Testimony before the House Committee on Oversight and Government Reform

Thomas D. Sydnor II,
Office of International Relations,
United States Patent and Trademark Office

July 24, 2007

Chairman Waxman and Ranking Member Davis, thank you for holding a hearing on the important problem of inadvertent filesharing. Together with Professor Lee Hollar and Mr. John Knight of the Department of Computer Science at the University of Utah, I am a co-author of the USPTO Report, *Filesharing Programs and "Technological Features to Induce Users to Share."*

Unbeknownst to many, users of popular filesharing programs are "sharing" files they do not intend to provide to thousands of strangers. These files may contain copyrighted works that users cannot legally distribute; they may also contain sensitive or proprietary data belonging to the user or a family member's employer. This problem can be called "inadvertent sharing."

Right now - and completely unknown to them – Americans are sharing sensitive personal data—their bank records, credit-card numbers, passwords, tax returns, and letters, to name a few. Without their knowledge, businesses are sharing confidential data about their customers, employees, and strategic plans. Federal, state, and local governments are also affected—and sensitive data has been exposed. Worse yet, Internet criminals know this, and they are data-mining filesharing networks.

Any program or service that lets users make files or data available to other users of the Internet *could* cause inadvertent sharing—regardless of whether it was a "centralized" server-based social-networking website or a fully "decentralized" peer-to-peer filesharing network.[1] In itself, the use of peer-to-peer networking should not affect whether users of a given program or service share or upload files unintentionally.

This Committee has shown great prescience in investigating filesharing. Back in 2003, this Committee investigated inadvertent sharing, even though the consequences seemed somewhat hypothetical: Then, it was unclear that inadvertent sharing could result in identity theft. Now, leading security experts, like Howard Schmidt, co-author of the Administration's *National Cyber-Security Policy*, conclude that inadvertent sharing is "a major part of the current identity theft problem." For example, Denver District Attorney Mitchell Morrissey recently indicted a gang of identity thieves who were buying crystal meth by downloading inadvertently shared financial data with LimeWire.

[1] For example, corporations and other entities often maintain complex networks of computers, network drives, and webservers in order to provide differentiated access to files and data: Some files and data are accessible to any user of the Internet, some only to those authorized to access a corporate "intranet," and others can be accessed only by particular employees or groups of employees. Even when such systems do not use peer-to-peer networking, files or data can be shared more broadly than was intended if permissions are managed incorrectly or if files or data are stored in the wrong location.

Surprisingly, inadvertent sharing by consumers has rarely been reported outside of the context of filesharing. The designs of popular social-networking, photo-sharing or blog-hosting sites explain why. Creators of these programs and services avoided designs that would tend to cause inadvertent sharing: Just like the developers of some early filesharing programs, they ensured that users would have to take multiple, affirmative steps before they would share or upload any given file. However, in recent years, distributors of file sharing programs have deployed features that may promote inadvertent file-sharing.

Four years ago, this Committee, and then the Senate Committee on the Judiciary, held hearings on *Usability and Privacy,* and inadvertent sharing. During both hearings, several legislators expressed concerns that unless distributors of file-sharing programs eliminated these features and their effects, their programs could compromise national security In response to these concerns, many distributors developed "voluntary standards and practices" to prevent inadvertent sharing. The resulting standards were complied in an industry *Code of Conduct*. This *Code* imposed three obligations to prevent inadvertent sharing:

- **The "Conspicuous Confirmation Requirement:** "[Our] software … shall conspicuously require the user to confirm the folder(s) containing the file material that the user wishes to make available to other users…."

- **The "Reasonable Design" Requirement**: "[Our] software … shall be designed to reasonably prevent the inadvertent designation of the contents of the user's … principle data repository … as materials available to other users."

- **The "Ready Uninstall" Requirement**: "A method by which [our] software … readily may be uninstalled shall be provided to users."

However, even with the *Code of Conduct,* inadvertent file sharing kept reoccurring—and causing the very problems that this Committee had documented or foreseen in 2003. For example, the Department of Homeland Security soon reported that inadvertent sharing was disclosing classified data: "Multiple organizations have ongoing investigations into disclosure of sensitive or classified material due to P2P."

When reports like this came to the attention of the USPTO, Jon Dudas, the Undersecretary of Commerce for Intellectual Property, directed me to find out why this supposedly solved problem was recurring. I then enlisted the computer-science expertise of my coauthors. We created a set of reporting criteria, and examined how the sharing-related features of five popular filesharing programs had evolved.

Our findings were presented in the USPTO Report, *Filesharing Programs and "Technological Features to Induce Users to Share."* It analyzed five popular filesharing programs, as well as *two* types of inadvertent sharing that could harm users.

Some users might inadvertently share *downloaded* files acquired through the filesharing program. Sharing of downloaded files can expose the user to a copyright-enforcement lawsuit because such files may be infringing: One study found that almost 97% of the files requested for downloading were infringing or highly likely to be.

Users might also inadvertently share *existing* files created by other programs and stored on the user's computer. Sharing existing files can expose families to identity theft, job loss, and an infringement lawsuit: Most computers contain sensitive personal data, employers' data, and large collections of audio files ripped from legally purchased CDs.

The USPTO Report concluded that the distributors of the five programs studied had repeatedly deployed five "features" that had a known or obvious tendency to cause inadvertent sharing of downloaded or existing files, or both:

- **Poorly Disclosed Redistribution Features**: By default, most filesharing programs will cause users to share files that they download. If poorly disclosed, these features can cause inadvertent sharing of downloaded files.

- **Share-Folder Features**: These features let a user select a different folder to store downloaded files—but they do not warn the user either that the folder selected will be shared or that its subfolders will be shared recursively. These features can cause users to share existing *and* downloaded files inadvertently: A user who tries to store downloaded files in an accessible location like "C:\" or "My Documents" will tend to "share" all of their personal files *and* their collection of audio files ripped from purchased CDs.

- **Search-Wizard Features**: These features search a user's hard drive, or drives, and either recommend or cause the sharing of folders that contain enough "media" files, including document, image, audio, and audiovisual files. They often recommend that new users share "My Documents" and all of its subfolders.

- **Partial-Uninstall Features**: These ensure that when a user uninstalls a filesharing program, the process will leave behind a data file. If another copy of that program is ever installed again on the user's computer, it will read that data file and share all folders shared by the "uninstalled" copy of the program. The user may receive no notice of this changed default behavior. These features can cause inadvertent sharing of downloaded or existing files.

- **Coerced-Sharing Features**: These provide misleading feedback that makes it look like a user has disabled sharing even though files are still being shared. These features can cause inadvertent sharing of downloaded files and inadvertent sharing of existing files if deployed with a share-folder feature.

Appendix A to this statement illustrates each of these features. While all can cause inadvertent sharing, the search-wizard and share-folder features criticized by *Usability and Privacy* are particularly troubling. In most programs, they cause *recursive sharing*: Not only will the user "share" most or all files stored in a folder selected by a wizard or used to store downloaded files, the user will also "share" most or all files stored in *all subfolders* of that folder. These share-folder and search-wizard features became *more* widely used and their implementations *more* aggressive *after* distributors had created a *Code of Conduct* that should have prohibited use of KaZaA-like share-folder or search-wizard features.

The continuing use of these five “features” is also troubling because they appeared and proliferated in waves: As users of filesharing programs learned how to disable some of these features, new ones appeared.

During 2002, share-folder, search-wizard, and partial-uninstall features appeared. By mid-2003, they were widely deployed in many filesharing programs. But then, the district-court decision in *Grokster* forced copyright holders to sue users sharing hundreds or thousands of infringing files. Predictably, users tried to stop sharing infringing files.

Then, coerced-sharing features began to proliferate. By July of 2005, four out of the five programs studied contained coerced-sharing features.

Certain “business models” worked only if many users of file-sharing programs shared many infringing files. When users were sued for doing that, their propensity to share infringing files plunged—and “technological features” that could “induce users to share” files inadvertently proliferated. As a result, the worst effects of inadvertent sharing—widespread identity theft and dangerous breaches of personal, corporate and national security—may have increased.

I will conclude by stressing two factors that make the prevention of inadvertent sharing particularly important. Each was stressed during this Committee’s 2003 hearing. Each remains valid today.

First, filesharing programs are designed to go where they are not wanted and to thwart the security measures that could exclude them. As Dr. Hale told the Committee in 2003, “P2P software is commonly designed to circumvent network security services.… Techniques such as tunneling, port hopping and push requests make it difficult to detect and filter P2P traffic. That is their intent; to foment user participation in spite of an enterprise’s security policy.… [T]here is no reason for [port-hopping] other than to allow network software clients to avoid detection.” LimeWire now agrees “that it is inappropriate for file-sharing programs … to be installed on any computer with highly sensitive information.” But it has made it difficult and expensive for computer owners to prevent this result. This makes it particularly important to ensure that users of its program never share any files inadvertently.

Second, as Chairman Waxman noted in 2003, “The users of file-sharing programs are predominantly teenagers.” Today, filesharing programs are still widely used by teenage or preteen children—and used to break the law: In the *Grokster* case, evidence showed that “[a]lmost 97% of the files actually requested for downloading were infringing or highly likely to be infringing.” Popular filesharing programs do have lawful uses, but many of their actual users use them to break the law much of the time.

This has safety implications: When teenagers or pre-teens use filesharing programs, they enter a shadowy network of anonymous strangers and mislabeled files that look like popular songs, but contain child pornography or dangerous spyware. The USPTO Report makes one point clear: When people enter these networks, no one will be looking out for them.

The conduct described in the USPTO Report is disturbing because it continued—in public—for nearly five years. Law-abiding adults did not detect it because they had no reason to use filesharing programs. So it was not detected by consumer advocates or the vast information markets that surround most popular consumer products. Even tech-savvy public-interest groups that focused on filesharing were blinded: They seem to have had no knowledge of how the public was being affected out on the electronic frontier.

Nor could users of filesharing programs complain to enforcement agencies when inadvertent sharing affected them. As the FBI told this Committee in 2003, when people are harmed while breaking the law, they have strong incentives to avoid involving law-enforcement agencies. If virtually every one using these programs is using them to break the law, then no one can complain if they are harmed.

For all of these reasons, it is important to understand why inadvertent sharing occurs and why the features known to cause it kept on being deployed. If the continued use of these features resulted from error, then the risk of inadvertent sharing might be expected to decrease: Over time, mistakes should tend to be fixed. But if these features were intended to dupe users, then the risk of inadvertent sharing might be expected to increase. People do not like to be tricked: Over time, duping schemes should thus tend to evolve, proliferate, and become more deceptive. The disturbing persistence of inadvertent sharing—the same "features" in the same programs repeatedly causing the same problems—thus raises important questions with broad implications.