak's methods and am prepared to testify on my conclusions about them.

**Documents reviewed**

6. In forming my opinions, I reviewed:

- a) The Rule 26(a)(2)(B) Report of Dr. Richard Waterman and all Exhibits
- b) The Rule 26(a)(2)(B) Report of Scott Zebrak (Exhibit C to the Waterman Report), all Exhibits and database materials produced by Mr. Zebrak in a timely manner
- c) The November 29, 2011 Transcript of the Deposition of Richard Waterman

---

<sup>1</sup> RULE 26(a)(2)(B) REPORT OF DR. RICHARD WATERMAN.

- d) The December 20th, 2011 Transcript of the Deposition of Scott Zebrak
- e) Prior testimony and reports of Dr. Waterman in other copyright matters attached as Exhibit A
- f) Elysium Digital's technical analyses of aspects of the Hotfile database, software questions, Internet issues, and the hard drive and databases provided by Mr. Zebrak. See Elysium's analysis summaries attached as Exhibit B, hereto.
- g) Sample counter-notices received by Hotfile attached as Exhibit C, hereto.
- h) Declaration of Charles J. Hausmann in Support of Plaintiff's Motion for Summary Judgment (*Grokster*), attached as Exhibit D, hereto.
- i) Case law, offline and online articles and websites, as identified below.
- j) Affidavit of Scott Wittenburg and Elysium Analysis of a Photography Podcast, attached as Exhibit E, hereto
- k) Email from Legal Counsel of Opera Software, attached as Exhibit F, hereto.
- l) Email and affidavit from Marc Schwegler from Farm Simulator / Giants Software and End User License Agreement, attached as Exhibit G, hereto.
- m) DirectX End User Licenses and printouts re: DirectX attached as Exhibit H
- n) Russian Book regarding weaving and embroidery from 1871 attached as Exhibit I

7. For the remainder of this report, I will focus on the quantitative picture that Dr. Waterman's report paints of Hotfile. I by no means agree, however, that the quantitative picture is the only relevant one, and I reserve my right, if called to testify, to comment on *qualitatively* important non-infringing uses of the Hotfile system. As an example, of what I mean by a qualitatively important non-infringing use I would point to the following incident. MIT's *Technology Review* recently published an article dealing with the role of digital services in the democratic uprisings collectively referred to as the Arab Spring.<sup>2</sup> The article recounts that one of the very important catalysts for the democratic demonstrations was a gory video of a hospital emergency room in Kasserine, Tunisia, dealing with individuals who had been beaten by the police. Denied access to other online services, one of the protest movements (Takriz) "smuggled a CD of the video over the Algerian border and streamed it via MegaUpload."<sup>3</sup> Al Jazeera picked up the video because of its exposure on MegaUpload and the excerpts showed on television catalyzed a wave of pro-democracy protests. Upon investigating this, I found that MegaUpload – like Hotfile – is a cyberlocker

---

<sup>2</sup> John Pollock, *Streetbook: How Egyptian and Tunisian Youth Hacked the Arab Spring* TECHNOLOGY REVIEW (Sept-Oct 2011) <http://www.technologyreview.com/web/38379/>

<sup>3</sup> *Id.*

site. (Interestingly, two hash-identical versions of the same video can be found on Hotfile, uploaded on Jan 11<sup>th</sup> 2011. Those versions were downloaded 21 times in January of 2011.)<sup>4</sup>

8. The importance of the site design here is that there was no approval required for posting, nor any editorial screening for what – in this case – was extremely disturbing, but nevertheless important material. In any quantitative study of a service like Hotfile, the video would count as a single non-infringing file. In terms of the *qualitatively* important non-infringing uses, a story like the Arab Spring one reveals the importance of open communication networks to free speech and First Amendment values in a way that transcends a single entry in an Excel spreadsheet quantifying infringement. In a final assessment, I presume that a court would want also to look at those qualitatively important non-infringing uses. In my remaining comments, however, I shall focus only on Dr. Waterman’s quantitative study and the flaws I found within it.

9. Dr. Waterman’s statistical review of Hotfile paints the following picture:

Based upon my review of the most recent data provided by Mr. Zebrak, approximately 90.3% of all daily downloads of files on Hotfile were downloads of infringing or highly likely infringing content; approximately 5.4% of the downloads of files per day on Hotfile were downloads of non-infringing or highly likely non-infringing files; and the remaining approximately 4.3% of the downloads of files per day on Hotfile were downloads of files whose copyright status could not be reliably determined in the time allowed.<sup>5</sup>

10. Dr. Waterman obtained this statistical snapshot by a procedure that includes several steps that deserve the court’s critical attention. I am not a statistician and cannot opine as to whether Dr. Waterman’s random number generator was properly calibrated. However, a key part of Dr. Waterman’s method is the choice of what files to exclude from the study, and how to weight those that remain. That choice – at least if the study is to be legally relevant to this trial – is one that is profoundly shaped by the law. With Dr. Waterman’s and Mr. Zebrak’s testimony, the plaintiffs are presumably attempting to provide the court with factual information relevant to the legal determination of

a.) whether Hotfile is a service with “substantial non-infringing uses” under *Sony*

and

b.) whether Hotfile is guilty of *Grokster*-style inducement liability.

11. In my opinion as a legal scholar, the method they have chosen to use has several fundamental flaws that cause it to present a misleading answer to both of those questions. In particular, by focusing purely on downloads, Dr. Waterman’s method entirely excludes one important use of the Hotfile system, a use that appears to be clearly non-infringing

---

<sup>4</sup> See Exhibit B, Massacre at Kasserine.

<sup>5</sup> RULE 26(a)(2)(B) REPORT OF DR. RICHARD WATERMAN, paragraph 5.

under *Sony* and which is obviously relevant to any analysis of inducement: namely, temporary personal storage and archival backup. A statistical analysis of the use of VCR's in the *Sony* case that, because of its design, implicitly excluded the time-shifting of TV programs from its analysis of VCR uses would paint a legally misleading picture. A court could not rely on such a study in making an assessment of contributory or vicarious liability, or in assessing whether there were substantial non-infringing uses. The same would appear to be true here. In reviewing District Court findings on substantial non-infringing uses, Courts of Appeal have made the rigorous requirements of such an inquiry very clear.<sup>6</sup> This study does not appear to satisfy those requirements.

12. My objections are grouped into three parts. The first is to Dr. Waterman's method as a general matter. The second is to the specific application of that method or protocol to the material found on Hotfile. The third goes to decisions that Mr. Zebrak made in assessing the copyright status of the files on Hotfile. I will deal with each of them in turn.

## I

### **GENERAL FLAWS IN DR. WATERMAN'S METHODOLOGY AS APPLIED TO ANY FILE-STORAGE AND TRANSFER OR "CYBERLOCKER" SITE**

13. To make clear the problems with Dr. Waterman's methodology it may be instructive first to imagine it being applied to an entirely hypothetical cyberlocker and file transfer site called Example.com. Example.com has 10,000 users. 9,900 of them use the site for storage and back up. Such users upload documents on which they are working, such as the PowerPoint files they use for work purposes. Since the users do not choose to share the URL's with others, and since Example.com does not provide a file listing search feature or allow other search engines to index content that is not linked to on the open web, those files are relatively inaccessible to anyone but the uploader. An average of 10 files is uploaded by each user. So long as no disaster occurs – the document does not get corrupted, or the folder does not get mistakenly deleted – they will never need to download those files and thus, the file will register zero downloads. Example.com's business model is to encourage these users to purchase the premium subscription by removing any content that has not been downloaded for 3 months. The premium subscription to Example.com

---

<sup>6</sup> As I pointed out in my initial Report, the Courts of Appeal have disapproved of District Court assessments of substantial non-infringing use on the basis of far more subtle mistakes, such as a focus only on current use rather than potential uses. "We depart from the reasoning of the district court that Napster failed to demonstrate that its system is capable of commercially significant noninfringing uses. *The district court improperly confined the use analysis to current uses, ignoring the system's capabilities.* Consequently, the district court placed undue weight on the proportion of current infringing use as compared to current and future noninfringing use." *A&M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004,1021 (9th Cir. 2001) (emphasis added.) To omit from one's statistical sample the method of usage most characteristic of a cyberlocker site – namely zero download storage – which is a direct analog to the substantial non-infringing uses that carried the day in *Sony v. Universal* - is an error of an altogether more obvious and fundamental type.

removes this limitation and allows storage for an unlimited period of time.

14. The remaining 100 users of Example.com use the system for file transfer. 10 use it for “space shifting” commercial content they have purchased so that they can watch it at another location when they are away from their home computers. (So long as they are space shifting their own content to themselves, this is a practice that is very probably a fair use.) Less salubriously, 70 use it for “sharing” their favorite pornography. Of those 70, 35 create edited excerpts featuring their favorite performers or scenes (raising a complex issue of legal analysis about whether there is a fair use, one that would depend on the substantiality of the portion used, the degree of transformation and the market for short form edited versions of pornography.) 35 simply copy the entire pornographic video file. (This would be infringing *unless* the pornographer gives express or implied license to distribute the pornographic video files to the web, perhaps to drive content towards a particular site whose watermark appears on the film.) 10 users utilize Example.com to share full length, commercial, (non-pornographic) copyrighted major studio films with the world. This use is infringing and provides 10% of the total downloads on the site. The total number of downloads combining the space shifters, the remixing and sharing pornography fans and the users illicitly copying commercial feature films is 90,000.

15. Finally, the last 10 users use the system to share open source software that they themselves have written and in which they hold the copyright. This is a popular use of Example.com and in fact includes the two most downloaded files on the system. (This is clearly a non-infringing use and many scholars, including me, would claim that this, by itself and without regard to any of the other clearly licit uses of the site, satisfies the *Sony* standard of a substantial non-infringing use.) There are a total of 10,000 downloads of the open source software.

16. As I understand Dr. Waterman and Mr. Zebrak’s methodology, they would classify the uses of Example.com as “90% infringing.” First Dr. Waterman’s methodology by focusing only on *downloads*, implicitly excludes the 9,900 users who utilize the site for storage and back up. Their 99,000 uploads have no downloads. This leaves him with a universe of 100,000 downloads. Based upon my review of Mr. Zebrak’s report and deposition transcript, it appears likely that he would classify all but the open source software downloads as infringing. If true, their conclusion would be that 90% of the uses of Example.com are infringing though the reality is very different. In fact, more than 99% of the *users* of Example.com are *not* infringing. More than half of the *uses* of the system – both uploads and downloads – are clearly non-infringing. And a significant percentage of the *downloads* on the system are either debatably a fair use, authorized by implied license or clearly non-infringing.

17. i.) Percentage of users, ii.) of uses and iii.) of uploads *and* downloads; these are all pieces of evidence that courts would presumably need in the process of determining whether services have a substantial non-infringing use – and given that *Sony* instructs courts not to look at *predominant* use, but rather current and potential substantial non-infringing uses, that evidence presumably needs to be comprehensive. Those same factors are also relevant to the multi-factor assessment of inducement liability that the Supreme

Court laid out in *Grokster*.

18. In short, there are crucial omissions in the universe of uses and users that Dr. Waterman's method captures. As a result in my opinion, his method – if used as *the* statistical snapshot on which a contributory, vicarious, or inducement liability assessment were to be carried out – would yield a legally misleading conclusion when applied to a cyberlocker and file transfer site.

19. I wish to stress that my claim is not that Example.com is Hotfile, though there are some obvious similarities. My claim is that the Example.com hypothetical shows why Dr. Waterman's method – as a general matter, not just in the case of Hotfile – will present a legally misleading picture of the facts about *any* cyberlocker/file-transfer site. I will now turn to his analysis of Hotfile in order to show in more detail the problems caused by the methodological choices he has made.

## II

### SPECIFIC FLAWS IN DR. WATERMAN'S METHODOLOGY AS APPLIED TO HOTFILE

#### i.) Files – And Types of Use – Excluded from Study

20. First, and vitally, by focusing only on downloads, Dr. Waterman excludes all files that have zero downloads from Mr. Zebrak's analysis of infringement. Working under my direction, the computer consulting company Elysium Digital examined the Hotfile database in order to discover how many files had zero downloads. They reported that, out of a total of 107,271,438 total files stored on the Hotfile system 57,923,301, or 54%, had no registered downloads. Thus in the case of Hotfile, Dr. Waterman's study actually excludes a *majority* of the files on the system.

21. Were many of those 57,923,301 files in fact being uploaded to Hotfile.com for file storage? That is something that neither Dr. Waterman, nor Mr. Zebrak nor I actually know because – by design – those files have been excluded from their statistical assessment of the uses of the system. Hotfile clearly can be and surely is used for file storage. Both Hotfile's architecture and its business model are consistent with it, particularly Hotfile's policy of capping (free) zero download storage at 3 months (14 days for anonymous users), while allowing unlimited storage time for Premium users. Offering a free "teaser" service that attracts users to a more feature-rich fee-paying premium service is such a standard business method on the Internet that it has attracted its own neologism: "freemium."<sup>7</sup> Further, given that Hotfile itself has no index to the files and the choice whether to share

---

<sup>7</sup> Nicolas Pujol, *Freemium: Attributes of an Emerging Business Model*  
[http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1718663](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1718663)

"Freemium is a business model that works by offering a product or service free of charge (typically digital offerings such as software, content, games, web services or other) while charging a premium for advanced features, functionality, or related products and services. The word 'freemium' is a portmanteau combining the two aspects of the business model: 'free' and 'premium'." <http://en.wikipedia.org/wiki/Freemium> [Last visited Dec 18, 2011]

the direct URL is the user's, the system appears well-suited to storage of a wide range of material<sup>8</sup> – such as a large PowerPoint file, for example.<sup>9</sup>

22. A user could store such a file on Hotfile, intending only to retrieve it personally if necessary, but would have the option of giving out the URL if subsequently she decided to share it with colleagues, who would then be able to access it without being given a personal password. As no one but the user has the URL to the file, and it is not indexed by search engines, the file is effectively private – yet the user can at any time share the file with colleagues or co-workers simply by giving them the URL. The Google search referred to in note 9 found more than 45,000 *publicly* listed PowerPoint files on Hotfile – that is, PowerPoint files that users have chosen to link to on the open web. Those files are presumably being shared – after a conference say. But a user can also use the system for storage or space shifting. Acting at my direction Elysium Digital found that there were more than 40,000 PowerPoint files on Hotfile, that have been downloaded either zero or one times. And of course, PowerPoint files are only one example of this kind of storage. A counter notice issued in response to an apparently faulty 'notice and takedown' request, for example, reveals that an architecture company was apparently using Hotfile to store drawings of the designs it created for clients.<sup>10</sup> One can imagine many other such examples.

23. One reason the plaintiffs have suggested that Hotfile is not used for storage is the absence of password protection on the files. The implication is that no one would store on a cyberlocker unless the file was protected by a password. However, once one understands the architecture of Hotfile, this particular objection is completely unconvincing, in my opinion. Files stored on Hotfile, if the user does not reveal or post the URL, are actually considerably *more* secure than files stored on common types of password-protected online storage. Consider files that are stored on a user's email or iTunes account. An outsider who wished to get access to that account and see the material would need to provide a username and password to do so. In both these cases, however, the username is the person's email address. Anyone who has had an e-mail from me or who has seen my e-mail posted on my website already has the username. Now the password alone protects the

---

<sup>8</sup> Hotfile URLs include the ID of the content and a randomly generated number. And the result is sufficiently long and complex as to be highly, highly unlikely for any other person to stumble upon by accident – actually more unlikely, as I will explain in a moment, than guessing a password on many typical forms of email or online storage. Hotfile does not index files. The large search engines such as Google only index a file on Hotfile if a user has chosen to publicly post the URL somewhere on the open web. If the user chooses not to do that, the file effectively cannot be accessed without the user's consent – the filename could not be discovered in any way.

<sup>9</sup> A search on Google on December 29<sup>th</sup> 2011 for “.ppt OR .pptx site:hotfile.com” (i.e. files with the PowerPoint file extensions .ppt or .pptx on the Hotfile site) returned 45,800 hits. These are the PowerPoint files on Hotfile that *have* had their URL's posted publicly. There are presumably more that have not had their URL's posted publicly and which could not be found without the storing user's consent.

<sup>10</sup> See Exhibit C.



content. Password rules vary. Assume here that the password can be composed of any number and any lowercase letter. If the password is a 7 character alphanumeric, longer than most passwords, the chances of “brute forcing” that password – that is of obtaining the password by random computerized guessing is 1 in 78 billion. Assuming 10 efforts a second, it would take a “brute force” attack (i.e. one that simply tries every different combination of letters and numbers) 248 years to gain the password and get access to the stored material. That certainly provides security. But how does it compare to a file posted to Hotfile where the user keeps the URL and never shares it with anyone? (Search engines do not index Hotfile files unless the user posts the link elsewhere on the open web.) An outsider who knows I have stored a file on Hotfile, but does not know the URL will need to guess the URL in order to get access to the content. He knows that the URL begins <http://www.hotfile.com> of course, but nothing else. A Hotfile URL is composed of two parts, a numerical upload ID and a second 7 character identifier. Together, they make up the URL. For example, <http://hotfile.com/dl/97361133/4bc1eqz/>. In order to guess the 7 character identifier, which is also composed of any number and any lowercase letter, the outsider would face the same odds as the person guessing my password – 1 in 78 billion. But *in addition* he would also need to guess the upload ID. In other words it is actually harder to get access to the URL of a particular Hotfile file than to get access to a typical kind of online password protected storage.

24. The fact that 57,923,301, or 54%, of the files on Hotfile have no downloads suggests that users are employing the system for something other than file transfer. Users who rely on Hotfile for temporary storage will most likely have zero downloads. Certainly many of them will. By excluding this central, and very probably legal, use, Dr. Waterman’s method, in my opinion, presents a legally misleading picture of Hotfile. I would note that Dr. Waterman’s testimony<sup>11</sup> in prior cases of alleged contributory and vicarious copyright infringement included a different statistical method as well as a download study – a study of files that were “made available,” that is, that were uploaded to the system. That was in the context of a peer-to-peer system where the possibility of storage effectively did not exist. Yet there, Dr. Waterman’s study effectively had two parts; one focused on the act of uploading and the other that of downloading. Had some variant of that technique been included here, in addition to the study of downloads, Dr. Waterman’s statistical picture could have included the storage function of Hotfile and Mr. Zebrak would have had to assess the legality of such storage. Dr. Waterman’s statistical analysis would thus not have neglected the possibility that Hotfile.com, the cyberlocker and file transfer site, was indeed being used as a cyberlocker. The method he uses here does neglect that possibility. In fact, he states in his deposition that, in this case, he was instructed by plaintiffs’ counsel to look only at downloads.<sup>12</sup> In my opinion, this is clearly an error.

25. At my direction, Elysium Digital examined the Hotfile database and found that an additional 6,182,360, or 5.76%, of the files on Hotfile have only one registered download, a number of downloads consistent with both storage and space-shifting – potentially licit uses. The ‘one download’ files do appear in Dr. Waterman’s sample, but they are given a

---

<sup>11</sup> See for example Exhibit A; Usenet Declaration paragraph 5.

<sup>12</sup> Waterman Depo. p. 212.

reduced weight relative to those that are downloaded more frequently.

Within each selected day, the sample frame was obtained by taking the dailydownload data and expanding the record of each file to capture the total number of recorded downloads of that file on that day. For example, if a file was downloaded 5 times in a day, the record would be expanded to reflect five separate downloads of that file.<sup>13</sup>

26. This method has a striking result. Imagine that there were to be only 10 files on the Hotfile system – eight an example of legal (no download) storage, one an example of legal (one download) space shifting of licitly purchased commercial content and one a commercial film that was illicitly uploaded and was then downloaded nine times. The eight files that were not downloaded would be ignored by the study, the file that was downloaded once would appear a single time, and the file that was illicitly downloaded would appear nine times. As a result, Dr. Waterman would classify the system as having at least 90% illicit uses. In addition, if Mr. Zebrak assessed the legal status of the file without considering the number of times that it was downloaded, as I believe he did, he would classify the single download space-shifting file as also being illicit, despite the fact there is a very strong argument this is a fair use under section 107.<sup>14</sup> In that case, the Waterman protocol would describe the system as 100% infringing.

27. A file downloaded nine times appears nine times in the total listing, in order to identify the relative percentage of illicit *downloads*. But if this is extrapolated into an assessment that 90% of the *uses* of the system are illicit, the conclusion becomes unsupportable. *Sony* directs courts to look at *types of uses* in assessing a system or product. It also rejects the conclusion that a system be classified as legal or illegal based on its predominant use. Thus, any study that merely includes statistical assessment of downloads, if not accompanied by other statistical surveys that include the zero download files, will fail to provide an assessment on which a court applying *Sony's* standard can rely. In this case, the focus on downloads alone actually excludes a majority of the files on the system from Mr.

---

<sup>13</sup> RULE 26(a)(2)(B) REPORT OF DR. RICHARD WATERMAN, paragraph 12

<sup>14</sup> See for example the explicit endorsement of such a position in the *Diamond* case. “The Rio merely makes copies *in order to render portable, or “space-shift,” those files that already reside on a user’s hard drive. Cf. Sony Corp. of America v. Universal City Studios*, 464 U.S. 417, 455 (1984) (holding that “time-shifting” of copyrighted television shows with VCR’s constitutes fair use under the Copyright Act, and thus is not an infringement). Such copying is paradigmatic noncommercial personal use entirely consistent with the purposes of the Act.” *Recording Indus. Ass’n of Am. v. Diamond Multimedia Sys., Inc.*, 180 F.3d 1072, 1079 (9th Cir. 1999). [Emphasis added.] Subsequent cases in the peer-to-peer context have cast doubt on whether this finding would hold true in a situation where a user sought to i.) claim fair use privileged access on a peer-to-peer network to *someone else’s* copy of a copyrighted work that the user himself had purchased, ii.) if that copy was being shared with the entire world. But in the context of a zero or one download storage or space shifting on a cyberlocker neither of those other factors obtains and *Diamond’s* premise would therefore strongly suggest fair use.

Zebrak's review and ignores a type of use that would clearly qualify as an actual current, and potential future, substantial non-infringing use.

## ii.) Questionable Decision to Include Pornographic Files

28. Dr. Waterman made the decision to include pornographic files though, as specified in his protocol, content that the Jenner and Block team classified as illegal or child pornography was removed from the database. Not all prior empirical studies in cases of alleged contributory, vicarious or inducement liability included pornographic files in the empirical assessments of copyright infringement. In the *Grokster* case, for example, all pornographic content appears to have been deliberately omitted.<sup>15</sup> But both Mr. Zebrak's description of his protocol,<sup>16</sup> and the plethora of content with tasteful titles such as "Wreck My Asian Virgin A\*\*" or "Big Wet T\*\*s # 10" in the Waterman study show that the opposite decision was made in this case. No explanation was given for that different methodology. Mr. Zebrak then proceeded to find the vast majority of that pornographic content "highly likely infringing." It is remarkable how many of the files listed in Dr. Waterman's study have salacious or disgusting file names, particularly in contrast to the relatively smaller percentage of the sample that actually contains verified studio content, that is to say, content in which the plaintiffs might actually have any copyright interest.

29. The impact of the decision to include pornographic works is significant. For example,

---

<sup>15</sup> The excluded "pornography" category in the *Grokster* study covered all pornography, not merely illegal (and particularly) child pornography, meaning that a much wider category of files was excluded from the infringement study. Dr. Hausman's report describes the classification as "'Porn,' meaning that the file was plainly pornographic, including files that, from their metadata, appeared clearly to constitute illegal pornography (e.g., child porn, etc.)" Declaration of Charles J. Hausman in Support of Plaintiffs' Motions for Summary Judgment at paragraph 22[Emphasis added.] Dr. Hausman's report is also clear that these files were then excluded from the study. "Once works were assigned to a particular category, spoofs, porn, junk/damaged/unintelligible, virus/malicious, KPL, and illegal files were removed from the sample per the protocol established by Professor Olkin, and the first 1,800 files obtained through Kazaa and through Morpheus (3,600 total) that fit one of the confirmed infringing/noninfringing; highly likely infringing/ noninfringing; or unknowable categories were analyzed for copyright infringement." *Id.* at paragraph 23. Thus Dr. Hausman excluded *all* pornography, illegal or not.

<sup>16</sup> Mr. Zebrak and Dr. Waterman, by contrast to Dr. Hausman, only excluded illegal pornography. "I understand that Dr. Waterman's protocol calls for exclusion of any file that, by its metadata, appears to contain child pornography or other illegal pornography, before the files are requested from Hotfile. Consistent with that approach, and in consultation with Dr. Waterman, I excluded any sample file from the study that, upon further review, I believed might likely contain child or other illegal pornography. All of these files were replaced in the sample set of 1750 files that I reviewed with another randomly selected file per Dr. Waterman's pre-established protocol." RULE 26(a)(2)(B) REPORT OF MR SCOTT ZEBRAK, paragraph 7. I could find no explanation for the variance in method from the Hausman study.

of the first 100 files in the Zebrak study, 25 seemed by their titles<sup>17</sup> likely to have pornographic content. Of those, Mr. Zebrak counted 22 as “Highly Likely Infringing” 2 as “Non-infringing” and one as “Child Pornography.” (The latter one being the only file which would be removed from the study.) In other words, 22 of the files tagged by Mr. Zebrak as “Highly Likely Infringing” in that 100 file stretch appear likely to be pornography – approximately 25% of the files identified as infringing in that set of files. Under the protocol used by Dr. Hausman in the *Grokster* case, all of those files would have been removed from the study. By contrast, in that same 100 file sample, only nine files are listed as “Confirmed Infringing (Studio),” that is, as being content in which the plaintiffs might actually have a copyright interest. The relatively small percentage of studio content is striking.

30. Pornographers certainly can have enforceable intellectual property rights and it is doubtless commendable to see the plaintiffs looking out for their interests so assiduously here. Nevertheless, there are reasons other than tastefulness why prior studies may have chosen to omit pornographic content, and why the court here might choose to put less weight on this fraction of Dr. Waterman’s statistics and Mr. Zebrak’s determinations.

31. One reason that pornographic content may sometimes be omitted from surveys of potentially infringing works is that it is very difficult, as compared to mainstream commercial content, to assess its copyright status. Consider the task that Mr. Zebrak and his team faced, forced to spend the holiday season going through what sounds like gigabytes of porn. Thankfully, I was spared this chore, but, as a legal scholar I am at a loss to think of how I could reliably determine the copyright status of so much pornographic content in such a short time-frame. Some producers of adult films clearly do *not* intend them to be spread freely and indeed litigate their claims of copyright infringement assiduously. This is an important point, one which presumably Mr. Zebrak and Dr. Waterman considered, and it should not be overlooked. On the other hand, the scholarly literature on the economics of pornography stresses that some of it is distributed free,<sup>18</sup> using indirect methods such as advertising, or the lure of longer versions or higher quality versions on a pay site to generate revenue. Indeed articles stress that some pornographers energetically push content at viewers, even when those viewers are unwilling,<sup>19</sup> and newspaper coverage has stressed the multiple business methods that the adult film industry has been using to generate revenue.

Michael Herman, director of business development at Adult Entertainment Broadcast Network — owner of PornoTube.com, a YouTube-like site with user-generated content — says exposure on the Internet is ideal for a company’s branding. PornoTube, started nearly a year ago, generates 10 million to 15 million hits a day — making it one of the 200 most-popular sites on the Web, according to

---

<sup>17</sup> I note for clarity’s sake that neither filename nor file title is a sure indicator of the contents of a file.

<sup>18</sup> Simon Bowmaker, *Economics of Pornography* in *ECONOMICS UNCUT* 174-175 (2000).

<sup>19</sup> Jerry Ropelato, *Tricks Pornographers Play* *Internet Filter Software Review*

<http://internet-filter-review.toptenreviews.com/tricks-pornographers-play.html>

Alexa, which tracks Internet traffic. Most of PornoTube's user-generated videos are free, but clips are limited to a few minutes. Consumers who want more must pay. PornoTube partners with others to sell subscriptions to paid websites, dating services and video-on-demand. "It's become an invaluable tool for us to promote business partnerships" with adult studios, Herman says. And it's a valuable outlet for adult performers. "I can do short clips just for the Internet," says Sunny Lane, an actress in Southern California who owns Sunnyslanelive.com. "It's a way to make more money and gain more exposure."<sup>20</sup>

32. Distribution of many of the types of pornographic content I have just mentioned on Hotfile would not be illicit, at least if it were expressly or impliedly licensed as it apparently sometimes is. Then, what of non-commercially produced videos by amateur exhibitionists – who now have access to high quality digital photographic equipment? Mr. Zebrak apparently did classify as non-infringing or unknowable some works tagged as amateur content, but how can one tell where the line is? And what of the user-generated remix containing excerpts from multiple films featuring favorite performers or positions? That would present a challenging fair use analysis though not one I would choose to put in a final exam. Finally, what of adult films where the copyright owner is not known or cannot be found? The term of art “orphan works” seems particularly inappropriate when dealing with such content, but it does not seem unreasonable to believe that many pornographic production companies are – literally – fly-by-night operations, where after several years the copyright owner may not exist as a corporate entity, or may have no interest in policing the rights to its work.

33. Given the difficulties in making any, let alone all, of these assessments in an objectively reliable manner, were I designing the legal protocols for the Waterman/Zebrak study, I would have omitted pornographic content from the analysis. I wish to stress however, that Dr. Waterman's choice to include pornography, unlike the decision implicitly to exclude zero download files from review, is not necessarily by itself an error. Reasonable minds could differ about whether it should be done or not – given the nature of the content. But once that decision has been made in the affirmative, a question is raised for the court about the reliability of that particular portion of the evidence if no *confirmation* of the copyright holder's objection to the sharing of the file is obtained. My own opinion is that little weight can be put on that portion of the files in the survey, at least without certification from the pornographers in question, similar to that that Mr. Zebrak received from the studios for their commercial content, that the file is indeed “confirmed infringing.” Thus, I believe that the Waterman/Zebrak study should either have omitted pornography altogether, or included it but only classified it as infringing if there was confirmation from the copyright holder. This is both because of the difficulty of identifying with certainty the copyright status of this particular content, and because of the reality that some purveyors of pornography may not object to having their work shared, particularly if it drives traffic to a

---

<sup>20</sup> Jon Swartz, *Purveyors of Porn Scramble to Keep Up With Internet USA Today* (Updated 6/12/2007) available at [http://www.usatoday.com/tech/techinvestor/industry/2007-06-05-internet-porn\\_N.htm](http://www.usatoday.com/tech/techinvestor/industry/2007-06-05-internet-porn_N.htm) (last visited Dec 30, 2011)

particular site, or increases demand for a longer commercial version. That possible diversity of viewpoint about the desirability of sites that allow for viral distribution of copyrighted content raises an additional issue in this litigation given the disparity between the high levels of pornography found on Hotfile and the relatively low levels of confirmed studio content. In the words of the *Sony* Court,

In an action for contributory infringement against the seller of copying equipment, the copyright holder may not prevail *unless the relief that he seeks affects only his programs, or unless he speaks for virtually all copyright holders with an interest in the outcome.*<sup>21</sup>

Because of the choice made by Dr. Waterman to include pornography and a number of other types of content when some copyright holders in those types of content have a different business model of digital distribution than that of the major studios, I question whether that last requirement has been satisfied.

### III FLAWS IN MR. ZEBRAK'S ASSESSMENT OF COPYRIGHT STATUS

34. Beyond the general methodological problems with the Waterman study, I have questions about specific decisions that Mr. Zebrak made in his review of the content to determine its copyright status. In my opinion, there are flaws in his methods.

35. First, because he is applying Dr. Waterman's protocol, he does not examine any zero download files in order to assess their copyright status. This excludes 54% of the files on the system – and one of the most important potential uses of the system – from consideration. I have pointed out the flaws this introduces to the study in Parts I and II of this Rebuttal Report and will not repeat those points here.

36. Second, it appears that, in those files that he *did* examine, Mr. Zebrak makes a clear methodological error. Effectively, his method seems to have focused intensively on the copyright status of the file itself, omitting full consideration of two key factors that one would need to examine in order to be able to classify a file as “highly likely infringing.”

- The type of use involved, including whether the conduct would constitute a fair use under section 107.
- The full range of possible forms of implied or express license by the copyright owner that would make the distribution legal.

#### **i.) Failure to Assess *Type of Use*: Fair Use**

37. I pointed out earlier that 5.76% of the files on Hotfile have only a single registered download. As with zero-download storage and backup, there is a very strong argument that

---

<sup>21</sup> *Sony Corp. v. Universal City Studios, Inc.*, 464 U.S. 417, 447 (1984). [emphasis added]

a user who purchases commercial, copyrighted content and “space shifts” a single copy of that content to a different computer, using Hotfile as the storage and download method, is making a fair use and is thus not violating the exclusive rights of the copyright holder. It was precisely a version of this argument that won the day in *Sony*. The content was copyrighted, commercially produced and was copied without permission – nevertheless the court, having considered all the aspects of the use, declared that it was a fair use. Space shifting was explicitly endorsed as a fair use in *RIAA v. Diamond Multimedia Sys., Inc.*<sup>22</sup> As I pointed out earlier, cases in the peer-to-peer context have cast doubt on whether this finding would hold true in a situation where a user sought to i.) claim fair use privileged access on a peer-to-peer network to *someone else’s* copy of a copyrighted work that the user himself had purchased, ii.) if that copy was being shared with the entire world. But in the context of zero or one download storage or space shifting on a cyberlocker neither of those other factors obtains. *Sony* and *Diamond* would therefore strongly suggest fair use. Mr. Zebrak’s deposition suggests that he took it to be black letter law that any file that is even theoretically available to others cannot thereby constitute space shifting or storage fair use.<sup>23</sup> In the context of a peer-to-peer network where numbers on downloads are unavailable this position might be credible. In a situation where we know the number of downloads to be zero or one, or in a situation where the link is not available on the open web, *Diamond’s* reasoning returns full force. At the very least, we cannot assume by the design of the study itself that such uses are not fair. The one download files return us squarely to the central category of uses in *Sony* and *Diamond*.

38. So far as I can tell, Mr. Zebrak’s analysis of downloads on Hotfile does not attempt to assess whether the use is of this type. Rather, from his description of his method, it would appear that his approach is one-dimensional. He looks at the legal status of the file in question and, if it is commercially produced and under copyright, with no evidence of formal open licensing, assumes that all copying is infringement. An analyst applying a similar method in the *Sony* case would have looked at the nature of the content in question – the movie *Shane*, say. The analyst would have discovered that *Shane* was commercially produced, was under copyright and was shared without permission. He then would have concluded, without looking at any other circumstances, including the number of copies made or by whom, that this was “highly likely infringing.” But an analysis with these assumptions would have found that almost *all* the uses of the VCR were “highly likely infringing.” In other words, it would have omitted the key variable on which *Sony* turned.

---

<sup>22</sup> *Recording Indus. Ass'n of Am. v. Diamond Multimedia Sys., Inc.*, 180 F.3d 1072, 1079 (9th Cir. 1999).

<sup>23</sup> “[W]e’re dealing with viral distribution of full-length commercial works, you know, without the authority of the copyright owner. That’s -- that’s what I concluded, and, you know, fair use is not applicable in that scenario. That’s well established.” DEPOSITION OF SCOTT ZEBRAK at 296. But in a situation where the file is only downloaded once, or the link to the file is not made available on the open web we cannot assume “viral distribution” of copyrighted works. We may well be dealing with exactly the kind of single copy, private storage dealt with in *Sony* and *Diamond*. Those uses – uses where there are zero or one copies made – do not somehow become unfair because the storage is “in the cloud” rather than in an iPod or on a dusty shelf behind the television.

39. There are a number of ways in which Mr. Zebrak's analysis could have been more accurate. The simplest would be to acknowledge that, in the case of 'one download' files, the fair use calculation made it impossible to say that the file was "highly likely infringing" and thus meant it must be included in the "unknowable" category. Other more complex methods that capture more of the factors relevant to fair use are also possible, such as classifying all single download files *that are not linked to on the open web* as "noninfringing" and those that are linked as "possibly infringing." A simple Google search would have enabled such a procedure, one that clearly distinguished between those files that were available publicly, and those that were effectively inaccessible to all but the uploader – itself further evidence of fair use. A failure even to consider the possibility of these forms of fair use renders the legal conclusions of the analysis particularly problematic in any study that purports to give the court relevant facts about the application of the test in *Sony*, a case that explicitly required attention to exactly such contextual issues. This appears to be a clear flaw in Mr. Zebrak's study. A study of a peer-to-peer network such as in the cases of *Napster* or *Grokster* would not need to pay as much attention to these factors, precisely because on a peer-to-peer network archival storage and backup is effectively impossible and space shifting less likely. A study of a cyberlocker site, however, has to pay attention to such issues. It is important to remember that Dr. Waterman's protocol excludes 54% of the files – the files with zero registered downloads – which could represent legal usage. When one adds to this the fact that Mr. Zebrak fails to consider fair use in looking at the 5.76% of files that were downloaded once, it seems that a total of nearly 60% of the files on Hotfile most likely to represent legal uses were either excluded from the study or classified using an incorrect procedure.

## **ii.) Errors in Classifying Content that Is Shared With Permission or Otherwise Legal to Distribute**

40. First, let me be clear that I am respectful of the daunting task that Mr. Zebrak faced in attempting to survey the copyright status of such a large number of files in a short period of time. Yet I have concerns about whether his method was accurate when applied to copyrighted content that was shared under an express or implied license. To his credit, Mr. Zebrak correctly identifies as non-infringing (and not illegal) those open source programs mentioned in my initial report that are found within his sample. That includes iReb and sn0breeze, the two most distributed files on the Hotfile system, and JDownloader, which is also very highly ranked. Yet beyond the world of software that is formally under an open source license, his method appears to have been tilted in the direction of finding content infringing even if there is strong evidence that it is shared with permission. Here are some examples:

41. **Orbit Downloader**<sup>24</sup> Orbit Downloader is a download assistant that is available for free download from <http://orbitdownloader.com>. The opening line in the site's "metatags" – the description of the site's content by the webdevelopers – is `<meta name="description" content="Orbit Downloader is a free social music, video and file downloader..">` Elsewhere

---

<sup>24</sup> Orbit Downloader is listed as file number 1510 in Mr. Zebrak's spreadsheet.



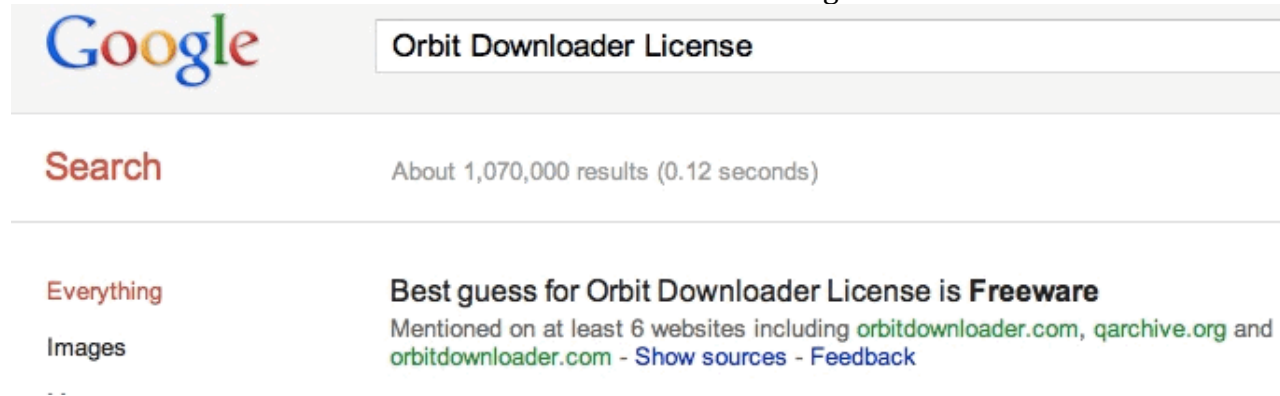
on the Orbitdownloader site, one can find the XML or Extensible Markup Language, data for the Portable Application Description.

<MASTER\_PAD\_INFO>Portable Application Description, or PAD for short, is a data set that is used by shareware authors to disseminate information to anyone interested in their software products. To find out more go to **Error! Hyperlink reference not valid.**><sup>25</sup>

The XML data provided by the developers of OrbitDownloader formally defines its qualities as follows. On each line a formal characteristic of the program is given between brackets that look like this <>....</>.

```
<Program_Name>Orbit Downloader</Program_Name>
<Program_Version>4.1.0.2</Program_Version>
<Program_Release_Month>06</Program_Release_Month>
<Program_Release_Day>28</Program_Release_Day>
<Program_Release_Year>2011</Program_Release_Year>
<Program_Cost_Dollars>0</Program_Cost_Dollars>....
<Program_Type>Freeware</Program_Type>26
```

That is, the developers of the software are explicitly identifying it as not merely zero cost but as “freeware.” If one goes to other popular and relatively authoritative download sites, such as CNet,<sup>27</sup> one will see the program listed as “freeware,” and available for free download. Finally, if one simply Google searches for “Orbit Downloader License,” Google, which has a “license search feature” will return the following:



Faced with this evidence, Mr. Zebrak classified Orbit Downloader as “Highly Likely Infringing.” The sources he gave to back up that conclusion included the Orbitdownloader .com site from which I have quoted, a site that clearly lists the program as both free and

<sup>25</sup> [http://dl.orbitdownloader.com/dl/pad\\_file.xml](http://dl.orbitdownloader.com/dl/pad_file.xml) (visited Dec 30, 2011)

<sup>26</sup> [http://dl.orbitdownloader.com/dl/pad\\_file.xml](http://dl.orbitdownloader.com/dl/pad_file.xml) (visited Dec 30, 2011) [emphasis added]

<sup>27</sup> [http://download.cnet.com/Orbit-Downloader/3000-2071\\_4-10600926.html](http://download.cnet.com/Orbit-Downloader/3000-2071_4-10600926.html) [last visited January 2<sup>nd</sup> 2012]

freeware. When asked about Orbit Downloader in his deposition, his reasoning appeared to be that – absent evidence that a company *specifically* authorized a particular distribution channel such as Hotfile – content is to be classified as “highly likely infringing,” even where the copyright owners themselves classify it as freeware.<sup>28</sup> Of course, a company might distribute at zero cost through certain sites and prefer not to distribute through others. Copyright gives them the legal right to make that choice. Yet if they formally classify their product as “freeware” and fail to include any End User License Agreement to the contrary, I think the argument for either express or implied license to reproduce is a solid one. At the very least, from this evidence one could not responsibly classify such a program as “*Highly Likely Infringing*.”

**42. Photography 101: Professional Photography Tips Tutorial**<sup>29</sup> Mr. Zebrak lists Photography 101 as Highly Likely Infringing. Photography 101 is in fact a popular set of podcasts by Scott Wittenburg, a photography teacher who distributes his podcasts freely online.<sup>30</sup> Mr. Wittenburg’s own site includes links to free versions of these podcasts (though, like many purveyors of free content, he also has a paid “app” that allows you to view the content more easily on your smartphone. This viral distribution of free content as an advertisement for other services is a common business method online and one that Mr. Zebrak’s working assumptions might lead him to misclassify.) Defendant’s counsel gave me an affidavit from Mr. Wittenburg. In that affidavit, which is attached, Mr. Wittenburg states “I know that by making my podcasts available for free on the internet, that people are able to download them and also repost them. So long as a person is not charging money for my podcast, I do not have any problems with this.”<sup>31</sup> Individuals clearly could download Mr. Wittenburg’s podcast from Hotfile without being charged money. This appears to be a legal reposting of Mr. Wittenburg’s podcast. It certainly cannot be classified, as Mr. Zebrak does, as “highly likely infringing.”

**43. DirectX**<sup>32</sup> DirectX is Microsoft’s collection of multimedia and gaming API’s (Application Programming Interfaces) that allow games and multimedia programs to play on, and thoroughly use, the capabilities of Windows platforms. Thus, developers of games frequently need to distribute the DirectX libraries together with their games. Microsoft freely distributes the DirectX libraries under an End User License Agreement (EULA)<sup>33</sup> that explicitly permits game developers to distribute the DirectX libraries with their games.

---

<sup>28</sup> “So, you know, the fact that it was doing that on its own web site doesn't -- doesn't make it less likely to me -- I'm saying that awkwardly. The fact that it's doing that on its own web site, if it's doing that, doesn't -- doesn't change my opinion that the distribution of it through Hotfile was unauthorized.” Deposition of Scott Zebrak at 307.

<sup>29</sup> Photography 101 is listed as file number 132 in Mr. Zebrak’s spreadsheet. It contains 7 of Mr. Wittenburg’s podcasts.

<sup>30</sup> <http://scottwittenburg.com/> [last visited Jan 3<sup>rd</sup> 2012]

<sup>31</sup> Affidavit of Scott Wittenburg 19<sup>th</sup> December, 2011. EXHIBIT E.

<sup>32</sup> DirectX is part of a set of files listed as file number 30 in Mr. Zebrak’s spreadsheet. See EXHIBIT H (Direct X Exhibits).

<sup>33</sup> END-USER LICENSE AGREEMENT FOR MICROSOFT SOFTWARE DirectX 9.0 Software Development Kit Update (October 2004) (Attached.)

That is, Microsoft not only makes the program available freely, it does so under a license that allows a game developer to *redistribute* the Direct X library in or with their game. Mr. Zebrak classifies what is apparently a ski jumping game – Skoki 2006 – as highly likely infringing and states that this is because the game folder includes DirectX. At my direction, the computer consulting company Elysium Digital examined the files that Mr. Zebrak classified as infringing. So far as I can tell, they are the files covered by Microsoft’s EULA. One of the files has a slightly different hash, but I think that it may simply be an earlier version of the program. In fact, if one installs DirectX from the game distribution, Elysium Digital confirmed that, during the installation, one is required to assent to a EULA binding the user to the terms of the DirectX EULA, thus complying with the requirements that Microsoft had set up for the “redistributable” portion of DirectX. Based on these facts, I do not think Mr. Zebrak can classify the software as “highly likely infringing” for its inclusion of the DirectX files. In fact it appears to be distributed in exactly the way Microsoft envisioned in writing the “redistributable” portion of the EULA. I would also note that in practice the software is widely available around the World Wide Web on reputable and highly visible sites where drivers or API libraries can be found – such as “Major Geeks.” Elysium Digital identified multiple examples of the DirectX software being made freely available by itself, suggesting, but not proving, that Microsoft tolerates distribution even more widely than the license suggests. That implication is not necessary for my analysis, however. Based on all of these facts, I would say the copyright status of the DirectX library, the software Mr. Zebrak focused on, is either “likely non-infringing” or, at the most conservative, “unknown.” It cannot in my view be classified as “highly likely infringing.”

**44. Farming Simulator “Mods”** This litigation has been an enlightening experience in many ways, but none perhaps more delightful than the discovery that there is a great interest in a game called “Farming Simulator.”<sup>34</sup> Farming Simulator, published by Giants Software is a simulator game akin to “Sim City.” The player makes certain choices and based on that, her farm thrives or fails to thrive. Farming Simulator, like many games, allows users to create new aspects to the game. Indeed it provides an editor program that assists the user to do so. In some other games, these new aspects consist of new levels or landscapes in which the game is played. In the case of Farming Simulator, users can create new features called “Mods” that generally consist of new types of farm equipment that the game will feature. Interestingly, and contrary to the assumption that commercial providers of copyrighted works would never relinquish any aspect of control over their works, makers of copyrighted games frequently allow and even encourage this practice. This practice has sufficiently fascinated scholars, that it has attracted its own academic literature, including articles such as *The Labour of User Co-Creators: Emergent Social Network Markets?*<sup>35</sup> The practice, and the academic literature, are directly relevant to this litigation in that they demonstrate another reason that one cannot assume that high quality, commercially produced online content is only shared illicitly. Sharing that content

---

<sup>34</sup> <http://www.farming-simulator.com/> [last visited January 6<sup>th</sup> 2012]

<sup>35</sup> John Banks & Sal Humphries, *The Labour of User Co-Creators : Emergent Social Network Markets?* 14 *Convergence: The International Journal of Research into New Media Technologies* vol. 401-418 (November 2008). The canonical book on the more general practice is Eric von Hippel, *DEMOCRATIZING INNOVATION* (MIT Press 2005).

licitly is in fact a central part of many business models.

45. Giants Software is one of the companies that encourages this practice, that is, it encourages its users to produce and to share “Mods.” As mentioned before, Giants Software actually includes an editor program in the game to make it easier for users to create Mods. They have even held competitions as an incentive to the practice, a fact noted on the website Mr. Zebrak cites in his reasons for claiming the Mods are infringing.<sup>36</sup> In the sample of 1750 files examined by Mr. Zebrak there are multiple examples of Farming Simulator Mods – that is, multiple file directories containing Mod files, each one of which Mr. Zebrak assessed and classified separately. Co-counsel in this case sent eight of those files to Giants Software, the copyright holders in and developers of Farming Simulator, to confirm that the company had no objection to the sharing of these Mods. Of the eight, Mr. Zebrak had found seven “highly likely infringing” and one non-infringing. Mr. Schwegler of Giant Software provided an Affidavit confirming that all of the files were not infringing. He also provided an email stating “I got the files. and checked them all. They are all free mods created by fans of our game and are legal to share anywhere on the web including Hotfile. The items do not infringe on our copyrights and do not contain cracks, serial keys or similar illegal software which would compromise our products.” It is hard to think of a more definitive answer.<sup>37</sup> In my opinion Mr. Zebrak’s classification of “Highly Likely Infringing” for seven of the Mods is clearly incorrect. Elysium Digital discovered two additional Mods, AA01 and BiginParadies, both listed as “highly likely infringing.” (All Mods are listed in the attached analysis.)<sup>38</sup> Given these facts, I would classify these files, too, as highly likely non-infringing, but I was not able to confirm this with Mr. Schwegler before this report was to be filed.

**46. Opera Portable Browser** Opera Portable Browser is a version of the Opera web browser that can be run from a USB stick. Mr. Zebrak classified it as highly likely infringing. There are both free and paid versions of the browser. An e-mail inquiry to Opera elicited a response<sup>39</sup> that seems to indicate they do not object to cloud storage of the free version. At the time this report was filed, I was still attempting to discover whether the version of software shared on Hotfile was the free version.

47. These examples are indicative of a larger point. Many, many copyright holders allow

---

<sup>36</sup> <http://www.farming-simulator.com/modContest2011.php> [last visited January 6th 2012]

<sup>37</sup> One could postulate the Mod creators objecting to the copying of their Mods but I think this far fetched since the only way for them to get online is for the user to post them himself. Further, the fact that users *like* sharing their Mods and showing their competence at creating them is well established in the scholarly literature cited earlier. Users do sometimes ask for attribution – one such request was included in the Mods mentioned here. The file posted on Hotfile included the requested attribution. Another Mod contained a claim for rights to an “Excerpt” of code. This too seems consistent with permission from the author of the Mod.

<sup>38</sup> See Elysium Farming Simulator analysis, EXHIBIT G.

<sup>39</sup> See EXHIBIT F.

redistribution of their work online, some of them as part of a profit making strategy, others because they have simply chosen to share the work. Mr. Zebrak stated several times in his deposition that his assumption is that if content was being generated as part of a profit making enterprise, then the copyright holder would object to it being shared and therefore he could classify it as “highly likely infringing.”<sup>40</sup> But on the world of the Internet, that assumption is a problematic one. Microsoft is a profit-making company but they want developers of games to embed Microsoft’s software platform, DirectX, inside their games and thus they give permission to distribute versions of games that include those files. Mr. Wittenburg makes high quality podcasts giving lessons in photography, but distributes those podcasts freely. Farming Simulator players create and freely share Mods – a practice encouraged by the copyright holder in the game. Mr. Zebrak’s assumption causes him to incorrectly classify all of these examples. To use an example that does not occur in Mr. Zebrak’s study, Nine Inch Nails distribute their album Ghosts I-IV under a Creative Commons license. It is legal to copy and redistribute non-commercially online. Yet at the same time, they sell CD’s and digital copies of their music and in fact that album was the best selling MP3 download album on Amazon.com in 2008.<sup>41</sup> Note the way in which this situation does not fit Mr. Zebrak’s background assumptions. Similarly, I note that there is music from little-known artists from Bulgaria and Turkey in Mr. Zebrak’s study.<sup>42</sup> Do the musicians, particularly those in countries with less well-developed music distribution systems than the United States, object to viral distribution of their songs or do they welcome it as a way of building recognition and increasing demand for concert performances? I would want more facts before I assumed that all this content was “highly likely infringing.” Finding this pattern of errors in the files I did examine makes me question whether the pattern continues in the ones I did not.

48. In the case of Hotfile, these concerns are not academic ones. I have attached to this study a collection of Counter Notices to Takedown requests received by Hotfile.<sup>43</sup> Those protesting include a musician who shares the musician’s own work online using Hotfile, a company that writes and freely distributes firmware updates for Samsung products and an architecture company that uses Hotfile for storage of drawings made for clients. All are complaining that their work has been wrongly removed when they intended to share it with permission, wanted to use the Hotfile service as one of their distribution channels, and had every right to do so. Classification of the copyright status of works shared online is extremely difficult and time-consuming – for content owners and online services alike. It is highly factually specific and easily distorted if one assumes that all of those distributing content online have the same business model or motivation. The Counter Notices, as well as the points made in my initial report, raise an additional point – one that was of particular interest to the *Sony* Court. There are clearly individuals and companies who

---

<sup>40</sup> Zebrak at p. 218-219; p. 257-258 (“professional artist” wouldn’t allow distribution through Hotfile); p. 276-278 (“antithesis”).

<sup>41</sup> Nate Anderson, *Free Nine Inch Nails Albums Top 2008 Amazon MP3 Sales Charts* <http://arstechnica.com/media/news/2009/01/free-nine-inch-nails-albums-top-2008-amazon-mp3-sales-charts.ars>

<sup>42</sup> See Zebrak Depo. pp. 209 and 280-84 (Turkish Rap) and 266-70 (Bulgarian Pop).

<sup>43</sup> See EXHIBIT C.

wish to use the Hotfile service to pursue entirely legitimate goals ranging from personal back up and storage to creating and distributing open source software and being compensated for it through the Affiliate Program to storing architectural drawings. These users wish to use this service to do things that are entirely in accord not only with the Copyright Act, but with the larger goals in Article 1 section 8 clause 8 of the Constitution. This lawsuit, and the plaintiffs' curiously narrow design of Dr. Waterman's study, implicate – and in the case of storage and space shifting, improperly ignore – those unquestionably legitimate uses. Discussing the analogy between contributory copyright infringement and contributory patent infringement, the Supreme Court had this to say;

When a charge of contributory infringement is predicated entirely on the sale of an article of commerce that is used by the purchaser to infringe a patent, the public interest in access to that article of commerce is necessarily implicated. A finding of contributory infringement does not, of course, remove the article from the market altogether; it does, however, give the patentee effective control over the sale of that item. Indeed, a finding of contributory infringement is normally the functional equivalent of holding that the disputed article is within the monopoly granted to the patentee.<sup>44</sup>

Because it believed that intellectual property holders should not be able to veto technological developments or services merely because those developments *could* be used to violate their intellectual property rights, the Court found that possibility unacceptable in the copyright as well as the patent context so long as the article had “substantial non-infringing uses.” I mention this legal background only to make the point that it is unfortunate that so few of those uses are reflected in Dr. Waterman and Mr. Zebrak's study.

49. Finally, I have more general concern about the accuracy of Mr. Zebrak's classification. It seems at times that his default assumption is that content is “highly likely infringing” and that considerable evidence is required to shift the needle on that point. He includes, for example, an 1871 Russian book<sup>45</sup> on the subject of weaving and embroidery techniques as “highly likely infringing.” The illustrations in the book are quite beautiful, but the idea that a book which carries the date “1871” on its cover is “highly likely infringing” in the United States is truly a strange one. (Published works from before 1923 are in the public domain in the United States.<sup>46</sup>) Mr. Zebrak links to a 1976 Dover Books edition on Amazon.com,<sup>47</sup> but this is not the version found on Hotfile. Dover Books is a publishing company that predominantly reissues works that have fallen into the public domain. The cover of the 1976 Dover Books edition is significantly different and the title is in English, making it easy to distinguish between the two at first sight. Dover's copyright would, in any event, only extend to any original material that they added, such as a new cover, not to the underlying

---

<sup>44</sup> *Sony Corp. v. Universal City Studios, Inc.*, 464 U.S. 417 440-441 (1984).

<sup>45</sup> A literal translation of the Russian title would be RUSSIAN ORNAMENT: SEWING, FABRIC, LACE (St. Petersburg: 1871)

<sup>46</sup> 17 U.S.C. § 304

<sup>47</sup> [http://www.amazon.com/Russian-Peasant-Needleworkers-Craftsmen-Pictorial/dp/0486232352/ref=sr\\_1\\_1?ie=UTF8&qid=1320680095&sr=8-1](http://www.amazon.com/Russian-Peasant-Needleworkers-Craftsmen-Pictorial/dp/0486232352/ref=sr_1_1?ie=UTF8&qid=1320680095&sr=8-1)

book or images, which remain in the public domain. Mr. Zebrak also links to another site that does contain the same file found on Hotfile and that correctly identifies the book's date of publication as 1871 and its place of publication as St. Petersburg,<sup>48</sup> so he must have been aware of its publication date. Given these facts, this book is clearly in the public domain and I am surprised to see Mr. Zebrak assert otherwise. Errors such as these in that fraction of his sample I did examine make me wary of the accuracy of Mr. Zebrak's assessments in the remainder of his sample.

#### IV

#### RELEVANCE OF THE FLAWS IN THE STUDY TO ANY INDUCEMENT LIABILITY CLAIM

50. Many of my comments have been directed to the way that the flaws in the Waterman/Zebrak study are problematic for any court investigating "substantial non-infringing uses" under *Sony*. Before concluding, I would like to highlight several key connections of those flaws to the factual analysis a court would perform in assessing any claim of *Grokster* style inducement liability.

51. First, and most obviously, the *Grokster* test is a multi-part one, with no single portion being sufficient. In applying such a test, a finder of fact will be guided by a sense of the overall *bona fides* of the service in question. By omitting key legal uses of Hotfile from the study, the Waterman report, in my opinion, provides a misleading starting place for such an assessment.

52. Second, the specific omissions from the Waterman study are relevant to particular components of the *Grokster* test. *Grokster* asks a court to engage in a complex study of whether a service is aiming to profit principally from infringement. If one omits storage or space shifting from one's picture of Hotfile, as the Waterman report does, then features of Hotfile's system – such as its removal of files that have not been downloaded after three months, for example, can be cast in a negative light. If one includes storage and space shifting, however, and the fact that Premium users are allowed to store their material permanently regardless of whether it is downloaded, then the business model looks rather different and altogether more benign.

---

<sup>48</sup>[http://translate.google.com/translate?hl=en&sl=ru&u=http://valhalla.ulver.com/f49/t11428.html&ei=\\_y20TqTMM5KRgQefm9iwBA&sa=X&oi=translate&ct=result&resnum=1&sqi=2&ved=0CCQQ7gEwAA&prev=/search%3Fq%3D%25D0%25A0%25D1%2583%25D1%2581%25D1%2581%25D0%25BA%25D0%25B8%25D0%25B9%2B%25D0%25BE%25D1%2580%25D0%25BD%25D0%25B0%25D0%25BC%25D0%25B5%25D0%25BD%25D1%2582.%2B%25D0%25A8%25D0%25B8%25D1%2582%25D1%258C%25D1%2591,%2B%25D1%2582%25D0%25BA%25D0%25B0%25D0%25BD%25D0%25B8,%2B%25D0%25BA%25D1%2580%25D1%2583%25D0%25B6%25D0%25B5%25D0%25B2%25D0%25B0.rar.html%26hl%3Den%26biw%3D685%26bih%3D300%26prmd%3Dimvns](http://translate.google.com/translate?hl=en&sl=ru&u=http://valhalla.ulver.com/f49/t11428.html&ei=_y20TqTMM5KRgQefm9iwBA&sa=X&oi=translate&ct=result&resnum=1&sqi=2&ved=0CCQQ7gEwAA&prev=/search%3Fq%3D%25D0%25A0%25D1%2583%25D1%2581%25D1%2581%25D0%25BA%25D0%25B8%25D0%25B9%2B%25D0%25BE%25D1%2580%25D0%25BD%25D0%25B0%25D0%25BC%25D0%25B5%25D0%25BD%25D1%2582.%2B%25D0%25A8%25D0%25B8%25D1%2582%25D1%258C%25D1%2591,%2B%25D1%2582%25D0%25BA%25D0%25B0%25D0%25BD%25D0%25B8,%2B%25D0%25BA%25D1%2580%25D1%2583%25D0%25B6%25D0%25B5%25D0%25B2%25D0%25B0.rar.html%26hl%3Den%26biw%3D685%26bih%3D300%26prmd%3Dimvns)  
Collection patterns of Russian folk ornamentation (embroidery, fabrics, laces).  
Title: Russian ornament. Sewing, fabric, lace Year: 1871 Publisher: St. Petersburg Series / Issue: A series or Issue: Pages: 42 Quality: good Size: 9.81 MB Format: DjVu  
Language: Russian

53. Third, I have argued here that Mr. Zebrak’s study is, in a number of ways, prone to make errors that predispose him to classify files as “highly likely infringing.” In other words, I have argued there is reason to believe that his statistics on infringement are too high and his identification of non-infringing content too low. Suppose for a moment, however that we accept Mr. Zebrak’s classifications as entirely accurate. One key feature of Hotfile’s business plan is to convert users to Premium status. Premium allows longer storage times, but it also allows faster downloads. Hotfile keeps a log heading (called “paidfor”) of what particular file prompted users to “convert,” that is, what file the user was so drawn to that he chose to subscribe to the paid Premium service. Using Mr. Zebrak’s own figures and classifications as the basis, Elysium Digital prepared a chart of the relative conversion rates for each of the types of content he identified. (That is, those users who converted to the Premium service on that type of file as a percentage of the total number of downloads of that category.)

Zebrak Category	Sum of <i>paidfor</i>	Dailydownload Total	Conversion Rate
Confirmed Infringing	44	215302	0.0204%
Highly Likely Infringing	1245	2123933	0.0586%
Noninfringing	699	650727	0.1074%
Unknowable	116	316235	0.0367%

Using Mr. Zebrak’s own classifications, which I have argued substantially underestimate the amount of Non-infringing work, we find that the Non-infringing category is nearly twice as good (1.82) at converting users to Premium as the Highly Likely Infringing. More notably still Non-infringing material is more than 5 times (5.264) as likely to cause users to convert to Premium as the Confirmed Infringing category – the major studio content that is the subject of this litigation. A rational Hotfile executive, knowing these numbers, would prefer Non-infringing content to Infringing content and, of all the content on the list, would be least interested in getting uploads of Confirmed Infringing content, that is of the copyrighted content owned by the plaintiffs in this case.

## V CONCLUSION

54. In my opinion, the study performed by Dr. Waterman and Mr. Zebrak has a number of flaws. These flaws are individually serious, and their effects are cumulatively more misleading.

- By focusing only on downloads, and thus ignoring files with zero registered downloads, the study omits one of the central potential uses of a cyberlocker site: personal storage. Zero download personal storage is likely to be legal, either



because the content was generated by the user, or because it is merely an archival or backup copy of licitly purchased content, not made available to others, and thus probably a fair use. The fact there are zero downloads is consistent with either of these scenarios. The result of the decision to focus only on downloads is to ignore the majority, 54%, of the files of the system and also to ignore one of the most important potential substantial non-infringing uses of the system. I would note that the possibility of personal storage is one of the qualities that distinguishes a service such as Hotfile from a peer-to-peer system such as Grokster. My claim here is not that I know what percentage of the zero download files on Hotfile are examples of personal storage. It is that it is a serious error in Dr. Waterman's study design to omit that possibility from consideration in the statistical picture he paints of the service.

- Turning now to the files he does study, Dr. Waterman chooses to include legal pornographic content in his file samples, excluding only illegal content such as child pornography. Other studies, such as that performed in the *Grokster* case, apparently excluded all pornographic content. This has large consequences for Dr. Waterman's study. 25 of the first 100 files in his sample are apparently pornographic as compared to nine files that are confirmed infringing examples of major studio content. Given the difficulties in assessing the copyright status, business methods and possible fair use claims of pornographic content, in my opinion this decision was erroneous and pornographic content should either have been omitted or only included when it was classified as "confirmed infringing." This means that, of the 46% of the files on Hotfile that were potentially a subject of Dr. Waterman's study, that is, those that were downloaded at least once, a substantial proportion and perhaps as many as a quarter are of a kind that might lead the court to doubt the accuracy of any assessment of their infringing status.
- Mr. Zebrak's legal classification of the files is erroneously incomplete in that it fails to consider information that would bear on fair use. His method appears to have been to examine the file to see if it was commercial copyrighted content and if so, and there was no immediate evidence that the copyright holder gave permission for copying, to classify it as "highly likely infringing." The 5.76% of the files on Hotfile that were downloaded once – a number consistent with both "space shifting" and backup or archival storage – are particularly germane here. Since the potential universe of files that Dr. Waterman looked at consisted of only 46% of the total files on Hotfile (that is, those with one or more downloads) the one download files constitute 12.5% of the total potential file universe from which Dr. Waterman's sample could draw. (I have no information on the *actual* percentage of one download files in Dr. Waterman's sample.) Mr. Zebrak could easily have made a separate classification for one download files found in Dr. Waterman's sample, perhaps by including them in his "unknowable" category. His failure to consider information bearing on potential fair uses here means that his classifications as "highly likely infringing" cannot, without more, be relied upon. Ironically, Mr. Zebrak's technique, had it been applied in the *Sony* case, would have excluded the

very non-infringing use on which the case turned – namely home taping, without permission, of single copies of commercially produced copyrighted content. This is a serious error in the study. In my opinion, this error alone precludes relying on Mr. Zebrak’s assessment of copyright status.

- Mr. Zebrak also appears to have operated on an assumption that producers of high quality copyrighted content would not have a business model that involved allowing the “viral” sharing of that content online. He does note, to his credit, that iReb and Sn0breeze – the two most widely shared files on Hotfile – are non-infringing and legal to distribute. Yet in his other assessments there were apparently a number of errors in legal classification, ranging from Mr. Wittenburg’s podcasts to the Orbit Downloader. Those errors, and the inclusion of such works as the 1871 Russian embroidery manual as “highly likely infringing,” concern me about the accuracy of his classification methods.

55. I am concerned about the effect of these flaws on the accuracy of the statistical picture that Dr. Waterman and Mr. Zebrak paint for the court. In particular I note that their methodology excludes a majority of the files on the system and ignores two of the central potential non-infringing uses, namely zero or one download storage, and one download “space shifting.” I wish to stress this point. Dr. Waterman’s protocol excludes 54% of the files on Hotfile – the zero download files – when they clearly could represent legal usage. When one adds to this the fact that Mr. Zebrak fails to consider fair use in looking at the 5.76% of files that were downloaded once, it seems that *a total of nearly 60% of the files on Hotfile most likely to represent legal uses were either excluded from the study or classified using an incorrect procedure.* I note also that the files that remain include a high percentage of pornographic content – omitted in some prior studies – and that there are a number of errors in classifying as infringing material that was actually shared with permission, was a fair use or was in the public domain. In my opinion, the cumulative effect of these flaws is to present a distorted and inaccurate statistical picture of the Hotfile service.

Signed,



James Boyle

UNITED STATES DISTRICT COURT  
SOUTHERN DISTRICT OF FLORIDA

CASE NO. 11-20427-WILLIAMS-TURNOFF

DISNEY ENTERPRISES, INC.,  
TWENTIETH CENTURY FOX FILM  
CORPORATION, UNIVERSAL CITY  
STUDIOS PRODUCTIONS LLLP,  
COLUMBIA PICTURES INDUSTRIES,  
INC., and WARNER BROS.  
ENTERTAINMENT INC.,

*Plaintiffs,*

v.

HOTFILE CORP., ANTON TITOV, and  
DOES 1-10.

*Defendants.*

---

**CERTIFICATE OF SERVICE**

I am a citizen of the United States and employed in San Francisco County, California. I am over the age of eighteen years and not a party to the within-entitled action. My business address is 235 Montgomery Street, 17th Floor, San Francisco, California 94104; my email address is dgracia@fbm.com.

I HEREBY CERTIFY that on January 6, 2012 I electronically served the following document(s) on all counsel of record on the attached Service List via their email address(es) as set forth on the Court's CM/ECF filing system, or in some other authorized manner for those counsel or parties who are not set forth on the Court's CM/ECF filing system. The documents served on this date is:

**THE REBUTTAL EXPERT REPORT OF JAMES BOYLE**

I declare under penalty of perjury that the foregoing is true and correct. Executed on January 6, 2012, at San Francisco, California.

/s/

---

Deepak Gupta

**SERVICE LIST: CASE NO. 11-CIV-20427-WILLIAMS-TURNOFF**

Daniel M. Mandil, Esq.  
Karen R. Thorland, Esq.  
Motion Picture Association of America, Inc.  
15301 Ventura Blvd., Building E  
Sherman Oaks, CA 91403  
Telephone: (818) 935-5812  
Fax: (818) 285-4407  
Email: [Daniel\\_Mandil@mpaa.org](mailto:Daniel_Mandil@mpaa.org);  
[Karen\\_Thorland@mpaa.org](mailto:Karen_Thorland@mpaa.org)

Duane C. Pozza, Esq.  
Luke C. Platzer, Esq.  
Steven B. Fabrizio, Esq.  
Jenner & Block  
1099 New York Avenue, N.W., Ste. 900  
Washington, DC 20001-4412  
Telephone: (202) 639-6094  
Fax: (202) 639-6068  
Email: [dpozza@jenner.com](mailto:dpozza@jenner.com);  
[lplatzer@jenner.com](mailto:lplatzer@jenner.com); [sfabrizio@jenner.com](mailto:sfabrizio@jenner.com)

*Attorneys for Plaintiffs*

*Party Name: Disney Enterprises, Inc.,  
Twentieth Century Fox Film Corporation,  
Universal City Studios Productions LLP,  
Columbia Pictures Industries, Inc., Warner  
Bros. Entertainment Inc.*

**Served via electronic mail by agreement**

Karen Linda Stetson, Esq.  
Gray-Robinson P.A.  
1221 Brickell Avenue, Suite 1650  
Miami, FL 33131  
Telephone: (305) 416-6880  
Fax: (305) 416-6887  
Email: [kstetson@gray-robinson.com](mailto:kstetson@gray-robinson.com)

*Attorneys for Plaintiffs*

*Party Name: Disney Enterprises, Inc.,  
Twentieth Century Fox Film Corporation,  
Universal City Studios Productions LLP,  
Columbia Pictures Industries, Inc., Warner  
Bros. Entertainment Inc.*

**Served via electronic mail by agreement**

*Attorneys for Plaintiffs*

*Party Name: Disney Enterprises, Inc.,  
Twentieth Century Fox Film Corporation,  
Universal City Studios Productions LLP,  
Columbia Pictures Industries, Inc., Warner  
Bros. Entertainment Inc.*

**Served via electronic mail by agreement**

BOSTON LAW GROUP  
Valentin Gurvits  
825 Beacon Street, Suite 20  
Newton Center, MA 02459  
Telephone: (617) 928-1800  
Fax: (617) 928-1802  
Email: [vgurvits@bostonlawgroup.com](mailto:vgurvits@bostonlawgroup.com)

*Attorneys for Defendants*

*Hotfile Corp and Anton Titov*

**Served via electronic mail by agreement**

Janet T. Munn, Esq.  
Rasco Klock  
283 Catalonia Avenue, Suite 200  
Coral Gables, FL 33134  
Telephone: (305) 476-7101  
Fax: (305) 476-7102  
Email: [jmunn@rascoklock.com](mailto:jmunn@rascoklock.com)

*Local Attorney for: Defendants*  
*Party Name: Hotfile Corp. and Anton Titov*

**Served via electronic mail by agreement**