

**EXPERT REPORT OF J. D. TYGAR PURSUANT TO FEDERAL RULE OF
CIVIL PROCEDURE 26(a)(2)(B)**

A&M RECORDS, INC. ET AL. V. NAPSTER, INC. CASE NUMBER C 99-5183 MHP (ADR)
JERRY LEIBER ET AL. V. NAPSTER, INC. CASE NUMBER C 00-0074 MHP (ADR)
UNITED STATES DISTRICT COURT, NORTHERN DISTRICT OF CALIFORNIA
(SAN FRANCISCO DIVISION)

FOR HEARING JULY 26, 2000

TABLE OF CONTENTS

INTRODUCTION AND BACKGROUND..... 3

SUMMARY OF CONCLUSIONS 6

OVERVIEW OF PERTINENT TECHNOLOGY – DIGITAL AUDIO AND COMPRESSION..... 10

CONCLUSION 1: NAPSTER ALLOWS USERS TO SHARE RECORDED MUSIC 15

CONCLUSION 2: NAPSTER ALLOWS USERS TO SHARE MP3 FILES 23

**CONCLUSION 3: NAPSTER DOES NOT HAVE ACCESS TO INFORMATION ABOUT
COPYRIGHTS..... 28**

CONCLUSION 4: NAPSTER CAN NOT CHECK AUTHORIZATIONS 34

**CONCLUSION 5: AUTHORIZATION WOULD CHANGE THE WEB TO A CENTRALIZED
UTILITY 35**

**CONCLUSION 6: WATERMARKING COULD CARRY RIGHTS INFORMATION WITH A
RECORDING 37**

**CONCLUSION 8: ID/PASSWORD MECHANISMS ARE A CUSTOMARY WAY OF ALLOWING
ACCESS..... 50**

CONCLUSION 9: REJECTING BOTS IS A WAY TO HELP MAINTAIN PERFORMANCE 56

APPENDIX A: COURT CASES FOR WHICH I HAVE PROVIDED TESTIMONY 59

APPENDIX B: MY CURRICULUM VITAE..... 60

Introduction and Background

My name is Justin Douglas Tygar. I am a tenured, full Professor at the University of California, Berkeley, with a joint appointment in the Department of Electrical Engineering and Computer Science (Computer Science Division) and the School of Information Management and Systems. Prior to joining UC Berkeley in 1998, I was a tenured faculty member in the Computer Science Department at Carnegie Mellon University. I continue to hold a nominal position at Carnegie Mellon University as a tenured faculty member on leave, and effective September 1, my Carnegie Mellon position will convert to an adjunct faculty member.

I am an expert in software engineering, computer security, and cryptography. I have taught courses in software engineering and computer security, at the undergraduate, master's, and Ph.D. level at both UC Berkeley and Carnegie Mellon University.

I am the recipient of the Presidential Young Investigators Award from the National Science Foundation and the Office of the White House. That award was in part for work on mechanisms to address issues of copy protection. I am currently a principal investigator or a co-principal investigator on software engineering related grants or contracts from the Defense Advanced Projects Research Agency, the National Science Foundation, and the United States Postal Service. I am currently a co-principal investigator on a grant from the National Science Foundation with my UC Berkeley colleagues Pamela Samuelson (a law professor working in the field of copyright law) and Hal Varian (an economics and business professor working in the field of Internet Economics). Under this grant, we are investigating mechanisms (both technical and policy mechanisms) to handle rights management issues.

I helped design the security standards for the US Postal Service's Information Based Indicia Program (cryptographic postal indicia). I was the co-inventor of a major electronic commerce payment system called NetBill which has been patented, implemented, and licensed to a commercial company CyberCash. I have consulted on a variety of software development projects at commercial companies, government organizations, and university research projects. I served on the National Academy of Science/National Research Council's information security group and helped author a book *Trust in Cyberspace* that discusses computer security issues. I recently have been appointed to the new International Federation for Information Processing (IFIP) Working Group on Internet Applications Engineering.

Appendix A contains a list of court cases for which I have provided testimony at trial or at deposition since January 1, 1996. Appendix B contains my curriculum vitae (including a full publication list for the last ten years).

Napster's counsel requested that I provide expert analyses in several lawsuits against Napster. In particular, I was asked to analyze issues relating to the structure of the Internet and the Napster application, and on associated security issues. I was also asked to opine on copy protection mechanisms, and the structure of Napster compared with existing commercially available audio and video copying systems. I am charging \$250/hour for this work and \$400/hour for any testimony associated with this work.

In reaching my conclusions, I have drawn on my experience and knowledge as a researcher in computer science, cryptography, computer security, and electronic commerce. As a computer scientist and an individual interested in audio applications, for many years I have kept abreast of developments in electronics and audio. I am also an

avid collector of recorded music on CD, and enjoy reading magazines and journals on the topics of recorded music and audio systems. Beyond my technical and personal background, I have used a variety of materials in preparing this report from both the World Wide Web and the Nexis news archive search system. Most of the material I used is explicitly cited in the text of this report; material that I used but that is not otherwise cited in this report includes:

- The Data Compression FAQ (<http://www.faqs.org/faqs/compression-faq/>)
- The Microsoft Developers Network (<http://msdn.microsoft.com>)

As I examine additional materials and perform further analyses, I reserve the right to revise and supplement my opinions.

Summary of Conclusions

I have reached the following conclusions in my analysis:

1. Napster allows users to share computer files that contain recorded music. In allowing users to reproduce music Napster may be compared to a number of consumer products including:
 - cassette decks;
 - dual dubbing cassette decks;
 - cassette radios;
 - CD/cassette combinations;
 - video cassette recorders (VCRs);
 - Minidisc recorders;
 - digital audio tape recorders;
 - Tivo and Replay digital hard disk video recorders;
 - CD burning software;
 - encoding or “ripping” software that takes source audio material and converts it into computer readable form;
 - MP3 portable storage devices that allow computers to download computer readable audio files for portable playback including MP3 portable storage devices using memory sticks, compact flash memory, and smart media memory; and
 - MP3 playing software.

2. Napster allows users to exchange and share MP3 files. In this way, it is similar to existing file sharing techniques including:

- electronic mail (e-mail);
- the World Wide Web;
- the File Transfer Protocol (FTP);
- search engines (such as Lycos or AltaVista);
- caching search engines (such as google.com);
- Gnutella;
- Freenet; and
- SpinFrenzy.

3. Napster can not distinguish between:

- material protected by copyright and restricted by the owner;
- material protected by copyright but for which the owner or the law permits free distribution; and
- material not protected by copyright.

Furthermore, to require Napster, any other search engine, or any file sharing utility to obtain an authorization from the rights holder prior to providing access to its material is technically infeasible and would prevent the effective operation of the utility.

4. Napster has no practicable way of checking that an authorization is from the party it purports to be from or that party holds any or all rights to any particular file.

5. Even if there were a way to check authorizations for all files that would be shared under Napster, the World Wide Web, or any other file sharing utility, the effect of

requiring authorizations would change the utility from a decentralized, ground-up information base to a centrally controlled top-down distribution device.

6. The Recording Industry Association of America (RIAA) and individual record labels have had active technical efforts in marking rights information in digital recordings at least since 1980. Despite this strong interest, RIAA has used poor engineering in choosing technical standards for recording rights information. Today there is no widespread use of rights marking technology that would allow Napster to identify protected recordings. In analogous fields, such as image protection, such technology is widespread.

7. Napster can not tell whether a particular use of its system is infringing.

8. The use of ID/password mechanisms to allow or restrict access to a service such as Napster is reasonable and customary and is superior to use of IP source addresses.

9. A “bot” is a computer program that automatically performs Internet accesses. For example, a bot might monitor all the files available for download through Napster. The use of bots can result in significant load and performance degradation of an Internet service such as Napster, and thus are sometimes blocked for performance reasons.

Below I discuss underlying digital audio technology, before moving on to a discussion of each of my conclusions listed above.

Overview of Pertinent Technology – Digital Audio and Compression

Audio information can be stored in analog or digital format. Analog formats, such as conventional cassette recordings or long-playing phonographic recordings (LPs), record sound waves directly onto the recording medium. When a needle moves in the grooves of an LP record, it vibrates in a way directly corresponding to the structure of the sound waves. Unfortunately, analog recordings are subject to gradual degradation over time – as the recording medium suffers physical wear, increasing amounts of noise or distortion may be heard in the recording.

Digital recording formats, such as compact discs (CDs), Digital Versatile Disks (DVDs), Minidiscs (MDs), digital audio tapes (DAT), and MP3 files address this weakness by storing information digitally – as a series of bits. Different media store this information in different fashions: CDs use optical storage, MDs use magneto-optical storage, DATs use magnetic storage, and MP3 files are stored using underlying memory technology associated with a computer or playback device. These media themselves may suffer physical wear and lose bits, but they usually include error-correcting codes that allow the playback device to recover lost bits. In normal usage, one expects the signal to remain considerably more stable than corresponding analog formats.

Digital recordings sample and record the sound waves of the audio source many thousands of times each second. The playback device reads this information (restoring damaged bits of information from error correction as necessary), and goes through a “digital to analog” stage to create electrical signals that can ultimately be amplified and played through speakers.

Digital recording is more effective than analog recording at making a permanent, non-degradable record of audio information, but it requires large amounts of storage.

Fully loaded conventional CDs contain approximately 650 million bytes (or 5 billion bits) of information. To help conserve the use of memory, several modern advanced formats use digital compression technology. Digital compression technology allows audio recordings to be stored using smaller amounts of memory. For example, by using compression, Sony's Minidisc format can store audio information on a disk that is far smaller in physical dimension than a normal compact disc.

There are two types of compression: "lossless" (or "perfect") compression and "lossy" compression. Lossless compression takes advantage of certain common patterns in the source material to store a digital signal in smaller space. When the signal is recovered, it is expanded to the full original source material. While perfect compression has some interesting features, it still usually results in rather high storage requirements. These storage requirements are often too large for a variety of portable and computer applications, since computer memory devices usually have a predetermined, fixed amount of memory. To make audio storage on computers and portable devices practical, further reduction in storage requirements are desirable.

Reducing storage does not only result in savings for memory. To transmit the data requires time, and by compressing data, transmission time can be significantly reduced.

Lossy compression can significantly reduce storage requirements that lossless compression imposes. Lossy compression takes the original source material and reduces it by eliminating some features in the original source. For example, the developers of a lossy digital audio compression scheme will develop a "psychoacoustic" model to estimate sound volume levels (called "noise levels") that are believed to be imperceptible

by human listeners as distinct sounds. These noise levels are dependent on the total sound picture. For example, many people can easily hear a distant cricket in a quiet meadow. Place the listener and cricket next to an active airport runway and the cricket may no longer be perceivable as a sound when airplanes are taking off and landing. The sound of the cricket is said to be “masked” by the sound of the airplane. By exploiting these types of psychoacoustic properties lossy compression can ignore certain portions of the signal from the source material when storing data. This allows for significant reductions in storage requirements. When the signal is decompressed, it will not be a perfect copy of the original digital source. However, if the compression scheme is well designed, a human observer should perceive the expanded signal as being quite similar to the original source material.

There are a variety of schemes proposed for lossy audio compression. One leading group investigating lossy compression is the Motion Pictures Expert Group (MPEG). The Group has published the MPEG 2 standard to allow video data to be digitally compressed, and this standard is used today in DVDs. Since video and film may contain a soundtrack, the Group has also investigated audio lossy compression techniques, and has proposed several approaches, including MP3. It is important to note that while MP3 produces adequate reproduction of sounds for many purposes, because it uses lossy compression, it necessarily results in loss of sound quality from original source material. The following review from the August 1999 issue of Computer Audio World is typical (see the full article at <http://www.hi-fiworld.co.uk/caw/cawreviews4.html> ; footnotes are added by me and not in the original article). The gist of the article is that

MP3 reproduction is good, but still has perceivable weaknesses in the quality of audio reproduction, even compared to commercially available devices such as MD players:

“Given a decent CODEC¹, slow speed copying and jitter correction, high bitrate MP3 can sound surprisingly good. MiniDisc is a good yardstick. First generation ATRAC² MDs sounded poor - flat, uninvolved and processed. Third generation ATRAC added musicality and pep, whereas fourth and fifth generation MD are impressively natural. By comparison, 128kBit MP3 lies close to third generation ATRAC MD. Although a little more artificial sounding, it's certainly lively and listenable with strong rhythmic drive and clean tonality. For those used to cassette, MP3 done properly can be impressive.”

While the degradation of audio quality associated with MP3 does not matter to some listeners, it does to others. As superior formats, such as the new DVD-audio format or Sony's Direct Stream Digital format become more widespread, MP3's audio degradation will matter increasingly to listeners.

Today a variety of vendors (including Sony and S3/Rio) have produced portable devices for storing and playing back MP3 audio data. To use these devices, a user must download audio information onto the MP3 storage devices. Here is a typical scenario: suppose a user wants to listen to a song through a personal audio device while she jogs. If she uses a portable CD player, she would need to carry the portable CD player as well as a disk with the recording. Many portable CD players can not handle the shocks and vibrations associated with jogging, so she may want to store the song on some other medium. There are a number of very compact portable MP3 players that are highly shockproof, so she may wish to compress and copy the audio information into the memory of an MP3 player. There are two approaches she may use to compress the information:

¹ A CODEC is a compression enCOder/DECOder, in this case for MP3 files.

² ATRAC is family of compression algorithms used to compress data on MDs.

- She could use widely available “ripping software” to make her own copy. (A partial listing of available ripping software for Windows can be found at <http://software.mp3.com/software/all/windows/ripper/>). This is often done on a computer – she places her CD in the computer, and the ripping software produces an MP3 file.
- She could obtain MP3 files that are already compressed and download those into her device. There is a wide variety of MP3 data on the Internet. Some artists allow their songs to circulate freely on the World Wide Web. Indeed, the Web acts as a distribution medium for many artists who do not have access to or choose to disassociate themselves from distribution through traditional channels for selling recordings. In other cases, material that the copyright owner has not explicitly authorized for Internet distribution is nonetheless distributed over the Internet. Napster is one of many tools that allows users to search for other recordings online. Napster allows a user to search for other individuals that have MP3 music files designated for sharing. Each MP3 file is described by a file name. This file name contains a shorthand that may indicate the nature of the recording. The user requesting the file will select it and download it from the server to the portable MP3 player.

Conclusion 1: Napster allows users to share recorded music

- *Napster allows users to share computer files that contain recorded music. In allowing users to reproduce music it may be compared to a number of consumer products including:*
 - *cassette decks;*
 - *dual dubbing cassette decks;*
 - *cassette radios;*
 - *CD/cassette combinations;*
 - *video cassette recorders (VCRs);*
 - *Minidisc recorders;*
 - *digital audio tape recorders;*
 - *Tivo and Replay digital hard disk video recorders;*
 - *CD burning software;*
 - *encoding or “ripping” software that takes source audio material and converts it into computer readable form;*
 - *MP3 portable storage devices that allow computers to download computer readable audio files for portable playback including MP3 portable storage devices using memory sticks, compact flash memory, and smart media memory;*
and
 - *MP3 playing software.*

Napster gives users a way to share music. Napster does this using the Internet, but the fundamental idea of sharing music is not new. There are a variety of consumer-oriented commercially available systems that allow users to share music:

- Cassette recorders, which were introduced in the 1960s, have long given users a convenient way to record sound material for later playback or sharing. By connecting a cassette recorder to other audio equipment such as a radio, LP player, CD player, or MD player, a user can easily make a recording. Dual dubbing cassette decks take conventional cassette tapes and provide a way for users to quickly duplicate the tape onto a fresh cassette. These devices permit individuals to share music from source materials. Note that the primary purpose of a dual dubbing cassette deck is to reproduce a cassette recording. Today there are a number of dual dubbing cassette decks available as consumer products in the United States, and their popularity testifies to the popularity of music sharing. Common experience likewise suggests that for many consumers the primary purpose of a single cassette deck is to reproduce musical recordings for sharing or later play back.
- Cassette/radio combinations offer the ability for a user to tape broadcasts off the radio. Products of this genre typically allow users to do this with a single button push or with a small number of button pushes. The ability to tape music directly from the radio is an important element of these devices. These devices allow a user to record a song for later listening or sharing, and taping provides a means for users to capture, listen, and share music that is not easily available. This may be appealing to individuals with limited means (e.g., the twelve year old music aficionado) or to individuals who wish to collect music that is difficult to obtain otherwise (such as live

broadcasts or broadcasts of music not widely distributed in the United States, such as Indian classical ragas). I personally recall cassette radio combination equipment being available in the 1970s. Even before modern cassette tapes, reel-to-reel tape decks with line-level input were available, and before that, acetate disk recording devices were available. My understanding is that much of our current legacy of recorded jazz music and old-time radio programs from the mid-century was made from radio broadcasts.

- CD/cassette combination equipment allows users to easily tape music from CDs for sharing. Equipment varies from portable “boombox” devices to more expensive “minisystems” that purport to provide high quality audio. I own a minisystem produced by Denon that includes a specialized feature for synchronizing the recordings of CDs onto cassette tapes. This feature ensures that a song is not divided between two sides of a cassette tape. CD/cassette recording devices provide a popular way of sharing music, and indeed, my local used CD store offers to buy back any used CD sold within a week of purchase at 75% of the purchase price, ideal for those who wish to conduct home taping. The idea of creating and sharing “mix tapes” with songs from a variety of sources arranged in a thematic or aesthetically appealing fashion has achieved currency as part of popular contemporary culture. And the presence of 74 minute long cassettes (the approximate maximum length of a standards-compliant compact disc) argues strongly that copying from CDs is recognized as one of the primary functions of cassette tapes.³

³ The CD standard specifies a maximum audio length of approximately 74 minutes. However, since the CD standard was published, some record labels have published non-compliant CDs with up to approximately 80 minutes of recording time. Please note that 80 minute long cassette tapes are also commercially available.

- Video cassette recorders (VCRs) allow users to record a video image from source material (such as a television broadcast). When a user sets his VCR to record a concert or other musical event to watch later or exchange with friends, he is sharing music. Similarly, recording devices such as Tivo or Replay record video directly onto a hard disk for later playback. This supports time shifting and provides a consumer-oriented digital storage format for video.
- Minidiscs (MD) and Digital Audio Tape (DAT) provide consumer oriented formats for digital storage and trading of music. Many of these products feature optical or cable links for directly receiving synchronized digital signals from a source CD player. These systems often feature protection using the Serial Copy Management System (SCMS) which purportedly prevents multigenerational copying of source material. In practice, it is technically easy to disable this protection in practice by playing recordings mastered with SCMS to generate an analog signal, and then digitally re-recording them to eliminate the SCMS indicator. A second approach for bypassing SCMS involves using a professional dual dubbing MD or DAT unit that allows one to explicitly turn off SCMS (such as Denon's Dual MD recorder/player). Despite the term "professional", these items are freely available from a number of mail order and Internet vendors for prices as low as about \$2000. A third approach involves some skill with electronics – a person with some electronics training can easily build a device that bypasses SCMS; the plans are fully available on the Internet, see, for example <http://www.stack.nl/~leon/scms/>.
- Some of the most innovative products use general memory devices also used in digital cameras. For example, Sony has introduced the Memory Stick Walkman

which uses the same sort of memory stick in both MP3 players and digital cameras (see

- <http://www.world.sony.com/Electronics/MS/products/index.html> and
- www.sel.sony.com/sel/consumer/ss5/car/networkwalkamrtm/memorystickrtmwalkmanrtmdigitalplayer/nw-ms7_specs.html).

Other MP3 devices use compact flash or smart media memory, both memory devices commonly used in digital cameras (see

<http://www.mp3shopping.com/english/memory.htm>). Sony's NW-MS7

<http://www.ita.sel.sony.com/products/vmc/index.html> is advertised as allowing one to “Log-on and download ATRAC3, MP3, or WAV files from your favorite music web sites”.

- CD recorders are now common in many consumer oriented PCs, and some PCs have both a conventional CD reader as well as a CD recorder. For example, the electronics retailer Circuit City maintains a web site (<http://www.circuitcity.com>) where they currently advertise two computers that feature both a CD reader and a separate CD reader/writer: The HP 8655C for \$999.99 and the HP 8665C for \$1249.95. Looking closely at the details of the former, less expensive machine, one finds that the description clearly envisions the proposed purpose of the machine (ellipsis in original):

“The Hewlett® Packard 8655C Desktop Computer gives you the power of an Intel® Pentium® III processor 533MHz, CD-RW Drive, 40X Max. CD-ROM Drive, and Intel® Direct AGP 3D graphics, just to name a few. Imagine... you can have all this in the comfort of your own home. You'll be able to surf the Internet, create and play your own music CDs, and relieve some stress with your favorite 3D video game. The included MusicMatch software also allows you to download and play a wide variety of digital music.” [sic]

Both models include the MusicMatch Jukebox4 software that “Allows You To Download And Play a Wide Variety Of Digital Music and Create Your Own Customized Playlists.” The HP 8670, retailing at a suggested price of \$1,599.00 is described at

http://hp-at-home.com/datasheets/datasheet1.cfm?model_number=8670C&country=us

as being

“Your Personal Music Machine: The HP Pavilion 8670C PC has serious musical talent. Among the cool musical features is the HP CD-Writer Plus, an integrated music tool that delivers a versatile set of features. It enables you to make your own customized audio CDs by pulling songs straight from the Internet. Then, you can store, organize and play those CDs right from your PC. The HP Pavilion 8670C PC also comes equipped with jukebox software from MusicMatch, allowing you to download and play a wide variety of digital music and create your own customized playlists.”

Features designed to copy CDs apply equally well to software packages. For example, Adaptec’s Easy CD Creator 4 Deluxe is described in Adaptec’s datasheet (<http://www.adaptec.com/products/datasheet/ecdc.html>) as featuring a software component called CD Spin Doctor that promises that it “removes hisses, pop and clicks from your favorite old records and cassettes.” It also features a whole CD to CD copy product called Disc-at-Once CD Copier. A warning notice at the bottom of the page states (emphasis in original):

“THIS PRODUCT OR SOFTWARE MAY BE DESIGNED TO ASSIST YOU IN REPRODUCING MATERIALS IN WHICH YOU OWN THE COPYRIGHT OR HAVE OBTAINED PERMISSION TO COPY FROM THE COPYRIGHT OWNER. UNLESS YOU OWN THE COPYRIGHT OR HAVE PERMISSION TO COPY FROM THE COPYRIGHT OWNER, YOU MAY BE VIOLATING COPYRIGHT LAW AND BE SUBJECT TO PAYMENT OF DAMAGES AND OTHER REMEDIES. IF YOU ARE UNCERTAIN ABOUT YOUR RIGHTS YOU SHOULD CONTACT YOUR LEGAL ADVISOR.”

In addition to CD copying through a computer, there are stand-alone “dual dubbing CD reader/writers” that allow an existing CD to be copied onto a blank CD. For example, Philips manufactures at least two models: CDR765 and CDR870.

- Ripping software (or encoding software) allows users to produce MP3 files from source material (such as a CD). Ripping software is common on the Internet and is available from links from web sites as varied as Microsoft.com to Sony.com.
 - For a partial listing of Macintosh based rippers, see <http://software.mp3.com/software/all/macintosh/ripper/> ;
 - for Unix rippers, see <http://software.mp3.com/software/all/unix/ripper/> ;
 - for Windows rippers see <http://software.mp3.com/software/all/windows/ripper/> (as mentioned above);
 - for rippers for other platforms, see <http://software.mp3.com/software/all/other/ripper/> .

MP3 software will allow one to play MP3 files directly over the speakers of a computer enabled for audio. One of the most popular MP3 software players is “Winamp” distributed by AOL, which in June 1999 bought its creator company, Nullsoft. AOL states on their corporate web site (<http://corp.aol.com>): “Digital Music Reaches Another Milestone As AOL's Winamp Celebrates 25 Million Registrants.” AOL features a web page with MP3 related software (<http://www.aol.com/webcenters/computing/multimedia.adp>) and points from that

page to an even more comprehensive C-NET web site containing a wide variety of MP3 related software.

It is important to note that Internet technology often facilitates music sharing from items listed above. A perusal of the musical discussion groups on the Internet wide bulletin board system known as “netnews” will yield a variety of music offered for trade or sharing in a variety of media: conventional cassette tapes, MDs, DATs, CDs, and MP3 files. Netnews predates the World Wide Web, but there are now some portals, including <http://www.deja.com> that provide access to netnews from the Web. Specialized consumer equipment facilitating sharing of musical recording has been popular long before Napster came on the scene.⁴

⁴In plaintiff’s “Notice of Joint Motion and Joint Motion of Plaintiffs for Preliminary Injunction”, a distinction is drawn between a product and service (paragraph beginning on page 23, line 15.) This is a distinction that is not widely accepted in the software field. The line between services and products is fairly confusing. For example, Intuit sells something called TurboTax and another thing called TurboTax for the Web. These share similar functionality and purpose, but one runs on the user’s computer and the second runs on a server reached over the Web. They share similar product names and family heritage, and are advertised together. Trying to call one a product and the other a service is not meaningful. Similarly Encyclopedia Britannica sells book copies of its encyclopedia but also makes the exact same material available over the Web. It is hard to see in this case how one is a service and the other a product. Even if one were to draw an artificial distinction based on where software was running, it is not clear how Napster would be classified. To run Napster, users download a special program and run it on their machines. To this extent, Napster has a number of qualities one would normally associate with a “product.”

Conclusion 2: Napster allows users to share MP3 files

- *Napster allows users to exchange and share MP3 files. In this way, it is similar to existing file sharing techniques including:*
 - *electronic mail (e-mail);*
 - *the World Wide Web;*
 - *the File Transfer Protocol (FTP);*
 - *search engines (such as Lycos or AltaVista);*
 - *caching search engines (such as google.com);*
 - *Gnutella;*
 - *Freenet; and*
 - *SpinFrenzy.*

Napster serves as a search engine that allows users to identify links to other servers that are willing to share files. Napster is not unique in offering Internet file sharing or searching capabilities. File sharing has long been a central function of the Internet, and there are long lists of technologies that support it. Perhaps the most common form of file sharing is e-mail. E-mail allows a user on a computer attached to a network to send a message – a file – to another user. Modern e-mail encoding systems permit the use of attachments that directly copy files for sharing among users.

The World Wide Web itself is a file sharing mechanism – every time one visits a web page, one is sharing the file that contains the contents of that web page. The Web is designed to be able to share arbitrary types of files. In particular, MP3 files are popular for sharing on the web, and according to a widely publicized 1999 report, “MP3” is said

to have replaced “sex” as the most popular search term on the World Wide Web (see, for example, <http://www.wired.com/news/mp3/0,1285,31834,00.html>).⁵ Before the WWW became such a dominant file sharing mechanism, there was an older technology known as the File Transfer Protocol (FTP) and a variety of search mechanisms for using FTP including Archie and Gopher (which acts as a front end to Wide Area Information Servers).

Search engines that focus on indexing web material include yahoo.com, altavista.com, and lycos.com. Some search engines, such as google.com, record a copy of web pages on their own server allowing access to the information even if the owner of a web page disconnects his web server. Much material on the World Wide Web may be the subject of copyright, including copies of lyrics of apparently copyrighted songs. Yahoo has a page that points to hundreds of such web sites, including some sites that purport to be comprehensive collections:

http://dir.yahoo.com/Entertainment/Music/Lyrics_and_Notation/Lyrics/

To fully grasp the number of lyrics referenced off this page, one needs to follow all the links. For example, 17 distinct sites are referenced that claim to offer the lyrics of the musical group Metallica⁶:

[http://dir.yahoo.com/Entertainment/Music/Artists/By_Genre/Rock_and_Pop/
Metal/Metallica/Lyrics/](http://dir.yahoo.com/Entertainment/Music/Artists/By_Genre/Rock_and_Pop/Metal/Metallica/Lyrics/)

Here is the claim of one of those pages referenced, “Mach5’s Complete Metallica Lyrics Page” (<http://www.oe-pages.com/ARTS/Rock/mach5224/indexframe3.shtml>):

⁵ In fact, the search engine site AltaVista (<http://www.altavista.com>) features a special MP3/audio search page directly off its home page.

⁶ NB, because of the length of the URL, I am forced to break it between two lines – it should be read as a single line.

“Hey all you Metallica fans! This is Mach5 coming to you from Lancaster, Pennsylvania. Enough said. I created this page so that you can find *ALL* the lyrics to *ALL* your favorite Metallica songs! I worked long and hard to get all these lyrics on here and I hope you appreciate it. I have every album on here with every lyric for every song, including the newest and highly anticipated, S&M ALBUM!!!!!!!!!!!!!! THE FIRST ON THE WEB!!!!!!!!!! Please try to get around to signing my guestbook if you have the time! Also, submit a vote on my poll! I just finished updating my S&M lyrics pageso everything is perfect now!” [sic]

The World Wide Web uses a “client-server” model of computation, which means that information is offered through explicit web servers. In contrast, Napster uses a “peer-to-peer” model in which users connect directly to each other and exchange files. There are other utilities that offer peer-to-peer functionality similar to Napster such as SpinFrenzy (<http://www.spinfrenzy.com>), Gnutella (e.g., <http://gnutella.wego.com/>) and FreeNet (e.g., <http://freenet.sourceforge.net/>). Note that since Gnutella and Freenet are widely distributed technologies and not products maintained by a single company, it is hard to imagine how their use could be controlled. Clearly, no single organization or small set of organizations could act as control points for those technologies.

While Napster’s initial focus is on MP3 files, I understand it supports Windows Media Audio (WMA) files as well and using a freeware package known as Wrapster (see <http://notoavian.tripod.com/>), one can exchange arbitrary files. Gnutella and FreeNet both offer general file sharing mechanisms. The ability to exchange arbitrary files supports functions that are universally accepted as important for society. For example, file sharing is vital to contemporary scientific research. As a researcher in computer science, I can attest to the importance of file sharing of research reports and technical reports to scientists. Another example mentioned by media reports is given by Lincoln Stein, a researcher at the Cold Spring Harbor Laboratory, who is investigating file

sharing mechanisms as a way of exchanging research information about the Human Genome Project: <http://www.laweekly.com/ink/00/19/cyber-heyman.shtml> .

Similarly, it is now common to use shared files to distribute important information to those who are interested. For example, user manuals are often distributed on the World Wide Web. One company that has been outstanding in using file sharing to distribute user manuals is Sony, which distributes a large number of its technical and user manuals freely over the Web. One interesting example of such a manual being distributed over the Web is Sony's Vaio Music Clip manual:

<http://www.ita.sel.sony.com/support/musicclip/VMCMAN.pdf>

which, incidentally, is a device that supports playing of MP3 files as well as Sony's "secured music format". This manual contains a copyright notice on the cover page. However, since Sony points to this work from its Web pages, it presumably wishes interested consumers to be able to freely download this page and print it out, despite its copyright notice.

Based on my lay reading of section 512(a) of Title 17, I believe that Napster qualifies for an exemption for liability and relief for infringement under the terms of that section. Napster provides service where (1) a party other than Napster initiates service; (2) Napster does not screen the material but provides connection and routing information automatically; (3) Napster does not choose the recipients of the material; (4) the copies lie solely on the two peers exchanging files – Napster does not maintain a copy of the file; and (5) the material is transmitted through the network without modification of content. Napster is a typical example of an Internet intermediary that allows communication between various individual parties.

From a technical standpoint, the idea of trying to control the distribution of MP3 files by restricting file sharing seems odd. File sharing is basic to the operation of distributed computers and the Internet, and is used for large numbers of legitimate and valid reasons.

File sharing is one of the oldest concepts in distributed computing, and the ability to share files among users through e-mail or file transfer protocols dates back to the 1970's origins of the ARPANET (the Department of Defense's Advanced Research Project Agency's Network, widely considered as the principal predecessor to the modern Internet). In the 1970s, a number of universities and research organizations used ARPANET to actively exchange files.

Moreover, file sharing can be done in a large variety of ways. Even though file sharing is one of the oldest concepts in distributed computing, it continues to surface in new and innovative ways. As new applications are added to the Internet, they often have file sharing components to them. For example, instant messaging systems (as exemplified by AOL's Instant Messenger or ICQ) include options for file sharing. Similarly, a number of different types of chat programs include options for file sharing.

If one wished to ban the sharing of MP3 files altogether (in my opinion, this would be undesirable both from a policy and technical perspective), the best technical approach would be to attempt to control MP3 players. While there are a number of MP3 players, they are designed primarily for the purpose of playing MP3 files, so attempting to restrict them would have a more focused impact than attempting to regulate the far more general technique of file sharing.

Conclusion 3: Napster does not have access to information about copyrights

- *Napster can not distinguish between:*
 - *material protected by copyright and restricted by the owner;*
 - *material protected by copyright but for which the owner or the law permits free distribution; and*
 - *material not protected by copyright.*

Furthermore, to require Napster, any other search engine, or any file sharing utility to obtain an authorization from the rights holder prior to providing access to its material is technically infeasible and would prevent the effective operation of the utility.

Napster has no way to tell the copyright status of any shared file, unless that information is explicitly declared to Napster. Given a sound clip, there is no practical algorithm for determining whether that sound clip is copyrighted or not.

At first glance, one might think of two possible techniques for distinguishing copyrighted materials. The first would use file names. The idea would be to build up a database of all file names corresponding to copyrighted recorded music (a truly Herculean task) and check files to be shared against that list. Unfortunately, this approach is doomed to failure. In the first place, one would need to create a comprehensive list of all recordings in which copyright is claimed. Next, file names are, at best, only a mnemonic chosen by the person offering the file for sharing. To anticipate all possible file names that cover copyrighted material would not be possible.

Suppose, for example, that user Jane Doe uses Napster to allow her to access her recording collection in a variety of locations (space-shifting) and that she is a fan of aboriginal African folk song. What would one make of a file name such as “Jane Doe/favorite Ituri chant.mp3”?

Ambiguity reigns supreme in file names – what does BS mean in an MP3 file name, for example: Britney Spears? Boz Skaggs? Bruce Springsteen? Barbra Streisand? Black Sabbath? The Backstreet Boys? The Boston Symphony Orchestra? Or a vulgar comment on the lack of artistic merit of the recording? A casual user will quickly discover the limitations of file names using Napster by referencing, for example, classical recordings.

Identification of pieces is often difficult, and file names often do not contain information about the ensemble, conductor, or date of the recording.⁷ Even if Napster tried to use file names to distinguish copyrighted material, users would quickly learn to work around them. If all files with file names containing the word “Metallica” were screened out, users might try referring to the musical group by the misspelling “Metalica”. (A casual search by me on Napster indicated a large number of hits on the misspelled name “Metalica”.) Indeed, some individuals actively advocate using misleading file names. The web site <http://www.stopnapster.com> advocates sharing sound files with misleading names; in particular, see:

⁷ In view of this shortcoming, some individuals have created a set of unofficial additional fields for MP3 files collectively known as an ID3 field (see, for example, <http://www.id3.org>). If ID3 information is included in an MP3 file, it is normally included at the time the MP3 file is created. ID3 information can help identify the source of a recording, but like file names, may contain information that is ambiguous, misleading, or incorrect. ID3 files may also contain information only meaningful to the creator of the MP3 file (again, consider the example of the user who uses Napster to space shift her music).

<http://www.stopnapster.com/trojans.html> ,

<http://www.stopnapster.com/bombs.html>

The latter web page contains the following (ellipsis mine):

“Just think of the reaction you'll get from users who think they're downloading the new Beastie Boys track but instead get four minutes of dogs barking, sirens going off, etc Basically, says Gunderson, a Napster Bomb is an intentionally mislabeled file masquerading as a song by a major artist, which is clearly protected by state and federal copyright laws.”

A second potential approach would use checksums (also called hash values) to distinguish copyrighted files. A checksum is produced by a mathematical algorithm and is an attempt to create a “fingerprint” of a file. Two different files will, with high probability, yield different checksums. The problem is that two recordings that are clearly from the same source material may have different representation in files, and thus different checksums.

Depending on how the song is converted to an MP3 file – what degree of compression is used, whether the song was put through an analog stage (which adds noise to the recording), how the song was mixed and mastered/re-mastered, the exact start and stop points of the recording, and the device or software used to create the MP3 file – one will end up with completely different files. Different files yield (with high probability) different checksums.

Napster has no way of identifying a particular file as corresponding to a particular recording. Consider music that has only been released in analog form. Every time an MP3 file is made of the music, it will yield a unique file. Similarly, personal experience has taught me that purportedly identical recordings released in different countries or even

by the same company within a country often have a different mix and thus would yield different MP3 files. (For example, I have heard a number of CD recordings of Miles Davis's well-known album "Kind of Blue" that have dramatically different audio characteristics. Similarly, I have heard multiple copies of Glenn Gould's 1956 recording of the Goldberg Variations with dramatically different audio characteristics.) Since the files are different, there is no way that Napster can automatically check whether a recording is on an approved list or not.

A dramatic illustration of this is contained in a 3 May 2000 letter from attorney Howard King to Sean Parker at Napster (Bates number NAP008871-90). The letter notes a number of items that are claimed to be improperly copied from Metallica recordings. The number of such items is 1,456,075. In a second letter dated 18 May, King gives a list of 2,280,474 items with a total of 470,846 distinct checksums (in this case, using MD5, a checksum algorithm developed by Professor Ronald Rivest of the Massachusetts Institute of Technology). This is clearly far larger than the number of recordings Metallica actually has issued.

Even if Metallica were able to create a comprehensive, error-free list of all checksums corresponding to files that have appeared on Napster to date, every time a new copy of a recording was ripped from source material, it would, with high probability, possess a completely new checksum. As a practical matter, the number of slight variations in potential copies in Metallica recordings would make checksums inappropriate as a way of identifying the list of all possible Metallica recording copies.

It is interesting to note that two of the items contained in the Metallica list entitled "Top 100 Distinct Digital Recordings" (Bates number NAP008875) appear to be the song

“Nothing Else Matters” performed by Chris Isaac, who I understand is not a member of the musical group Metallica. I also note a number of the files on the top 100 files are indicated as being live recordings. Since King explicitly states in his cover letter that “Metallica makes no claim of copyright infringement with respect to recordings of their songs made by fans at Metallica live concerts” one is led to question whether these live recordings represent bootleg copies for which Metallica does not claim infringement. Certainly one can not tell from the file names whether Metallica claims these files to infringe or not.

A third approach would be to assume that all recorded music is protected by copyright and require Napster to collect pre-authorization for songs it distributes. Suppose one were to make the assumption that all recorded music is protected by copyright.⁸ Under this assumption, one would need to collect some sort of clearance from a rights holder.⁹ But the recordings themselves would be offered on servers that in general are not necessarily run by the rights holder. This would mean that Napster would need to check that offered files matched recordings for which clearance was held. For reasons discussed above, this would be technically infeasible if file names or checksums were used to identify material. The alternative seems to be to have a human being check the recording and make sure it matches a recording on the approved list. As the list of recordings that are authorized for Napster distribution would grow, they would exceed human memory capacity and could not effectively be identified. This is doubly true for classical recordings – how many people can listen to an arbitrary recording of a warhorse

⁸ Note that this assumption is false – for example, I understand recordings such as 1940’s V-Disk recordings made for the US Military by popular artists of the day are explicitly claimed to be in the public domain. Similarly, I understand that many government recordings are in the public domain. I further understand that many early recordings have had copyright expire and are now in the public domain.

symphonic piece such as Tchaikovsky's 1812 Overture and state with certainty the conductor and orchestra performing the recording? Indeed, evidence of the difficulty of identifying pieces is given through a regular column in *Downbeat*, a magazine for jazz musicians and enthusiasts. For decades, *Downbeat* has featured a "blind" test where a well-known, highly talented musician listens to a number of musical selections and is asked to identify them. Musicians often fail miserably (and in interesting ways), making this feature consistently entertaining. If professional musicians can not accurately identify the source of pieces in genres in which they work day in and day out, what chance would a "Napster human checker" have?

Even if accurate identification of the source of a recording were apparent to a listener, it would be straightforward for a user to modify his server to offer a different recording than the one advertised on Napster. For example, perhaps a user has a submission checked as being an "authorized for distribution" recording of a Fireside Chat of President Franklin Roosevelt. But once the submission is checked, the user modifies his server to present a track from the musical group Metallica in place of Roosevelt's speech.

⁹ This raises a natural question of how one would identify the party presenting the clearance as actually being the rights holder. I discuss this question in Conclusion 4 below.

Conclusion 4: Napster can not check authorizations

- *Napster has no practicable way of checking that an authorization is from the party it purports to be from or that party holds any or all rights to any particular file.*

If parties were to submit authorizations to Napster, how could Napster check that the party purporting to submit the song for reference by Napster's engine actually had approved the use of Napster? Clearly, submission of a copyright registration notice would not suffice; I understand that copyright registration forms are public documents and copies are available to any interested party. One could easily imagine a scenario where an ordinary individual might falsely represent himself to be a representative of the musical group Metallica. This individual could submit a Metallica recording, complete with copyright registration, to Napster with a statement that the recording was approved for distribution. It is not clear how such an individual falsely presenting himself as a Metallica representative could be distinguished on the computer from a true representative. Clearly, a return e-mail address would not suffice since anyone can sign up with electronic mail services such as Microsoft's hotmail.com and, if they are the first party to claim the user name, claim a user name of, for example, james_hetfield@hotmail.com. (I understand James Hetfield is a member of the musical group Metallica.)

Conclusion 5: Authorization would change the Web to a centralized utility

- *Even if there were a way to check authorizations for all files that would be shared under Napster, the World Wide Web, or any other file sharing utility, the effect of requiring authorizations would change the utility from a decentralized, ground-up information base to a centrally controlled top-down distribution device.*

The World Wide Web and the Internet are revolutionary because they turn traditional information distribution methods on their head. For decades, if not centuries, conventional distribution of information has been dominated through a set of intermediate publishers and distributors. However, on the WWW, anyone can publish any material and have it be instantly available to all WWW users. The WWW consists of millions of computer users who post material on their own servers or servers operated by any of a very large number of third parties. Moreover, there is no delay in publication; the moment that information is updated on a Web server, it is available for access by everyone on the WWW. The WWW transcends national boundaries and traditional publication boundaries. There is no central authority or intermediary who approves or keeps track of material that is posted on the Web or the Internet.

In particular, the technical standards for the World Wide Web are governed by an engineering group known as the Internet Engineering Task Force (IETF). The IETF oversees the technical development of the World Wide Web making a number of particular technical decisions.

A system which required each file or web page to be pre-authorized would not fit with this model. It would add delay to publication. It would act as an official gatekeeper. Such a model would completely change the nature of the Internet. But more seriously, it

would change the underlying technical model of how the WWW works. The performance model of the Internet would change, almost certainly for the worse, and probably with severe technical difficulties. It is not even clear that the World Wide Web could technically continue to function in such a top-down model.

Conclusion 6: Watermarking could carry rights information with a recording

- *The Recording Industry Association of America (RIAA) and individual record labels have had active technical efforts in marking rights information in digital recordings at least since 1980. Despite this strong interest, RIAA has used poor engineering in choosing technical standards for recording rights information. Today there is no widespread use of rights marking technology that would allow Napster to identify protected recordings. In analogous fields, such as image protection, such technology is widespread.*

In 1980, the Recording Industry Association of America wrote to 38 universities and research centers across the US, requesting a technical solution to home taping. In 1982, the CBS¹⁰ technology center in Connecticut introduced “Copycode”, a system for protecting recordings by introducing an audio notch. An audio notch represents a range of audio frequencies that are reduced in volume. Recording devices could check for the notch and prevent the material from being recorded. CBS and the RIAA initially proposed a 250 Hz notch at 3.84 kHz that was 60 dB deep¹¹, although they subsequently modified their request by narrowing the notch by half.¹²

By 1987, RIAA was aggressively pushing for mandatory use of Copycode in US distributed DAT recorders. According to a news article published in the *Wall Street*

¹⁰ CBS's recorded music division was purchased by Sony Corporation in the mid-1980s and became Sony Music Entertainment.

¹¹ I am using standard notation here: Hz is hertz (a measure of cycles per second or frequency), kHz is kilohertz (thousands of hertz), and dB is decibels (a measure of the volume of the sound).

¹² A brief history of the early origins of Copycode can be found in 3 September 1988 *Daily Telegraph*, “Connected - Technoturkey: Copycats live to tape another day” by Barry Fox. Some material in this paragraph was paraphrased from that article.

Journal on 30 June 1987, “Japanese Consumers Are Slow to Acquire Nation’s Digital-Recorder Breakthrough” by Stephen Kreider Yoder (ellipses mine):

“Because the recorders can copy CDs and tapes, Western record companies are afraid the machines will make it easier to pirate songs. So they have pushed for modifications in the DAT recorders that would prevent them from duplicating recordings. The Japanese are resisting. . . . Still, some industry executives believe the outlook for DAT recorder sales may be improving. . . . Record-industry officials also are retreating from their initial insistence that DAT makers use a "spoiler" device developed by CBS Inc. to prevent record copying When Japan introduced DAT prototypes last year, the Recording Industry Association of America and other industry groups demanded that the Japanese place the CBS-designed spoiler device, named Copycode, in their machines. Otherwise, record companies say, music pirates would have a field day churning out near-perfect copies of commercial tapes and records. This year, several U.S. lawmakers sought to amend a trade bill to block the import of DAT players without spoiler devices. The International Federation of Phonogram and Videogram Producers wants similar restrictions in Europe.”

RIAA lobbying¹³ eventually yielded to tests by the Department of Commerce’s National Bureau of Standards (NBS, today called the National Institute of Standards and Technology). This article “U.S. Agency Rejects Device to Counter Digital Recorders: Tests Show Sony’s Copycode is Unreliable, Can Hurt Quality and Be Bypassed” by Jeffrey A. Tannenbaum in the 2 March 1988 *Wall Street Journal* summarizes the results of those tests:

“Copycode , an anti-taping system designed to cripple digital audio tape recorders, flunked government tests.

“The National Bureau of Standards, part of the Commerce Department, said the system is unreliable, sometimes hurts audio quality and, in any case, ‘can be easily bypassed.’

“The Home Recording Rights Coalition, a lobby favoring unrestricted use of digital recorders, said the government finding will ‘effectively kill legislation seeking to block the DAT (digital audio tape)

¹³ See *Wall Street Journal*, 24 August 1987, “Critique: Audio Wars, CDs, and ‘Slivers of Sound’ ” by Gregory Sandow.

format.’ The Recording Industry Association of America, which wants to restrict DAT, pledged to resolve the dispute through ‘negotiation, legislation or litigation.’

“Copycode initially was promoted by CBS Inc.'s records unit, now part of Sony Corp. of Japan.

“The Japanese-made DAT machines, already on sale in Japan and Europe, bring compact-disk quality to tape recording. Some owners of musical copyrights fear that the machines would lead to an increase in both home taping (which is legal) and commercial piracy (which is illegal), both to the detriment of recorded-music sales.

“These DAT foes had endorsed Copycode , which puts a special ‘notch’ in protected recordings so that machines can't copy them.

“But according to the government tests, Copycode sometimes works when it isn't supposed to, while sometimes failing to work when it is supposed to. Also, the notch can make a discernible difference in the recording, and five different methods can defeat the anti-recording device. Other anti-taping systems have been proposed, but they weren't included in these tests.

“Play-only DAT machines already are reaching the U.S. market, for use in cars. It still isn't clear when DAT machines that record may arrive.”

Somewhat surprisingly after this engineering disaster, the principal inventor of the Copycode scheme was reportedly rewarded with a senior vice-president post at RIAA, as reported in this article in the trade press (19 January 1996, Audio Week, “Audio Notes”); N.B. the telegraphic style of writing is from the original text):

“Father of controversial CBS Copycode copy-prevention system on DAT has been tapped as RIAA senior vp to head up newly formed New Technology Div. David Stebbings moves to RIAA after 14-year engineering stint at company formerly called CBS Records, now known as Sony Music Entertainment. Establishment of division ‘is an important move for us, and we’re confident that this new division will be of great service to our membership, as well as to the industry overall,’ RIAA Pres.-COO Hilary Rosen said. ‘We’ve been monitoring technological advances for years, but decided that the time had come to put a name on the division and to significantly expand its scope. It will serve to protect the copyrighted works of our record companies as they venture into the complex business environment of the future.’ Rosen praised Stebbings as ‘a renowned visionary, as well as an experienced, business-oriented engineer. As our association positions itself for the future, I feel sure that David will take us there with confidence.’ Statement didn't mention

Stebbing's work on Copycode, which reached height of activity in 1987 when bills surfaced in House and Senate to bar imports of DAT recorders that didn't contain Copycode circuit. Copycode was killed when independent analysis by govt.'s National Bureau of Standards (now National Institute of Standards & Technology) supported opponents' claims that technology degraded audio quality of music and didn't work as effectively as designed in barring digital dubs. RIAA Chmn.-CEO Jay Berman said no reference to Copycode was made in Stebbings announcement because his trade group considers that issue 'gone and forgotten.'" [sic]

After the engineering debacle of Copycode, RIAA still felt that the engineering challenges presented by the issue of copying were not difficult to technically address. In an October 1989 Office of Technology Assessment Report, *Copyright and Home Copying: Technology Challenges the Law*, contains the following information identified as being based on information from the RIAA (ellipsis and footnotes are mine, quote begins on page 58):

"To identify the range of technically feasible alternatives to prevent or limit copying, the information summarized below was provided in April 1989 by the RIAA Engineering Committee. Neither the Engineering Committee nor the RIAA intended the information outlined below as an endorsement of any particular system or approach.

"According to the RIAA Engineering Committee, copy-protection systems could be designed to prevent copying of prerecorded and/or broadcast material, to limit copying, or to allow copying with remuneration.¹⁴ Copy-protection systems of these types might be implemented in the analog domain, the digital domain, or both According to information provided by the RIAA Engineering Committee, efforts are ongoing to develop a system of this type."

The report goes on to briefly outline some the technical possibilities provided by these systems.

To make copy protection systems a reality, the RIAA contracted with the consulting firm Bolt, Beranek, and Newman to develop a better engineered system. An

excerpt from a 2 March 1996 article in the science news weekly *New Scientist*, “Noisy Dilemma Over How to Beat Pirates” by Barry Fox summarizes the BBN system:

“KEEPING track of its tunes is a major headache for the music industry, especially now that anyone can send recordings whizzing around the world with the help of a computer and a modem. The Recording Industry Association of America and the International Federation of the Phonographic Industry have kept quiet until now about their latest proposal for a technical solution to the problem. But *New Scientist* has uncovered a patent that gives the game away - and reveals potential flaws in the proposed security system.

“Until five years ago, the RIAA's favoured anti-copy system was Copycode - an idea developed by CBS, which is now part of Sony. But independent research showed that the system was unreliable and could spoil the quality of the sound. So the RIAA commissioned Bolt Beranek and Newman (BBN), an acoustics company based in Cambridge, Massachusetts, to develop a new system.

“An international patent application (WO 93/12599) filed by the company in 1991 describes the BBN process. The patent shows that the inventors bury a string of digital code words in the analogue music signal, like a watermark. The code represents the artist's name and the title of the music, and this watermark should always travel with the music, no matter how it is recorded or transmitted.

“To disguise this extra layer of noise, the code words are spread across a range of frequencies from around 1.9 to 10.7 kilohertz. To make this ‘spread spectrum’ signal less noticeable, the encoding equipment relies on a principle called psychoacoustic masking. It continually monitors the music and adjusts the level of the added noise to make sure that the music disguises the noise at any given frequency. The patent proposes that the code should always be 19 decibels quieter than the music.

“The result, claims the patent, is a ‘composite audio signal which is not readily distinguishable’ or is ‘essentially indistinguishable’ from the original recording.

“But BBN faces a problem because psychoacoustic masking is also the principle underlying most digital sound compression systems. The systems that squeeze sound over modems, transmit digital radio and permit digital consumer recording all do the same thing - they save on the number of bits by throwing away parts of the music signal that the ear will not hear.

“Here’s the catch. If record companies make the BBN code loud enough to survive compressed recording and transmission, it may spoil the sound of the original music. But if the code level is reduced to make it

¹⁴ The original text of the OTA report contains a reference to a sidebar box discussing these options in greater detail.

genuinely inaudible on the original, it may be lost when the signal is compressed for recording or transmission.

“To add to BBN's troubles, Sony demonstrated a system last week that is designed to deliver the lowest levels of noise and distortion ever. Sony expects its Direct Stream Digital system to be used initially as an archiving tool, making perfect copies of existing master tapes. The same system could be used with the forthcoming high density CD formats to deliver higher quality sound to consumers.”

By late 1999, RIAA had teamed with recorded music companies and organizations for other countries to form the Secure Digital Music Initiative (SDMI). A press release dated 26 February 1999 announces the kick-off this organization and quotes Cary Sherman, who is identified as being “senior executive vice president and general counsel of the RIAA.”

The SDMI quickly shot down the BBN proposal, as was reported in The Daily Telegraph on 23 September 1999 by Barry Fox in “Connected – Technoturkey: Copy Protection is a Game of Musical Chairs – Barry Fox on ill-fated schemes and dreams”.

Here is an excerpt from that article:

“The music industry and its many trade bodies continue to breed copy-protection turkeys.

“In the mid-Eighties, CBS invented CopyCode, which sucked identifying notches from the music. The Recording Industry Association of America, and the world trade body the International Federation of the Phonographic Industry, wanted laws to make recorders reject CopyCoded music. In 1988, a US government committee said it did not work - and spoilt the music. The RIAA abandoned notching and switched to an approach from acoustic consultants Boulton, Beranek and Newman.

“BBN worked on the noise smear principle. An encoder in the recording studio adds modulated noise which is always 19dB below the music.

“The RIAA claimed that psychoacoustic masking made the BBN code inaudible. ‘Far from being a problem,’ an RIAA expert assured, ‘high definition or high signal-to-noise ratio systems benefit from the proposed RIAA/BBN system.’”

“Earlier this year, the Secure Digital Music Initiative took on the job of testing systems for controlling music delivery over the Internet. The

4C consortium of Intel, IBM, Matsushita (Panasonic) and Toshiba also asked ‘golden-eared’ experts to choose an audio watermarking system that can be used to protect music released on the DVD-Audio disc.

“The SDMI/4C have now said ‘no’ to BBN.”

The SDMI Call for Proposals for screening digital content (dated 5 May 1999) can be found at¹⁵

- <http://www.sdmi.org/dscgi/ds.py/CheckOut/File-364/pdwg99050504-TransitionsCFP.doc>

Proposals were due by 23 May 1999. A summary of responses dated 30 May 1999 can be found at

- http://www.sdmi.org/dscgi/ds.py/Get/File-438/PDWG99053001R01-Summaries_of_Proposals.doc

Companies as varied as Microsoft to Sony responded to the call. Sony’s response, dated 22 May 1999 can be found at

- http://www.sdmi.org/dscgi/ds.py/Get/File-432/PDWG99052713-Sony_Corp_Proposal_to_Screening_CfP.doc

Sony’s response contains the following statement: “Sony has been engaged in development of high potential watermarking technology for several years and the system based on our watermarking technology will satisfy the needs of creators, record companies, service providers and consumers in sound quality, reliability etc.”

However, Sony’s system was not the one selected by the SDMI. The SDMI reportedly chose Aris’s “Musicode”¹⁶ as a transitional technology. By designating

¹⁵ The URLs (web page addresses) mentioned on this page are long, and they are thus split over several lines. Note that they are also taken from the “members only” portion of the SDMI web site.

¹⁶ See, for example “Questions and Delays Beset Digital Audio Watermarking” in 13 September 1999 *Audio Week*, or “Digital Watermark Chosen for DVD Audio and SDMI” in 16 August 1999 *Audio Week*.

Musicode as a “transitional technology”, the SDMI was able to collect a large number of viable proposals and select a good standard for music watermarking within just a few months of its being formed. SDMI allowed itself breathing room to develop further technologies. Later systems can use a later technology (referred to by the SDMI as a “Phase II technology”) but copy protection information using Musicode can be included in recordings starting now.

Musicode is an example of a watermarking technology. Watermarking records rights information in low-level bits in a digital recording. In a well-designed watermarking system, the additional small level of noise added by the watermarking signal should be imperceptible to the listener. By looking for the watermark, a player device or software could determine rights information associated with a recording. It could refuse to play information that had been improperly copied. (This would only apply to recordings which explicitly had the watermark added. It would not address “bootleg” recordings made by individuals during concerts.)

Watermarking technology has been known for some time. On 2 and 3 April 1993, I attended a workshop on “Technological Strategies for Protecting Intellectual Property in the Networked Multimedia Environment” at Harvard’s Kennedy School of Government sponsored by Harvard, MIT, the Coalition for Networked Information, and the Interactive Multimedia Association. During this conference Professor Kineo Matsui presented a paper on digital steganography, i.e., embedding information in an image. Matsui’s system explicitly addressed images, but during the discussion at the conference, the idea of extending this to other types of digital representations was discussed, including digital representations of audio. At the conference, the organizers distributed a paper co-

authored by Matsui and Kiyoshi Tanaka. In January 1994, this paper was re-published as part of the proceedings of the conference as Volume 1 Issue 1 of *The Journal of the Interactive Multimedia Association Intellectual Property Project*. I believe that similar work was published even earlier; Matsui references two of his 1990 works “Embedding the Attribute Information into a Dithered Image” (published in *Scripta Technica*) and “Embedding Secret Information into a Dithered Multi-Level Image” (published in the *IEEE Military Communications Conference*). The application of these techniques to digital audio recordings would be obvious to someone familiar with digital technology – indeed, dithering is a fundamental technology used not only in digital images but in digital audio.

Digital watermarking technology has been quickly adopted in other media. For example, Digimarc packages software with Adobe Photoshop for adding watermarks to images. This technology has been adopted by publishers who wish to protect pictorial content. For example, a 20 June 1997 press release reports that Playboy, a publisher of adult magazines and an adult web site, will use Digimarc to protect images (<http://www.digimarc.com/news/pr97-12.html>). The press release reports:

“Digimarc's new innovative MarcSpider™, the first service to search the World Wide Web for digitally watermarked images, enables Playboy to track images that have been re-posted on the Web. The MarcSpider crawls the Web, looking at hundreds of millions of pieces of information, locating Digimarc watermarked images and reporting back where and when they were found.”

In short, watermarking technology has been available for a decade. It is now so well understood that it has been the topic of several undergraduate senior theses. And outside of the field of digital audio, watermarking has enjoyed a fair success. It is true that sophisticated attackers can remove or obscure watermarks (and I believe that the Aris

system adopted by the SDMI is vulnerable to the same class of attacks). But against ordinary users, watermarking technology is effective at keeping digital content safe. If this technology had been adopted by the record industry in the mid 1990s, producers of MP3 player devices and programs would be able to include screening information to make sure that protected content was not played without authorization.¹⁷

Copy protection and rights management recording has been a subject of intense concern by record companies, starting with RIAA's call for help in 1980, CBS's and later Sony's work in the area. It is fair to say that RIAA's effort in this work is marked by a number of poor technical decisions and engineering bumbles. In 1989, RIAA's Engineering Committee clearly believed that a marking scheme for protecting recordings was in its near grasp. I must say that I agree with this 1989 assessment. While the state of the art has advanced and continues to advance, a technology adopted a decade ago would have resulted in a decade's worth of recordings with copyright information clearly marked in the audio of the recordings themselves. Even if the RIAA adopted a system in the late 1980s or early 1990s that was not perfect, it could have been used in a transitional fashion, just as Musicode is being used as a transitional watermark by SDMI.

I find it interesting that SDMI, within 3 months of being formed, was able to collect a large number of excellent technical proposals for realizing marking of copyright information.¹⁸ Had RIAA applied similar engineering management to its task back in the

¹⁷ Similarly, several observers have proposed distributing content in encrypted form. This content would be read and decrypted by a tamper-resistant device which would enable the information to be played or displayed. This technology was well understood in the 1980s.

¹⁸ Similarly, other major companies have had no difficulty in selecting secure formats for distributing music recordings. For example, see 27 July 2000 article in the *New York Times*, "AOL to Distribute Software to Secure Music Copyrights: Intertrust Deal Could Break Online Logjam" by Amy Harmon, containing the following excerpt:

"In a deal that may break a logjam in promoting the secure downloading of music from the Internet, America Online said yesterday that it had agreed to include

late 1980s, the picture today would be very different. Most recordings would have audio indicators of copyright information which could be used by MP3 players and by Napster software.¹⁹ If such screening would be considered appropriate based on legal and policy considerations, it would be easy to technically effect.

software from the InterTrust Technologies Corporation on compact discs containing AOL 6.0, the newest version of its own popular software, later this year.

“Like several other companies providing digital rights management, InterTrust allows a recording company to wrap music in software that ensures it will only be played according to rules specified by the company. Before downloading a song, customers would agree to terms like how much they would pay, or how long they would be able to play it.”

¹⁹ Meanwhile, there have been commercial proposals for copy protection as well. For example, Midbar, an Israeli company, markets a product called Cactus Data Shield that purports to limit improper copying of digital audio. Midbar’s home page (<http://www.midbartech.com/>) contains the following statement: “Midbar’s Cactus Data Shield protects music CDs against unauthorized digital duplication. Transparent to the music provider, this breakthrough technology is easily implemented in the CD manufacturing process.” To the best of my knowledge, Midbar uses a proprietary technology, and I am not familiar with the details of the system, but if Cactus Data Shield does what Midbar claims it does, it could prevent sharing of information from CDs. I am not aware of any systematic use of this technology by any American music distributor or CD publisher.

Conclusion 7: Napster can not tell whether a use of the system is infringing

- *Napster can not tell whether a particular use of its system is infringing.*

There are many legitimate reasons why a user may wish to use Napster with copyrighted materials. These reasons certainly appear to be acceptable, and may be permissible without the consent of the copyright holder.

- Consider a user owning a recording on a vinyl LP record. Every time she plays the LP, it will cause wear on the vinyl recording and degrade the quality of the recording. Such a user would certainly seem to have a legitimate right to use Napster to access digital copies of the recording for later playback.
- Consider a CD collector who wishes to play his CD collection both at home and at work. As an alternative to transferring the physical CDs on a daily basis, he may wish to use Napster to make the recordings available in both locations.
- Consider a creator of a musical recording who wishes to freely distribute the recording to attract publicity for her musical group. Napster offers a simple way of doing so.
- Consider a musical conservatory instructor who is teaching a course on instrumental play. He may wish to play some of his own copyrighted recordings to his class to illustrate a point; Napster is a particularly easy and inexpensive way for him to make the recordings available both during lecture and for private study by the students.
- Consider a music critic who wishes to review several recordings. Napster provides a convenient way for her to quickly access those recordings for review.

- Consider a student who has a limited budget for his CD purchases. Napster provides an easy way to preview the material to help him choose the items he wishes to purchase.
- Consider the jogger mentioned in the introduction who wishes to carry recordings with her for her morning jog. Napster provides an easy way for her to carry the recordings on her MP3 playback device.

These examples are not meant to be exhaustive, but merely to illustrate that Napster has a variety of uses, many of which appear to be perfectly legitimate, even if they involve copyrighted material. There is no way for Napster to distinguish the purpose for which it is used.

Conclusion 8: ID/password mechanisms are a customary way of allowing access

- *The use of ID/password mechanisms to allow or restrict access to a service such as Napster is reasonable and customary and is superior to use of IP source addresses.*

Napster uses a user ID and password to determine whether a user should be allowed to access the Napster utility. To access Napster, the user must present a valid user ID and password. In the Beta 6 version, Napster further stores values in the user's operating system. On Windows systems, there is a structure called the "registry". Values can be stored in the registry under specific labels. I understand that Napster stores the user ID in the operating system's registry. This is similar to cookie technology widely used by Internet sites.

If a user is denied access to the Napster site (perhaps because he has been identified as dealing improperly with copyrighted material), he will not be allowed to log in. In addition to denying the user access at Napster's server, the Napster client will also store a special additional data in the registry indicating that a user on this computer has been blocked. If the user attempts to reestablish a second account with a new user ID and password, the previous data stored by Napster in the registry will be discovered, and he will still be denied access. To gain access, the user will need to erase his registry (which will cause a number of problems in the operation of his computer) or to manually edit the registry. While it is possible that a user could edit his registry, it would take a fair level of technical expertise to do so.

The ID/password scheme used by Napster is similar to schemes used throughout the Internet. It provides an appropriate level of protection and is superior to alternatives. To see this, consider some alternatives:

- *Use of IP addresses.* Every user on the Internet has an IP address for her computer. This value identifies the computer for the duration of the session on the Internet and is widely used by communications software. Thus, some people might think that IP addresses are an appropriate way to identify users. Unfortunately, this is not the case. Many IP addresses are assigned on a rotating basis using “Dynamic IP addresses.” For example, an Internet Service Provider (ISP) may wish to support more users than it has IP addresses. In this case, it will conserve IP addresses by reusing them. Every time a user logs in, she will receive an IP address assigned out of the pool. If she logs in two different times, she is likely to have two different IP addresses, and thus they are not an effective long-term identifier of a particular computer. Worse, two different users may at different times be assigned the same IP address. Suppose identification were based on IP address and user A were to have her Napster account suspended. If user B later logged in and coincidentally was assigned user A’s former IP address, he would also have access to Napster denied, even though he may have done nothing wrong. If one’s IP address is blocked, it is easy to get new IP addresses by signing up with a new ISP, including the services offering IP addresses for no or little charge. One does not even need to switch ISPs to get a new IP address, one may be assigned or request a new IP address. (I have personally had this experience with PacBell’s ISP.)

Network Address Translation (NAT) adds further difficulties for the IP address blocking approach. Firewalls and a variety of consumer products (including stand-alone black boxes starting at about \$150 and the Linux operating system, which is available for free) provide a facility for multiple users to simultaneously access the

Internet through a single IP address. Messages coming from the Internet to the NAT device have a single IP address. The NAT device translates that single IP address into local IP addresses that are recognized only on the local area network. The bottom line is that a large number of distinct computers may be sharing a single IP address. Blocking one IP address would block every user sharing the same NAT device. With increased concern over denial-of-service attacks and computer security in general, use of firewalls and NAT devices continues to grow dramatically.

- *Names.* Another approach would be to use names, or mailing addresses, as ways of indicating unique identity. Unfortunately, this is likely to be ineffective. First, people write their names in many different ways. My legal name on my birth certificate and passport is Justin Douglas Tygar. However, I commonly go by the name Doug Tygar and professionally sign my papers as J. D. Tygar. I often visit Asia, particularly Japan and Taiwan, and I have Japanese and Chinese names that are not easily correlated with my English name (e.g., my legal Chinese name is transliterated into Roman characters as HU Dao-ge.) More generally, names written in non-Roman alphabets such as Cyrillic or Chinese often have multiple transliterations into Roman characters. But, more importantly, there is nothing to prevent me from using a pseudonym or false name on the Net, and indeed, many users of the Internet appear to regularly do so. A recent book *The Hundredth Window* explicitly recommends using false names for privacy reasons.²⁰ Similarly, addresses are not effective unique identifiers. I have appointments in two departments on the

²⁰ A 16 August 1999 CNET article “Are ‘Registered User’ Figures Worth Anything?” by Jim Hu casts further doubt on how accurately users report their names when registering for sites, including the following quote by Patrick Keane of Jupiter Communications: “How many of those Yahoo users are using ‘Elvis’ as their user name? I’d say a pretty huge number.”

UC Berkeley campus, and correspondingly two offices. Each of those addresses, as well as my home address, can be written in multiple ways that are accepted by the US Postal Service. And again, this all supposes I use my correct address. E-mail addresses are equally poor as unique identifiers. In short nothing about personal names or associating alleged names with user names would assist Napster in blocking users who are alleged to infringe.

- *Biometrics.* User's identity may be distinguished by physical characteristics, such as hand geometry, fingerprints, or the pattern of blood vessels on the retina of the user's eye. Using a variety of techniques, these can be measured using specialized equipment (such as retinal scans). In practice, biometrics is vulnerable to problems with both false positives (inaccurately reporting more than one persons as having the same identity) or false negatives (inaccurately reporting one person as having more than one identity). Moreover, it is often easy to get these systems to report two different identities (for example, ordinary people have two eyes, and thus two possible retinal scans.) Perhaps in the future, more reliable techniques may become available. But this technology is not in widespread use today, and biometric equipment is hardly standard on consumer PCs.
- *Smart cards.* Smart cards are physical tokens that can be used to identify a user. Well designed smart cards take advantage of public key cryptographic techniques to provide positive identification of users. However, in the United States, there is no official organization that checks the identity of individuals and issues smart cards to them to prove their identities. Of course, because of the international nature of the Internet, a program to issue identifying smart cards would ultimately have to be

international in scope. Such a program would need to agree on a common standard for identification, develop ways to systematically and uniquely identify all residents, issue exactly one card securely to each resident, and providing a secure way to revoke cards in case they are lost. Moreover, smart card reading equipment would need to be standard on consumer-oriented computers. Finally there are some fundamental and unsolved security problems with smart cards, including their vulnerability to “differential power analysis” attacks: <http://www.cryptography.com/dpa/index.html> .

- *Public key cryptography.* Companies such as Verisign (<http://www.verisign.com>) offer “public key cryptographic certificates” that are digitally signed by “certificate authority” offering a way of identifying users through a variety of well-known authentication protocols. This is an elegant, effective solution to authentication. Unfortunately, personal Verisign certificates are not commonly held by ordinary World Wide Web users, and it begs the question of how Verisign would uniquely authenticate individual users’ identities to issue them a certificate. Moreover, these techniques depend on each user keeping his “private key” (corresponding to the “public key” in the certificate) secret. If a user accidentally discloses his private key, the authentication would be lost.
- *Credit cards.* Some have suggested using credit card numbers as a method of authentication. However, this suffers from many drawbacks. First, many people (such as young people, the poor, people with bad credit histories, resident aliens, people overseas, etc.) may not have a credit card. Second, many people have more than one credit card, so a credit card number would not uniquely identify a user. A banned user could simply use an alternate card. Third, there would be legitimate

privacy concerns with a large database of credit card numbers being kept by a company. Fourth, it is not clear to me that Visa International, Mastercard International, or other companies would consent to having their credit card numbers used in this fashion. Credit card companies may be concerned with the potential liability associated with misuse of the technology. Fifth, it is not hard to steal another person's credit card number. I often observe people being very casual in their use of credit card numbers. For example, they may not properly destroy receipts, leading to a practice known as "dumpster diving" where people will try to steal credit card receipts from trash. Sixth, credit card numbers are usually disclosed in several contexts when an individual makes a purchase. When I buy a meal at a restaurant and charge it to my credit card, the server usually takes my card and the bill to a back room. I do not know if the server may be copying my card – but if she were a Napster fan and wanted to establish a fake identity using the credit card, she would have a clear motivation for doing so.

The bottom line is that while user ID/password authentication is not perfect, it is preferable to the alternatives.

Conclusion 9: Rejecting bots is a way to help maintain performance

- *A bot is a computer program that automatically performs Internet accesses. For example, a bot might monitor all the files available for download through Napster. The use of bots can result in significant load and performance degradation of an Internet service such as Napster, and thus are sometimes blocked for performance reasons.*

“Bot” is a term used to refer to a program that searches the Web and tries to determine information. For example, shopping bots are sometimes used to compare prices to find the cheapest seller of a given item. The web site <http://www.bookfinder.com> is an example of such a system – it finds available new and used copies of books and allows comparison of prices and quality of the product by searching many different web sites.

However, bots can induce tremendous load on a system. Because they are constantly querying web servers, they may induce significant load. In a well-publicized case, the auction site Ebay recently implemented software to detect and ban bots from their site, citing performance reasons (see, for example, <http://technews.netscape.com/news/0-1007-200-1948171.html>).

Napster similarly bans bots from accessing the search engine functions of Napster. They assert that this is to maintain the performance of their search engine, and this seems to me to be a perfectly sound reason. This has the side effect of banning bots that might search for appearances of a certain word (such as “Metallica”) in file names.

It may seem contradictory at first that Napster could effectively detect and screen out bots and still not be able to always screen out users on a negative “do not allow

access” list. However, more consideration will quickly lead one to the conclusion that this is completely natural. Bots are caught by Napster in the act of being bots – they are caught while their IP address is known to Napster and their address can be immediately dropped to terminate their activity. By contrast, if a user is to be blocked by the system for alleged infringement, he would have to be banned when he logged back on, since the identification generally occurs at the time of log on. As discussed above, IP addresses for users may vary, and can not be used as a reliable way of authenticating the users. Users authenticate themselves to Napster, and if they are able to falsely authenticate themselves, they can not be identified on a rejection list.

The following analogy may help illustrate this point. Consider a small merchant who is concerned about shoplifting in his shop. If he is observant, he will be able to see shoplifters and catch them in the act of shoplifting. However, he will probably not be able to effectively keep people from his store if he is armed with a list of hundreds of thousands of names of known shoplifters – how would he know if someone on the list entered the store?

The topic of detecting and deterring bots has attracted attention recently in the computer science community. Bots are now being used in a variety of distributed computation contexts, including computer gaming (see, for example, <http://www.botepidemic.com>). Their presence is often unwanted on systems, so a number of approaches for rejecting bots have been developed. This in turn, has inspired a number of computer programmers to attempt new techniques for bots, resulting in yet more sophisticated bot rejection techniques.

Signature

I declare, under penalty of perjury, under the laws of the United States of America, that the foregoing is true and correct. Signed this ____ day of June 2000, in Peitou, Taiwan.

Justin Douglas Tygar, Ph.D.