CyberEthics

Morality and Law in Cyberspace

Third Edition

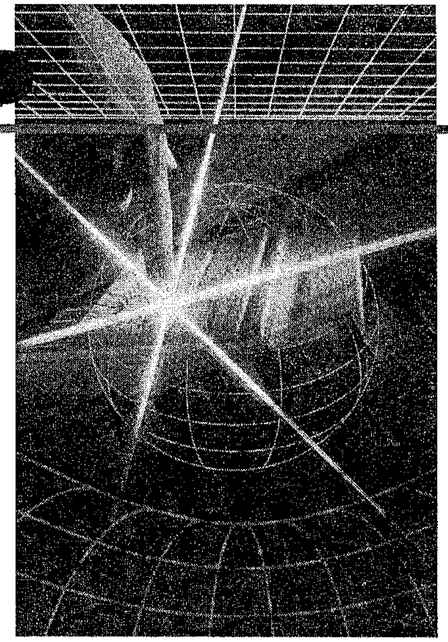
(2006)

Richard A. Spinello
Carroll School of Management
Boston College

JONES AND BARTLETT PUBLISHERS
Sudbury, Massachusetts
BOSTON TORONTO LONDON SINGAPORE

CHAPTER TWO

Regulating and Governing the Internet



Although there has been much written about the perils of overexposing children and teenagers to the Internet, a headline in *The New York Times* sounded especially ominous: "A Seductive Drug Culture Flourishes on the Internet." The article explained how the Internet is now rife with web sites that endorse illegal drugs or provide explicit instructions for making, growing, and consuming such drugs. Many of these web sites make drugs sound exciting and alluring, and never even hint about the risks of addiction. Also, the problem is compounded because "the Internet lacks a quality control mechanism to separate fact from hyperbole or from outright falsehood, even in discussions that may ultimately encourage an activity that remains illegal for Americans of all ages."[1]

This is a disturbing but certainly not a surprising development, and for some it does not augur well for the future of this ubiquitous technology. But from its earliest origins a free-wheeling spirit has dominated the rules of discourse in cyberspace. According to Jonathan Katz, "it is the freest community in America."[2] Hence one of the most formidable issues faced by public policy makers throughout the world is whether or not to impose some limits on this free and unencumbered flow of information in cyberspace—to restrict, for example, the dissemination of pornography or perhaps to ban these nefarious web sites that promote illicit drug use. Even if a decision were made to do so, implementation of that decision would be quite challenging given the Internet's complexity and its vast global reach.

The debate about pornography on the Internet or about web sites advocating illicit activities reflects deeper questions about how the Internet should be regulated or governed. Although the Internet's anarchy and lack

of structure has led to some excesses, many users are loath to see it replaced by tighter, centralized controls. Most civil libertarians, for example, contend that the Internet thrives precisely because there is no central governing authority. Consequently, they favor the continuation of decentralization and self-governance instead of any form of government intervention, believing that traditional forms of regulation would interfere with electronic interactions and the free flow of ideas. They argue that the Internet should be able to develop its own unique political structure, set appropriate standards, and even handle its own disputes.

Civil libertarians, however, are slowly losing this struggle. During the last decade there have been many regulations imposed on the Internet such as the Can-Spam law in the United States. Europeans have been particularly active with their extensive data protection policies and an international treaty on cybercrime. As these regulations proliferate, the haze of legal ambiguity in cyberspace will steadily dissipate.

Nonetheless, the residual confusion and anarchy can expose users to legal landmines and potentially costly liabilities. Take the case of the Blue Note jazz club in Greenwich Village versus the Blue Note nightclub in Missouri. While the Blue Note club in New York has a federal trademark for its name, the Blue Note in Missouri obtained "the right to use the trade name locally in Missouri"—what is known in trademark law as a "geographical carve-out." But when this Blue Note designed its own web page the Blue Note in New York protested, claiming that its trademark had been violated by the Missouri club's worldwide presence on the Internet. As a result of this and similar cases the legal system is still trying to decide whether business on the Internet "falls under the laws of some, all, or any of the jurisdictions from which that Internet site can be reached."[3]

Can governments begin to resolve the numerous jurisdictional problems created by the Internet? What will happen when different sovereignties begin to assert their authority in cyberspace? Will increased regulation stifle the Internet's innovative spirit? Is there an alternative to a cyberspace that is burdened by more rules and restrictions?

Before plunging into a discussion of these complex matters it is instructive to review the history and technology of the Internet, and so we devote a portion of this chapter to that purpose. It is important to understand the architectures of the Internet to appreciate the various possibilities for regulation and government intervention. This overview includes a cursory treatment of the World Wide Web and the recent surge in electronic commerce along with some of its attendant social problems. It is also instructive at this point to consider the separate but related issue of governance, that is, managing of mundane tasks such as assignment of domain names. This process too has triggered ethical controversies that are considered in later chapters.

Our primary purpose in this chapter, however, is to discuss the appropriate regulatory response to the social problems that have emerged in the

online world. Can market forces handle these problems or is government intervention essential? What about the possibility of a more decentralized bottom-up approach? Perhaps the optimal approach is finding the right interaction of policy and technology?

This discussion sets the stage for the more in-depth treatment of speech, property, privacy, and security in the remaining chapters. For each of these broad issues it is necessary to evaluate how underlying technologies change our ability to establish and enforce policy.

A Short History of the Internet

This summary of the Internet's creation is not a mere indulgence in nostalgia. We investigate the past to understand the present—by looking at the Internet's technological evolution we can better appreciate its present architecture and perhaps uncover some clues about its future.

The origin of the Internet's basic architecture can be traced back to the search for a "survivable communications" system. During the late 1950s the U.S. Department of Defense (DOD) was concerned about the need for a failure-resistant communications method. In 1961 Paul Baran developed such a method, which has become known as *packet switching*. Baran admits that "the origin of packet switching itself is very much Cold War."[4] Package switching (originally called "message switching") works by breaking up a message into fixed-sized units or "packages"; each package is "labeled with its origin and destination and is then passed from node to node through the network."[5] This technology was also being separately developed by Donald Davies, a British expert on computer security, who was the first to use the term "packet" in reference to data communications. Davies also built an experimental packet switching network in the mid-1960s.

The first large scale packet switching network that was developed based on the insights of Baran and Davies was the work of the Advanced Research Projects Agency (ARPA), a research agency of the DOD, which financed high-tech research. In the late 1960s, the DOD provided generous grants to universities and corporations to establish a communications network between major research centers in the U.S., including universities such as MIT and Stanford. It recruited Lawrence Roberts of MIT's Lincoln Laboratory to oversee the construction of the ARPANET, the first incarnation of what is now known as the Internet.

The basic infrastructure of the ARPANET consisted of several time-sharing host computers, packet switching interface message processors (IMPs), and leased telephone lines. The host computers were already in place at the universities and research centers that would be part of the network; AT&T provided the telephone lines. The IMPs were needed to perform key network functions such as sending and receiving data, error

checking, and message routing. The responsibility for building these systems was delegated to Bolt, Beranek and Newman (BBN) a research and consulting firm in Cambridge, Massachusetts.

By the end of 1971, the primitive ARPANET was up and running. Its primary goal was supposed to be resource sharing, that is, enabling connected sites to share hardware processing power, software, and data. But the network's users soon discovered another function: electronic mail. Instead of using the network primarily to leverage remote hardware resources, users began sending huge volumes of e-mail. As a result, this popular application soon began to dominate traffic on this fledgling network. According to Abbate, "Network users challenged the initial assumptions, voting with their packets by sending a huge volume of electronic mail but making relatively little use of remote hardware and software. Through grassroots innovations and thousands of individual choices, the old idea of resource sharing that had propelled the ARPANET project forward was gradually replaced by the idea of the network as a means of bringing people together."[6]

In the early 1980s, this system was subdivided into two networks, the ARPANET and Milnet. Furthermore, connections were developed so that users could communicate between the two networks. The interaction between these networks came to be known as the *Internet*. *Internet* was actually first used in a research paper written by Cerf and Kahn in 1974; that paper described a "network of networks" that would eventually link together computers all over the world. In the late 1980s, the National Science Foundation network (NSFNET), which relied on five supercomputers to link university and government researchers from across the world, replaced the ARPANET. The NSFNET began to encompass many other lower level networks such as those developed by academic institutions, and gradually the Internet as we know it today, a maze of interconnected networks, was born.

In these early days the federal government generously subsidized the Internet, and as a consequence there were restrictions on any commercial use. The Internet was the exclusive domain of government researchers, scientists, university professors, and others who used it primarily to share their research findings or other academic information.

But the NSF no longer subsidizes the Internet, which has assumed a strong commercial character during the last decade. During the early 1990s the Internet quickly became available to corporate users; e-mail providers such as MCI and Compuserve opened up e-mail gateways. By 1993, 29% of the host computers connected to the Internet belonged to corporations. Commercial use now accounts for the vast majority of all Internet traffic. Management of the network has been transferred to private telecommunications carriers that manage the backbone, that is, the large physical networks that interconnect. Thus, the network's vitality depends on the cooperation and goodwill of these telecom providers.

The global diffusion of Internet usage during this period has been an extraordinary phenomenon. In 1983 there were a mere 500 host computers (computers with unique Internet Protocol addresses) connected to the Internet. That number has grown to over 200 million host computers at the beginning of this new millennium. By 2005, the number of world Internet users was estimated at 1 billion, approximately 15% of the population; the growth rate between 2000 and 2005 has been 160%.[7] Although the rapid development of the global Internet has been extraordinary, there is still a disparity between developed and developing countries. However, in some countries that is also beginning to change. In Latin America, there were fewer than 30 million Internet users in 2000, but that number has grown to over 68 million in 2005.[8]

This global connectivity provided by the Internet is perhaps its most attractive feature. It brings together millions of people and thousands of organizations all over the world, and has helped to achieve what the *Economist* calls "the death of distance," that is, the overcoming of geographic proximity as a barrier for conducting business.

The Internet's Current Architecture

How does this all work? As intimated, there is actually little physical substance to the Internet. There are a few dedicated computers at key connection junctures, but "like a parasite, the Internet uses the multi-billion dollar telephone network as its hosts and lets them carry most of the cost."[9] Data is fluidly transferred over this network by means of a network protocol called TCP/IP. The TCP/IP protocol allows for complete interoperability on the Internet so that computers can communicate with one another even if they have different operating systems or applications software. TCP/IP therefore makes the network virtually transparent to end users no matter what system they are using, and it allows the Internet to function as a single, unified network.

TCP/IP consist of two elements: the IP or Internet Protocol, which establishes a unique numeric address (four numbers in the form **nnn.nnn.nnn.nnn** ranging from 0 to 255) for each system connected to the Internet. IP is a means of labeling data so that it can be sent to the proper destination in the most efficient way possible. If a user connects to the Internet through an Internet Service Provider (ISP), that user is normally assigned a temporary IP address, but users which connect from a local area network (LAN) in their organizations are more likely to have a permanent IP address.

The second piece, TCP, or Transmission Control Protocol, enables network communication over the Internet. As discussed, the data are broken up into pieces called "packets," with the first part of each packet containing

the address where it should go. The packets are then sent to their destination by a system of routers, that is, servers on the Internet that keeps track of Internet addresses. These packets can take completely different routes to reach their goal. Once all the packets arrive, the message or data will be reconstructed, based on the sequence numbers in the headers to each packet, and redirected to the appropriate application.

The Internet's physical infrastructure is composed of many large, interconnected networks which are known as Network Service Providers (NSPs). NSPs include IBM, SprintNet, and PSINet as well as several others. According to Hafner, these backbone providers "adhere to what are known as peering arrangements, which are essentially agreements to exchange traffic at no charge."[10] Each NSP connects to three network access points, and at those points packet traffic may be transferred from one NSP backbone to another. NSPs also sell bandwidth to smaller network providers and to ISPs.

Routers, also known as "packet switches," perform much of the work in getting data transmitted over the Net to its ultimate destination. When a packet arrives at a router, the router looks at the IP address and checks the routing table, and if the table contains the network included in the IP address the message is sent to that network. If not, the message is sent along on a default route (usually to the next router in the backbone hierarchy). If the address is in another NSP, the router connected to the NSP backbone sends the message to the correct backbone where it is sent along by other routers until it reaches the correct address.[11]

As we survey the Internet's technical and social evolution, the most distinctive features of its network architecture should be apparent. Perhaps the Internet's most important characteristic is its *openness*, thanks to an open-ended network architecture the Internet has supported an extraordinary level of innovation: e-mail, blogs, instant messaging, and MP3 music files are just some of the many applications this technology has enabled. According to Castells, "the openness of the Internet's architecture was the source of its main strength: its self-evolving development, as users became producers of the technology and shapers of the whole network."[12]

Second, the Internet is *asynchronous.* Unlike telephone communication, there is no need for coordination between the sender and recipient of a message. An e-mail message, for example, can be sent to a mailbox that can be accessed at any time by its owner. Third, the Internet permits a *many-to-many format of communications*[13]: many users can interact with many other users through electronic mail, bulletin boards, web sites, and other vehicles. Unlike traditional media such as newspapers the Net is interactive; users can speak back. Fourth, the Internet is a *distributed* network instead of a centralized one, whereby data can take any number of routes to their final destination. There is no center to the Internet, that is, there is no central server or single controlling authority, because information can travel from

one location to another without being transmitted through a central hub. This gives users more control over the flow of information. Because it is a decentralized, packet-based network, it is more difficult to censor that information. Also, this resilient design makes the Internet's structure more durable. As Hafner points out, "that deceptively simple [packet-switching] principle has, time and again, saved the network from failure."[14] When a train fire in Baltimore damaged a critical fiber optic loop, Internet data easily circumvented the problem. Finally, the Internet is highly *scalable,* that is, it is not directly affected when new computer links are added or deleted. Thus, it allows for much more flexible expansion or contraction than many other proprietary network technologies. Its basic architecture encourages universal access and participation.

The Internet, then, should really be conceptualized as a flexible and open infrastructure. It is designed to maximize interoperability, that is, to be completely independent of software programs, hardware platforms, and other protocols. As a result, it is well suited to new applications and can easily accommodate revolutionary developments in both software and hardware. Because of its malleability, however, it is naïve to assume that the Internet of today will be the Internet of the future. The architectures of cyberspace could conceivably undergo a major transformation in the next few years. As we pointed out in Chapter 1, if the state chose to influence those architectures by mandating digital identity or otherwise controlling access through ISPs, cyberspace could become a very different place.

The World Wide Web

The most recent surge in the Internet's popularity can be attributed to the emergence of the World Wide Web. The web is a collection of multimedia documents that can be easily accessed through the Internet. The web was developed at the European Particle Physics Lab as a means of exchanging data about high-energy physics among physicists scattered throughout the world. This group developed a standard known as Hypertext Markup Language (HTML) that supports a procedure whereby "tags" or triggers are attached to a word or phrase that links it to another document located anywhere on the Internet. The documents created by HTML are stored on computers known as servers and can include straight text, visual images, streaming video, and audio clips. Documents belong to a web site that has a specific address such as "www.avemaria.edu." The last three letters represent a "top level" identification (for example, "edu" stands for education and "com" stands for a commercial enterprise) and the middle part of the name designates the actual site (Ave Maria University).

Net browsers such as Navigator provided by Netscape (based on the original Mosaic model) or Microsoft's Internet Explorer enable users to

"explore" the web rather effortlessly. They are highly versatile navigational tools that enable users to access, display, and print documents; they also give users the ability to link to other documents at any location on the web. Hyperlinks can create a maze of interconnected documents and web sites that can sometimes confuse users but also greatly expand opportunities for research and investigation.

The web has transformed the Internet into a user-friendly medium because the web page is an intuitively obvious interface for even the most novice user. More significantly, according to Samuelson and Varian, "the back-end protocols for authoring and distributing web pages (HTML and HTTP) were easy to understand and use as well, facilitating the rapid deployment of web servers."[15] The diversity and heterogeneity of current web sites is evidence of the accuracy of this assessment.

Despite its brief history, the World Wide Web itself has already become a vast, tangled network. Web sites were first deployed at major universities and research centers, but now proliferate throughout cyberspace at schools, hospitals, corporations, and many other organizations. According to the Internet Systems Consortium, there were approximately 400 million active domains operating on the web in 2005.[16] Even individuals or small businesses have established their own web pages. These web pages will undoubtedly be the vehicle for the acceleration of electronic commerce and many other network-based activities like education or fund raising. Web-based marketing is beginning to show significant results, and as a consequence ad banners and commercial messages can be found now in almost every region of cyberspace.

The plethora of web sites has created a density of information that can make it difficult for users to locate a particular site. Search engines such as those provided by Microsoft or Google can help in this process, but even they are sometimes ineffectual in the face of such voluminous data. Part of the problem, of course, is that the web is just too large and too volatile to index properly, but these search engines have made great strides in this regard.

Regardless of the difficulties that users encounter trying to navigate their way through cyberspace, the web continues to rapidly gain in popularity. It is quickly becoming its own unique institution, taking the place of libraries, print catalogs, and even traditional news media for many users. It can be a rich source of research, news and information, and entertainment. And as more and more users develop their own sites, it has helped bring about the democratization of information predicted by many Internet visionaries.

Finally, there are predictions that we will soon witness the birth of a more sophisticated next-generation web thanks once again to the work of Berners-Lee and his colleagues. This is being called the *Semantic Web*, because it will be able to understand human languages, thanks in part to the extensive use of XML, a language that can tag words and phrases so that computers know what they mean. For example, if a bot travels to a site and sees an 11:05 AM departure time the bot does not know what 11:05 means unless it is accompanied by a tag <departure time>. The ultimate objective is to transform the web into a giant intellect so that "every computer connected to the Internet would have access to all the knowledge that humankind has accumulated in science, business, and the arts since we began painting the walls of caves 30,000 years ago."[17]

Electronic Commerce

Electronic commerce (or e-commerce) refers to trade that occurs on the Internet. Thanks to the infamous dot.com debacle in 2000, euphoria about e-commerce and the Net has faded, as we have come to appreciate that many e-commerce ventures were no more than phantom edifices. But no one is dismissing the likelihood that this global network will be a main thoroughfare of commerce in the near future. In fact, some business experts predict a "new wave of disruption" as the web transforms industries such as jewelry, hotels, and real estate. Amazon.com has already jumped into the jewelry business, buying diamonds wholesale for as little as $500 and reselling them for $575.[18]

What are some of the general benefits of e-commerce? First, it eliminates the constraints of time and space and thereby provides extraordinary convenience for consumers. As noted, the Net is a fundamentally asynchronous technology so users can do their browsing and shopping at any time. Second, the Internet is a low-cost communications technology so it can greatly reduce overhead and transaction costs. According to Mandel and Hof, "the Internet is a tool that dramatically lowers the cost of communication, [and] that means it can radically alter any industry or activity that depends heavily on the flow of information."[19] It is, of course, much less expensive to operate one virtual book store (amazon.com) than a chain of physical stores. And the more digitizable the product, the lower the cost structure, because those products do not require an infrastructure for distribution. It is even more advantageous for a company such as Monster.com to provide an online exchange for jobs than it is for amazon.com to sell books; the online exchange does not have to worry about warehousing a physical product and delivering that product to its customers. The fewer the service requirements and the lower the logistic requirements, the more scalable the online business, that is, the easier one can grow the business without additional costs.

A third advantage of online commerce is the ability to customize sales and advertising to each individual consumer. A web shopper's every move in cyberspace can be traced and this allows vendors to compile a profile of

a consumer's preferences. According to the *Economist*, "With this feedback, online merchants can further differentiate themselves from their physical world competitors by customizing their shop or service for each customer."[20] For example, Amazon uses collaborative filtering technology that enables it to analyze a customer's purchases and to suggest other books the customer might like based on what people with similar purchase histories have bought.

The e-business landscape is complicated, but it is useful to follow Applegate's helpful distinctions. She categorizes some companies as digital infrastructure providers including IBM, Cisco, AT&T, and Microsoft; they provide the servers or physical networks that make electronic commerce possible. In a second category she includes companies that operate in the Internet's distribution channel: focused distributors and portals. Focused distributors provide products and services primarily on the web and portals serve as gateways to the Net.[21] Under the category of focused distributors there are four basic digital business models: business to consumer (B2C), consumer to business (C2B), business to business (B2B), and consumer to consumer (C2C).

The B2C model involves direct sales to consumers and includes companies such as Amazon.com. With 1 billion million people already online, the potential here is obviously vast. For the consumer, the big attraction is convenience—with a single click of the mouse an order for clothes, books, fine wine, or groceries can be placed at a web site. For the retailer a key advantage is lower costs. Hence the B2C model allows for quick scalability of one's business. Only one web site is needed to service customers all over the world. There may be a high initial investment for a computer system, but unlike traditional retailers there is no need to continually invest in new stores and other physical assets in order to increase revenues.

The C2B model is epitomized by Priceline, the company founded by Jay Walker that allows customers to name their price for various objects such as airline tickets or hotel rooms. According to Priceline's founder, Jay Walker, "In the traditional model of commerce, a seller advertises a unit of supply in the marketplace at a specified price, and a buyer takes it or leaves it. Priceline turns that model around. We allow a buyer to advertise a unit of demand to a group of sellers. The sellers can then decide whether to fulfill that demand or not. In effect, we provide a mechanism for collecting and forwarding units of demand to interested sellers."[22]

The third model is C2C, and the prime example is eBay, the online auction service that acts as an intermediary for customers who want to auction off various goods to other customers. The eBay operation illustrates how online commerce can function with extraordinary efficiency. The buyers and sellers do all the work: sellers pay a fee to eBay for the opportunity to auction their wares, and when the auction ends the seller and buyer negotiate over the payment and shipping. For its role as an intermediary eBay

normally receives between 7% and 18% of the sale price. Because its customers do most of the work and take most of the risk, some have concluded that eBay is the perfect virtual business.

Finally, B2B refers to electronic commerce between two organizations and includes procurement, inventory management, sales service and support, and so forth. Perhaps the greatest potential in the B2B market is the remarkable growth of trading sites that range in complexity from online catalogs to public exchanges where buyers and sellers come together to exchange goods. Ventro, formerly known as Chemdex, was a public exchange for the medical equipment industry. These exchanges are open to a much larger group of buyers than a private network, and this greatly enhances the potential market for the sellers.

In addition to focused distributors we find *portals*. *Horizontal portals* such as Yahoo and Netscape function as gateways to the web by providing an initial point of access from which users can connect to various sites. *Vertical portals* such as Quicken.com in the area of financial services provide "deep content" in one area. The boundaries between portals and focused distributors have become increasingly murky. According to Applegate, many portals now "serve not only as gateways but also as destinations where people stop and conduct business."[23]

Despite the success of companies like Amazon.com, eBay, and Expedia, some consumers are still reluctant to embrace electronic commerce. Some fear becoming the victims of fraud or scams that are easier to execute thanks to the anonymous nature of Internet transactions. Others are concerned about the tenuous security of the Internet and worry that the price of convenience may be a loss of privacy. The accretion of social trust will be accelerated by "trustworthy" systems that are secure and respectful of privacy rights. These systems go a long way to inspiring more confidence in web-based transactions.

Although electronic commerce web sites have made the greatest progress in the United States, they are also proliferating in many other countries such as China. The Chinese government has encouraged the private sectors to develop web-based businesses, and entrepreneurs are responding enthusiastically. As these sites gain in popularity they could transform China's outdated retail business and enable manufacturers to automate purchases from suppliers. According to Einhorn, "the global implications of this are enormous—as web-based e-commerce spreads, traders around the world could link directly with suppliers and retailers across China."[24] Chinese-language portals such as sina.com and china.com are also emerging to serve as gateways to the Net for Chinese citizens.

The Internet can never return to its halcyon days when it was frequented only by technology buffs and academic researchers who formed an intimate and knowledgeable online community. As electronic commerce intensifies, the Net will continue to evolve, and to a large extent its future is

in the hands of many different stakeholders who were not involved in the Internet's early days and who have a much more pragmatic and profit-oriented attitude about the Net than its early founders. We turn now to a few remarks about the downside of this phenomenon.

Social Problems and Social Costs

The Internet's popularity and commercialization has led to some familiar social problems and frictions in cyberspace. The erosion of privacy, the emergence of perverted forms of speech, and illegitimate copying of music and video files represent just some of these problems. At the same time, e-commerce vendors have been victimized by fraud and attacks by hackers. In the remaining chapters of this book we diagnose and analyze these problems, and review some equitable resolutions.

At this stage, however, it would be instructive to consider broad philosophical differences for dealing with these difficulties. In Lessig's terms, is the optimum solution to be found in law, the market, code, or social norms? It is naïve to think that any one of these four modalities of regulation such as law can constrain a problem such as privacy erosion in cyberspace. For complex problems the proper solution will undoubtedly be found in the interplay of law, code, and the market. The question becomes which of these forces should have primacy? Which one should generally take the lead in controlling the Net? One's answer to this query depends on one's faith in the market forces or on one's ideological assumptions about the efficacy of government regulation.

In economic terms, some of these frictions in cyberspace can be described as *social costs* or *negative externalities*. Coase explains that social costs are generated by "those actions of business firms which have harmful effects on others."[25] Certain social harms we have been discussing can be viewed from this perspective, that is, as harmful byproducts of certain Internet transactions. For example, the erosion of privacy, which often results when information is exchanged between two parties, or the transmission of disruptive forms of speech like spam, would fall into this category. Social costs represent failures of the market system or "market imperfections"; they are costs borne involuntarily by others that are not reflected in the price of the good whose production created those costs. In the case of privacy, the sale of personally identifiable data to third parties is an externality because the cost is imposed on the individual whose data are sold, and hence that cost is ignored by the seller. When regarded from an economist's viewpoint, the issue becomes one of weighing the economic benefits of the sale of this data against the social costs of privacy erosion.

Consider the problem of unsolicited electronic e-mail (spam). The production and distribution of spam messages advertising some product or service are minimal, but the real costs are shifted to other parties in cyberspace such as ISPs and even the recipients of those messages. Because the true costs of spam are not internalized by its "producers" spam is overproduced resulting in a lack of allocative efficiency, or efficiency in the allocation of society's economic resources.

But what should be done about these market failures, these externalities that now plague cyberspace just as they plague real space? Let us review several ideologies for how best to deal with market imperfections and those social harms we find on the Net.

The Invisible Hand

The vast majority of users agree that spam is a menace, but when it comes to doing something about spam opinions diverge. Do we need more government intervention, public polices, to contain the flood of junk mail and restore economic efficiency? Economists such as Ronald Coase of the Chicago School are skeptical of the government and tend to put more faith in the invisible hand of the marketplace to solve problems like spam. The impersonal forces of the market can often do a better job of fixing market imperfections than the vested interests of other marketplace participants (such as government regulators). Much of Coase's work has drawn attention to the limitations of government regulation as a solution to the problem of negative externalities. The government's regulatory agencies often do not understand the industries they are trying to regulate, a problem that could be exacerbated in contexts where sophisticated technologies are involved. Coase and others have also frequently noted the inefficiencies of large, centralized bureaucracies and the persistence of organizational inertia. Finally, there is always the potential for *capture*, a process whereby those being regulated influence regulators so that they no longer act in the public interest. According to Coase, there are "few more unpleasant sights than an unholy alliance between the regulators and the regulated industry to solve the problem of competition by suppressing it."[26]

What Coase and others favor as an alternative is greater reliance on the marketplace. Consider, for example, the problem of privacy erosion in cyberspace. We can certainly enact laws to deal with this problem but maybe the markets will effect a more efficient, welfare-enhancing solution. Those who favor this approach presume that market pressures will force vendors to respect privacy rights at a level consistent with the needs and interests of consumers. According to Reidenberg, the U.S. approach to privacy relies in part on just such a market-based solution; data protection in the U.S. is a "question of economic power rather than political right."[27]

But the marketplace has often proved to be an inadequate forum for addressing social problems. The market's reaction to those problems is often reactive and inequitable. Lessig, for one, sees the invisible hand threatening

"liberty and openness" in cyberspace.[28] Economists like Pigou categorically reject the viability of market-based solutions to these externality problems: "No 'invisible hand' can be relied upon to produce a good arrangement of the whole from a combination of separate treatments of the parts. It is, therefore, necessary that an authority of wider reach should intervene and should tackle [society's] collective problems. . . ."[29] According to this view, the marketplace always functions as an important constraint on behavior, but it should not take priority over other regulatory forces such as law, norms, and code.

Regulating the Net: The Visible Hand

As Pigou suggested, the alternative to the market as primary regulator is the greater reliance on policy constraints imposed by government. But can the Net be regulated—is it really "regulable" in the same way that the physical world can be subjected to rules and regulations of local sovereigns? Can the unrestricted freedom of cyberspace be reined in by government forces?

The Internet defies regulation for several key reasons. First, its distributed architecture and resilient design makes the Net hard to control. Packet-switching technology, for example, has meant that it's not so easy to stop the flow of information. As John Gilmore puts it, "Information can take so many alternative routes when one of the nodes of the network is removed that the Net is almost immortally flexible. . . . The Net interprets censorship as damage and routes around it."[30] The Internet's lack of a physical center means that it has no moral center that can be held accountable for information flowing over the network.

Second, there is the Internet's content, digital information, 1s and 0s that can be transmitted through cyberspace with ease and stored on the recipient's hard drive. As Negroponte observed, "The information superhighway is about the global movement of weightless bits at the speed of light."[31] All forms of information including images and voice can be digitized, and a digital file is especially difficult to contain. One consequence of this is that digital file sharing technologies such as those developed by KaZaA and Gnutella are threatening to undermine the economics of the music and movie industries.

Finally, governments that seek to control or regulate the Net face an array of jurisdictional conundrums. As we have seen, a fundamental problem with a particular sovereignty imposing its will on the Internet is that laws and regulations are based on geography—they have force only within a certain territorial area, for example, a state, county, or nation. As one jurist put it: "All law is prima facie territorial."[32] Moreover, because the Internet is a borderless global technology, it is almost impossible for any country to enforce the laws or restrictions it seeks to impose on this sprawling region of cyberspace. If the U.S. decides to outlaw pornography, it can only effectively enforce this restriction among U.S. purveyors of pornography. It cannot restrict vendors located in Europe or the Caribbean from making pornography available on the Internet for everyone to see. It can, of course, put the burden on Internet providers and hold them liable for transmitting the illicit material no matter where its source is located. But this seems to be an unfair and unworkable solution because it is expensive and difficult for ISPs to detect and properly filter out all communications with pornographic elements.

Despite these obstacles, local sovereignties will not be deterred from regulating the Net. Consider France's efforts to prevent Yahoo from allowing Nazi memorabilia to be sold on its auction web sites, despite the fact that the server hosting these sites is located in the U.S. Those bringing the suit against Yahoo claimed that the company violated local French law. But to what extent should the global Internet be subjected to local law? The potential problem, according to Zittrain, is that "anyone posting information on the Internet is unduly open to nearly any sovereign's jurisdiction, since that information could have an effect around the world."[33]

In addition to the control of content, governments may pursue other forms of Internet regulation, such as the regulation of the information infrastructure or regulation of e-commerce. For example, a particular sovereignty might be concerned with preserving open and equitable access on the Internet, but give free reign to content providers along with the focused distributors and portals engaged in e-commerce.

Governments that do seek to regulate e-commerce might do so by regulating privacy or data protection or by insisting on certain security standards for a web site. The European Union Privacy Directive, for example, lays out strict privacy rules for companies doing business within the European Union (E.U.). In the U.S., however, the preferred solution has been self-regulation.

All sovereignties must make decisions about the scope of Internet regulations. Should they aspire to developing regulations to protect the infrastructure or focus on content controls? Once the appropriate scope is defined, sovereignties must decide whether they should apply existing laws or craft new ones. For example, should the U.S. apply existing intellectual property laws to the web or is it necessary to develop new ones? According to Samuelson, "Although some commentators have suggested that copyright law is outmoded in the Internet environment, the general view in the U.S. and the EU is that copyright law can be applied and adapted to protect expressive works in digital form."[34]

Some countries, unfortunately, have been overly aggressive in Internet regulations. Despite its encouragement of web-based business, the Chinese government remains exceedingly anxious about the Internet, and they have made it quite clear that "by linking with the Internet, we do not mean

the absolute freedom of information."[35] Chinese officials use a firewall to block access to pornographic and other objectionable web sites, such as those operated by human rights groups. China's iron grip on political discourse has been tested by Internet access, but this country has responded with its usual heavy-handed and repressive tactics.

Finally, as we have implied, there are perils in having each local jurisdiction impose its own laws on the Net. Different privacy laws, for example, could disrupt the flow of e-commerce or impede other information exchanges. If this borderless global technology is to be properly regulated, shouldn't there be a set of international standards? Don't we need a global law for this global technology?

There has been some effort made to harmonize laws pertaining to the Internet. Consider, for example, the WIPO Copyright Treaty, which stipulates how copyright laws will be applied to digital works. In addition, countries outside the E.U. have begun to embrace its Privacy Directive. Argentina, Australia, Canada, Switzerland, and New Zealand have either adopted the E.U. Directive or they are working on a set of rules that are heavily influenced by that Directive.[36] Some in the U.S. see this trend toward convergence as a threat to national sovereignty, but others believe that a simple global standard is the only way to ensure that privacy rights are recognized and enforced throughout the world.

Although harmonization is sound in theory, it will be immensely difficult to accomplish in practice thanks to deeply embedded cultural and legal differences between most countries. Given this difficulty, Samuleson recommends that nations should instead strive for "policy interoperability" instead of full harmonization, that is, "agreeing on goals a policy should achieve, while recognizing that nations may adopt somewhat different policy means to implement goals."[37]

A "Bottom-Up" Approach: The Sovereignty of Code

In Chapter 1, we alluded to the Net's empowerment of the individual through its code. Thanks to strong encryption programs, for instance, it is more difficult for the state to conduct surveillance on confidential electronic communications. Similarly, filtering technologies give individuals the power to limit content or format the information they wish to receive. Electronic anonymity also frustrates lawmakers' efforts to hold individuals accountable for their online actions. The Net empowers individuals through technology. It is shifting control from the state to the individual, and this is a source of great consternation for many government leaders.

The individual's empowerment through code makes possible a more bottom-up approach to regulation that some users and civil libertarians favor. But can a case be made for letting the Internet organize and moderate itself as much as possible? According to David Post, "there are some problems on the Internet best solved by these messy, disordered, semi-chaotic, unplanned, decentralized systems, . . . and the costs that necessarily accompany such unplanned disorder may sometimes be worth bearing."[38] This messy bottom-up approach Post describes is not a panacea for the Internet's various externalities, but it may be an adequate means of regulating conduct and addressing some aspects of the social problems described in the chapters ahead.

There is surely much to be said for reliance on the constraints imposed by technology in the hands of individuals. In some ways it seems preferable to the regulatory regime of government. It's nonintrusive, simpler, less expensive, and gives users the ultimate choice about what they want to see or not see. Bottom-up constraints also avoid the expensive government infrastructure that inevitably accompanies a regulatory scheme. In addition, this approach fits with the cultural shift now taking place in countries like the U.S. whose citizens are increasingly anti-bureaucratic. Instead of reliance on bureaucracy and public policy to solve society's ills, they favor individual empowerment and local control whenever possible.

As we observed in Chapter 1, however, some legal scholars have perceptively made the case that technical solutions implemented by private parties can sometimes be more restrictive than actions taken by a democratic state. As Seth Finkelstein writes, "because of a perspective that might be rendered 'government action bad, private action good' there's great unwillingness to think about complicated social systems, or private parties acting as agents of censorship."[39]

In his critique of filtering systems such as PICS, Lawrence Lessig has made similar observations. PICS, which stands for Platform for Internet Content Selection, is a labeling standard that provides a way of labeling and blocking online material. It can be used by parents or schools to block access to a web site with pornographic material or one filled with virulent hate speech. According to Lessig, the widespread deployment of this technology can yield a "tyranny of the code" as those in positions of authority impose their own standards on unsuspecting users.[40]

The power and potential of blocking software like PICS has not been lost on civil libertarians who have begun to better appreciate how these technologies can undermine the free flow of information far more effectively than government-imposed censorship. The threat to freedom may be more subtle and dispersed, but the end result is still the sort of social domination, now effected by private parties, that the Net is designed to resist.

The French philosopher Michel Foucault appreciated the import of this difference as well. In his writings on the nature of power, he differentiated between explicit state commands emanating from the sovereign power and a more covert and implicit exercise of domination. The latter normally has taken the form of surveillance, but it can take other forms as well. Accord-

ing to Foucault, "we have the emergence or rather the invention of a new mechanism of power possessed of a highly specific procedural technique. It is a type of power which is constantly exercised by means of surveillance rather than in a discontinuous manner by means of a system of levies or obligations distributed over time."[41] This clearly echoes Lessig's concern about the "tyranny of the code," a tyranny that can come from different and nonobvious sources.

We are left then with a provocative but seminal question—should control and regulation of the Internet for the most part be left in the hands of private parties and the corrective technologies that they create and distribute in the marketplace? Or should we embrace a more top-down approach? Should the Internet be regulated more directly to contain its social costs without the collateral damage that can accompany the bottom-up approach? Are the sinews of Internet stability best found in the rational laws and regulations emanating from a sovereign power or an international body? Or are they found in the architectures of the Net responsibly deployed by individuals?

Internet Governance

Although there is some disagreement on how the Internet should be regulated through government intervention, no one questions the need for some type of governance and technical coordination. No matter how opposed one is to regulatory oversight, the Net cannot survive without this type of coordination. There must be governing bodies that handle ordinary and routine technical matters such as the determination of technical standards and the management of domain names and IP addresses. For our purposes, *governance* refers to managing these matters rather than regulating the Net through content controls or other mechanisms.

Two major policy groups that provide such governance are the World Wide Web Consortium, an international standards setting body, and the Internet Engineering Task Force (IETF), which develops technical standards such as communications protocols. According to the *Economist*, a culture of "cautious deliberation" prevails within the IETF, which strives to be democratic in its decision-making processes. Anybody can join the IETF and any member can propose a standard "and so start a process that is formal enough to ensure that all get a hearing, but light enough to avoid bureaucracy."[42]

The Domain Name System (DNS) also needs coordination. The DNS maps the domain names of organizations such as eBay to the actual numeric Internet Protocol address (e.g., 709.14.3.26). The DNS is a hierarchical system divided into separate domains. When a domain name is

invoked by a browser the request is forwarded to the DNS server, which is normally operated by an ISP, and that server locates the databases for each subdomain. If the domain name is www.loyola.edu, the DNS server first locates the server for ".edu," which is the Top Level Domain (TLD); it then finds the server for "loyola," the Second Level Domain, and so forth. Using this method the web page is found and transmitted back to the recipient.

This system was formerly administered by a small private company called Network Solutions International (NSI), which charged $50 for the registration of a domain name and usually awarded the name on a first-come, first-served basis. As the Internet became commercialized, disenchantment with the NSI arrangement escalated. As a result, after some political maneuvering, the domain name system is now in the hands of the Internet Corporation for Assigned Names and Numbers (ICANN). ICANN is an international, nonprofit organization with full responsibility for the DNS. ICANN itself does not actually distribute domain names. That task is delegated to domain name registrars such as VeriSign. ICANN determines the policies for domain name distribution, and it has the final say for selecting firms that qualify as registrars.

Domain names were introduced to impose some order on the Net, and originally there were six TLDs: .com, .net, .org, .edu, .gov, and .mil. ICANN has recently decided to create several new TLDs, such as, .aero. (air transport companies), .coop (cooperatives), .biz (business), .museum (museums), .name (individuals), .pro (professionals such as lawyers), and .info (nonrestricted use). The purpose of these new extensions is to handle the overusage of popular TLDs such as .com and .org. It remains to be seen whether these new extensions (like .biz) will be embraced by the public and become as popular as the original TLDs such as .com.

To its credit, ICANN has acted swiftly and deliberately to deal with the issue of cybersquatting and other domain name disputes. In October, 1999, it established the Uniform Dispute Resolution Policy for adjudicating such disputes and protecting legitimate trademarks. That policy is discussed in more detail in Chapter 4, in the context of the treatment of trademark law and the Lanham Act.

ICANN is currently governed by a board of eighteen members; nine of those members are elected by the at-large membership. Critics of ICANN contend that despite its claims to represent an international constituency, Americans dominate ICANN. Moreover, they insist that its structures are not democratic enough and that it does not give average users enough say in its governing procedures. Whether these criticisms will undermine ICANN's authority is anyone's guess at this point, but its supporters say that ICANN has the potential to emerge as a model of consensus building and international cooperation the global Internet community demands.

Internet Regulation and Ethics

At this stage of the Internet's rapid evolution, it would be presumptuous to predict which, if any, of the regulatory approaches described here might actually prevail. There will undoubtedly be some mixture of bottom-up controls combined with top-down regulations. The real question is how expansive a role the government will end up playing in regulating the Internet. There is a case to be made that this role should be substantial given the importance of cyberspace for the future of commerce and for many other social interactions. There is also understandable wariness about the unpredictability of trusting the Internet to regulate itself. Without the government's sustained efforts to ensure a level playing field, companies like Microsoft and Google could exert undue influence on e-commerce and monopolize essential facilities. Also, supporters of more extensive government regulations raise legitimate concerns about the poor state of efforts in the U.S. to handle online privacy through self-regulation.

However, if the state does intend to expand its regulatory role in the recalcitrant region of cyberspace, it will confront at least two formidable challenges. First, it must try to apply its own territorially based laws to a global entity. For example, anti-Semitic hate speech is illegal in Germany, but purveyors of neo-Nazi web sites have relocated their servers to the U.S. to avoid German jurisdiction. Second, it must contend with code that has radically empowered the individual—individuals have at their disposal programs like anonymizer.com which can block IP address tracking. These obstacles appear to have weakened the state's sovereignty and given the individual the upper hand.

It would be premature, however, to underestimate the power of the state and to toll the death knell for its sovereignty. As Michel Foucault writes, "wherever there is power, there is resistance."[43] The state will certainly resist this state of affairs and seek to retrieve its lost dominance and diminished sovereignty. It may, for instance, use its vast power to tightly regulate ISPs or to demand that other private surrogates carry out its regulatory regime. Public policy makers too recognize the power of code as a constraint in cyberspace and might be willing to mandate the use of certain codes (such as filtering and IP address tracking) to counteract the difficulties of regulating cyberspace through fiat alone. As Lessig observes, the state will work to increase the very regulability of cyberspace by exercising control over its code.[44]

What we are left with, then, is a power struggle between a frustrated state and a newly empowered Internet community. At the epicenter of that struggle is the code of cyberspace. In many respects, the code is a far more effective constraint than law, norms, or the marketplace. One can envision many possibilities on both sides for using that code to gain control. For example, the architectures of the Internet currently facilitate electronic anonymity, but the state could respond by requiring that ISPs use code that mandates digital identity and the traceability of all Internet transactions.

What makes this struggle so perilous is the facility with which the code of the Internet can be manipulated. Andrew Shapiro describes the Internet's capacity for empowering individual users as the "control revolution." He argues that the state's resistance to that revolution will "become more refined as governments become more adept at influencing code without running afoul of constitutional limitations or public opposition."[45]

Code is such a powerful regulator in the hands of the state or individuals because of its malleability and obscurity, its flexible ability to regulate or shape behavior gradually and inconspicuously. Code does not always constrain or influence behavior openly and directly, in a way that is transparent to those it affects. This contrasts sharply with the constraint of law, because the process of crafting laws through democratic procedures is subject to considerable public scrutiny.

Hence the paramount importance of ethics in all of this. Although we do not take a stand on the preferability of a bottom-up or a top-down regulatory philosophy, we contend that whatever approach becomes dominant, there must be careful attention paid to basic human values such as autonomy, privacy, and security. Informal social controls abetted by technology may have the potential to provide effective and fair-minded regulations of cyberspace conduct, but only if those stakeholders involved are committed to responsible behavior. This will help to minimize any negative effects on human rights that these corrective technologies (such as filtering) can bring about if they are carelessly deployed.

Likewise, a top-down legislative process must be guided by these same core values. Governments must not overreact to the control revolution with restrictive laws or bypass the democratic process and manipulate Internet architectures to curtail basic human freedoms and rights merely for the sake of greater order and stability in cyberspace. They too must behave responsibly in their attempts to regulate cyberspace.

As we argued in Chapter 1, what is of primary and utmost significance is the preservation in cyberspace of those transcendent human goods and moral values, which are so basic for the realization of human flourishing. *Moral values must be the ultimate regulator of cyberspace, not the code of engineers.* This orientation helps to ensure that abuses of the code are kept to a minimum. If Internet stakeholders, including public policy makers, software developers, educators, and corporate executives, act prudently and responsibly, they will be vigilant and conscientious about respecting these values. As a result, they will find themselves guided by a moral wisdom that encourages care for others and a sense of measure concerning the public affairs of the Internet. This will also help to achieve a reasonable equilibrium between the state and other Internet stakeholders.

In the next several chapters we discuss what constitutes responsible approaches to cyberspace regulation. In the course of that discussion we consider how code can be responsibly designed, developed, and utilized. We also focus on how the core moral values can be applied to some of the troubling dilemmas now emerging in cyberspace. To be sure, the application of those values is not an exact science and there will often be room for reasonable people to disagree. But if there is a shared conviction that the Internet must be governed by these broad moral standards, it will be easier to equitably resolve these inevitable conflicts.

Discussion Questions

1. Discuss the pros and cons of extensive government regulation of the Internet either by a local sovereign government or by an international body specifically constituted for this purpose.
2. Evaluate the "bottom-up" approach to regulation as it was presented in this chapter.
3. In what ways does the structure and present architecture of the Internet affect the choice of an optimal regulatory structure?
4. What is ICANN and what does it do?

Case Studies

L'Affair Yahoo!

Company Background

Yahoo! was founded in 1994 by David Filo and Jerry Yang, two Ph.D. students at Stanford University. It was originally developed as a portal, that is, a gateway or guide to the web and as a way to keep track of web site addresses. It also incorporated search functionality. This fashionable guide to available web sites quickly evolved into a commercial site and thriving business. In 1995 Yahoo took on Tim Koogle as its CEO. From the outset, Yahoo under Koogle's guidance saw itself as a media company and not just as a search engine. During 1996 and 1997, Yahoo added considerable content and communication facilities as it evolved into a full-fledged Internet portal. Yahoo's primary services are called *properties*. These properties included *navigational services*, which help users find web sites and other information more easily. It also includes *community properties*, which help users communicate with one another. For example, users could access the Yahoo Address Book, which allowed them to use an address book from any connected system. There were also e-commerce properties for shopping or making travel arrangements. Millions of people now use Yahoo for e-mail, instant messaging, scheduling, personal web pages, chat rooms, job searches, and auction sites.

Yahoo generates most of its revenues through advertising and deals with e-commerce partners. The company reaches 60% of all Net users worldwide, and it tracks the visits of 166 million users. Yahoo has also expanded mightily into overseas markets. Foreign users now amount to 40% of Yahoo's customer base. Yahoo has the biggest global reach of any Internet brand—it offers 23 local versions in 12 different languages. Yahoo has prided itself on good relations with foreign governments. According to Forbes, Yahoo devotes much energy to "hitting the international conferences and meeting heads of state to talk Internet policy and plead Yahoo's local interests."[46]

Thus, Yahoo provides a variety of means by which people from all over the world can communicate and interact over the Internet. Yahoo's auction site allows anyone to post an item for sale and solicit bids from any computer user around the world. Yahoo sends an e-mail notification to the highest bidder and seller with the respective contact information. Yahoo is never a party to the transaction, and the buyer and seller are responsible for payment and shipment of goods. Yahoo informs sellers that they must comply with company polices and may not offer items to buyers in jurisdictions in which the sale of such item violates the jurisdiction's applicable laws. Yahoo, however, does not actively regulate the content of each posting and individuals have posted offensive material including Nazi-related propaganda and material.[47]

The French Resistance

During the spring of 2000, Yahoo's relations with the French government ran into serious problems. In April, two French associations, the French Union of Jewish Students and the International League against Racism and Anti-Semitism (La Ligue Contre Racise et L'Antisemitisme [LICRA]), filed suit against Yahoo, demanding that they remove swastika flags and other Nazi memorabilia from their American web site. French law expressly prohibits the display or sale of objects that incite racial hatred and this includes any World War II Nazi memorabilia. The French Court cited 1,000 Nazi and Third Reich–related objects for sale on Yahoo auction sites including Hitler's autobiography, *Mein Kampf*, and *The Protocol of the Elders of Zion*, an infamous anti-Semitic book. Any French citizen could access these materials on Yahoo.com directly or through a link on Yahoo.fr. (Yahoo's regional web sites such as Yahoo! France [http://www.yahoo.fr] use the local region's primary language, targets the local citizens, and operate under local laws.)

In May 2000, Judge Jean-Jacques Gomez of the Tribunal de Grande Instance de Paris ruled in favor of these two groups. He concluded that Yahoo

had violated French law and offended the "collected memory" of France. He ordered Yahoo to make it impossible for French users to access any auction site that contained illegal Nazi items such as relics, insignia, and flags. He also ordered Yahoo "to eliminate French citizens' access to web pages on Yahoo.com displaying text, extracts, or quotations from *Mein Kampf* and *The Protocol of the Elders of Zion*."[48]

Yahoo's lawyers claimed that the company was powerless to obey this order, maintaining that it would not be technically feasible to accomplish the task of identifying web users by national origin and blocking access to the contested sites. Yahoo also claimed that the French court lacked jurisdiction, because its principal place of business was in Santa Clara, California. The Judge dismissed the latter claim, and he assembled a panel of three experts to determine whether or not Yahoo's assessment regarding technical feasibility was correct.

The panel, consisting of three individuals representing France, Europe, and the U.S., was charged with answering this question: is it technically possible for Yahoo to comply with the court order, and, if not, to what extent can compliance be achieved? The panel concluded that foolproof 100% compliance was impossible. But it also concluded that Yahoo could block up to 90% of French users by using several levels of detection. Over 60% could be blocked by the same technology that Yahoo used to customize the site for French users, that is, by providing French users with French banner ads. This entailed tracking their IP address, which in most cases reveals the physical location of the user. This would not work, however, for the subscribers of some ISP services (such as AOL customers), because the ISPs assign temporary IP addresses. However, it was estimated that another 20% to 30% could be identified by requiring users to fill out a "declaration of nationality."

Of course, each method of detection could be easily circumvented. One could employ an anonymizer such as www.anonymizer.com to prevent the IP address from being revealed. And one could also lie about one's nationality on the declaration form.

But Judge Gomez was satisfied that Yahoo could identify most of the users logging on from France. Hence, in November 2000, the judge reissued the preliminary injunction (*Ordinance en referé*) against Yahoo that he had first issued in May. Yahoo was ordered to install a filtering system (or equivalent technology) to block French citizens from these problematic sites that auction Nazi objects or that present any Nazi sympathy or holocaust denial. Yahoo was informed that it had 90 days to comply with the court order or face a fine of up to 100,000 francs (about $13,000) per day. In his ruling the judge referred to Yahoo's ability to detect French web users because it already preselects them for its French-language banner ads. The judge also pointed to Yahoo's other restrictions, citing its policy "of not allowing the sale of drugs, human organs or living animals on its auctions sites."[49]

This unique case triggers many difficult jurisdictional issues. On one hand, France has the right to assert jurisdiction over its citizens and to enforce its own laws. But how can it enforce its laws against a company located in the U.S.? One of Yahoo's lawyers predicted "that any effort by French authorities to enforce Judge Gomez's judgement in a United States court against Yahoo's United States assets would fail because of the First Amendment, which protects hate speech."[50] Other commentators such as the Center for Democracy and Technology in the United States immediately criticized the decision as a grave threat to freedom of expression on the Internet.

Yahoo's Dilemma

Yahoo officials must now decide whether or not to comply with the French Court's order. They had several options. They could adopt a defensive posture: ignore the Court order and continue to allow its auctions sites with these controversial items to be made available to French citizens. The company might combine this strategy with an appeal of the French Court's decision. Or it could take blocking measures to shut out French residents from the contested sites to ensure compliance even if they are not fully effective. It also has the option of banning hate material including these Nazi-related items from all of its auction web sites. This might be accomplished by using software that scans the items before they are made available for sale. This course of action would be the most drastic; it would be a departure from Yahoo's longstanding policy against the monitoring of its web properties.

As the November decision began to sink in, Koogle and his colleagues realized that they faced an insuperably difficult decision. How could it balance the interests of its diverse stakeholders without getting embroiled in a protracted legal battle with the French government?

Questions

1. In your opinion, what should Yahoo do about this situation? Should it make concessions to the French Government?
2. Do you agree with the French court's efforts to enforce local laws against anti-Semitic hate speech against Yahoo?
3. What are the broader implications of this case for the future of free speech on the Internet?

A Case of Libel?

Mr. Joseph Gutnick, a prominent Australian business man, was quite shocked when he came across some unflattering remarks about himself in an online article in *Barron's*:

> Some of Gutnick's business dealings with religious charities raise uncomfortable questions. A *Barron's* investigation found that several charities traded heavily in stocks promoted by Gutnick. Although the charities profited, other investors were left with heavy losses. . . . In addition, Gutnick has had dealings with Nachum Goldberg, who is currently serving five years in an Australian prison for tax evasion that involved charities.[51]

Gutnick decided to file suit for libel. *Barron's* is owned by Dow Jones & Company, publisher of the *Wall Street Journal*, which has corporate headquarters in the U.S. But Mr. Gutnick and his lawyers wanted to file the libel suit in his home state of Victoria where the libel laws are quite strict.

Dow Jones, on the other hand, sought to have the case heard in the U.S., where *Barrons online* is written and disseminated. The company feared the precedent that would be set if the case were heard in Australia. In the future, posting material online could leave them open to multiple law suits in many different jurisdictions. Accordingly, Dow Jones' lawyers argued that the U.S. jurisdiction was the fairest place to hear this dispute.

But the High Court of Australia ruled that Gutnick could sue in his home state of Victoria, reasoning that this "is where the damage to his reputation of which he complains in his action is alleged to have occurred, for it is there that the publications of which he complains were comprehensible by readers."[52] According to Zittrain, the Australian High Court dismissed Dow Jones' "pile on" argument "that Gutnick could next sue the company in Zimbabwe, or Great Britain, or China," or where ever he read the allegedly libelous remarks.[53] The Court observed that Gutnick lived in Victoria and this was where the alleged harm occurred. It also noted that Dow Jones profited from the sale of *Barrons Online* to Australians.

Nonetheless, the Australian court's ruling was unsettling for many in the publishing world. According to one lawyer for the publishing industry, "The problem is that rogue governments like Zimbabwe will pass laws that will effectively shut down the Internet."[54]

Question

Do you agree with the ruling in this case? Why or why not? Are Dow Jones' fears unfounded or do they have some merit?

References

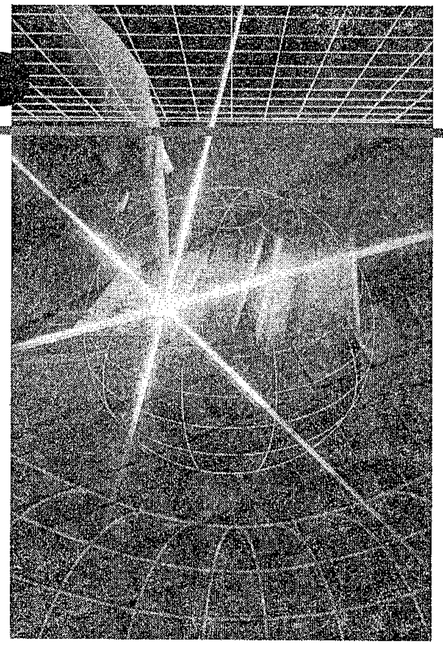
1. Christopher Wren, "Drug Culture Flourishes on Internet," *The New York Times*, June 20, 1997, p. A19.
2. Jonathan Katz, "Birth of Digital Nation," *Wired* (April 1997):186.
3. Geanna Rosenberg, "Trying to Resolve Jurisdictional Rules on the Internet," *The New York Times*, April 14, 1997, p. D1.
4. Stewart Brand, "Interview with Paul Baran (Founding Father)," *Wired* (March 2001):145–153.
5. Janet Abbate, *Inventing the Internet* (Cambridge: MIT Press, 1999), p. 11.

6. Ibid., p. 111.
7. World Internet Usage Statistics. http://www.internetworldstats.com Updated, July 2005.
8. Ibid.
9. "The Accidental Superhighway: A Survey of the Internet," *The Economist*, July 1, 1995, p. 6.
10. Katie Hafner, "The Internet's Invisible Hand," *The New York Times*, January 10, 2002, p. E1.
11. This discussion is adapted from Rus Shuler, "How Does the Internet Work," available: http://rus1.home.mindspring.com/whitepapers.
12. Manuel Castells, *The Internet Galaxy* (New York: Oxford University Press, 2001), p. 28.
13. Jonathan Zittrain, "The Rise and Fall of Sysopdom," *Harvard Journal of Law and Technology* 10 (1997):495.
14. Hafner, p. E5.
15. Pamela Samuelson and Hal Varian, "The 'New Economy' and Information Technology Policy" (Working Paper, University of California, Berkeley, July 18, 2001).
16. Internet Systems Consortium, ISC Internet Domain Survey, July, 2005; available at: http://www.isc.org.
17. Otis Port, "The Next Web," *Business Week*, March 4, 2002, p. 97.
18. Timothy Mullaney, "E-Biz Strikes Again," *Business Week*, May 10, 2004, p. 80.
19. Michael Mandel and Robert Hof "Rethinking the Internet," *Business Week*, March 26, 2001, pp. 116–141.
20. "Survey of Electronic Commerce," *The Economist*, May 10, 1997, p. 6.
21. Lynda Applegate, "E-Business Models," in *Information Technology and the Future of the Enterprise*, ed. G. Dickson and G. DeSanctis (Upper Saddle River, NJ: Prentice-Hall, 2001), pp. 49–101.
22. N. Carr "Redesigning Business," *Harvard Business Review*, November–December, 1999, p. 19.
23. Applegate, p. 64.
24. Bruce Einhorn, "China's Web Masters," *Business Week*, August 2, 1999, p. 49.
25. Ronald Coase, "The Problem of Social Cost," *The Journal of Law and Economics* 3 (1960): 1–44.
26. Ronald Coase, "The Theory of Public Utility Pricing and its Application," *Bell Journal of Economics and Management Science* 1(1970):113–128.
27. Joel Reidenberg, "Privacy Protection and the Interdependence of Law, Technology, and Self-Regulation," *Cahiers du C.R.I.D.*, 2002.
28. Larry Lessig, *Code and Other Values in Cyberspace* (New York: Basic Books, 1999), p. 60.
29. A.C. Pigou, *The Economics of Welfare* (London: Macmillan, Ltd., 1962), p. 195.
30. Howard Rheingold, *The Virtual Community: Homesteading on the Electronic Frontier* (Reading, MA: Addison-Wesley, 1993), p. 7.
31. Nicholas Negroponte, *Being Digital* (New York: Knopf, 1995), p. 12.
32. *America Banana Co. v. United Fruit Co.* 213 U.S. 347, 357 (1909).
33. Jonathan Zittrain, "Be Careful What You Ask For: Reconciling a Global Internet and Local Law," in *Who Rules the Net?* ed. A Theirer and W. Crews, (Washington, D.C.: Cato Institute, 2003).
34. Pamela Samuelson, "Five Challenges for Regulating the Global Information Society" (paper presented at Communications Regulation in the Global Information Society Conference, University of Warwick, June, 1999).
35. Tony Walker, "China's Wave of Internet Surfers," *The Financial Times*, June 24, 1995.
36. P. Thibodeau, "Europe's Privacy Laws May Become Global Standard," *Computerworld*, March 12, 2001, p. 77.
37. Samuleson, "Five Challenges," p. 7.
38. David G. Post, "Of Horses, Black Holes, and Decentralized Law-Making in Cyberspace" (paper delivered at Private Censorship/Perfect Choice Conference at Yale Law School, April 9–11, 1999).
39. Seth Finkelstein, "Internet Blocking Programs and Privatized Censorship," *The Ethical Spectacle*, August, 1998, http://www.spectacle.org/896/finkel.html

40. Larry Lessig, "Tyranny in the Infrastructure," *Wired*, July 1997, p. 96.
41. Michel Foucault, *Power and Knowledge: Selected Interviews and Other Writings*, trans. C. Gordon (New York: Random House, 1980), p. 111.
42. "Regulating the Internet—The Consensus Machine," *The Economist*, June 10, 2000, p. 73.
43. Michel Foucault, *The History of Sexuality, Volume I*, trans. R. Hurley (New York: Vintage Books, 1978), p. 95.
44. Larry Lessig, "The Laws of Cyberspace," p. 11.
45. Andrew Shapiro, *The Control Revolution*, (New York: Century Foundation Books, 1999), p. 73.
46. Quintin Hardy, "Yahoo: The Killer Ad Machine," *Forbes*, December 11, 2000, p. 174.
47. *Yahoo, Inc. v. LICRA*, U.S. District Court for N. Cal, Case #C-00-21275 JF, 2001.
48. Ibid.
49. John Tagliabue, "French Uphold Ruling Against Yahoo on Nazi Sites," *The New York Times*, November 21, 2000, p. C8.
50. Carl Kaplan, "Ruling on Nazi Memorabilia Sparks Legal Debate," *CyberLaw Journal*, November 24, 2000.
51. Quoted in Felicity Barringer, "Internet Makes Dow Jones Open to Suit in Australia," *The New York Times*, December 11, 2002, p. C6.
52. *Dow Jones & Company, Inc. v. Gutnick* (2002) 194 A.L.R. 433, H.C.A. 56.
53. Jonathan Zittrain, "Be Careful What You Ask For," p. 197.
54. Barringer, p. C6.

CHAPTER THREE

Free Speech and Content Controls in Cyberspace



The Internet has clearly expanded the potential for individuals to exercise their First Amendment right to freedom of expression. The Net's technology bestows on its users a vast expressive power. They can, for instance, disseminate their own blogs, publish electronic newsletters, or establish a home page on the Web. According to Michael Godwin, the Net "puts the full power of 'freedom of the press' into each individual's hands."[1] Or as the Supreme Court eloquently wrote in its *Reno v. ACLU* decision, the Internet enables an ordinary citizen to become "a pamphleteer, . . . a town crier with a voice that resonates farther than it could from any soapbox."[2]

But some forms of expression, like pornography or venomous hate speech, are offensive. They provoke a great sense of unease along with calls for limited content controls. Many resist this notion, however, insisting that the state should not interfere with unfettered access to online content.

As a result, the issue of free speech and content controls in cyberspace has emerged as arguably the most contentious moral problem of the nascent Information Age. Human rights such as free speech have taken a place of special prominence in the past century. In some respects these basic rights now collide with the state's inclination to rein in this revolutionary power enjoyed by Internet users. Whereas the U.S. has sought to suppress online pornography, the target of some European countries such as France and Germany has been mean-spirited hate speech.

In addition, speech is at the root of most other major ethical and public policy problems in cyberspace including privacy, intellectual property, and security. These three issues are discussed in Chapters 4, 5, and 6, where the free speech theme continues to have considerable saliency, but it is instructive at this point to consider how these issues are interconnected.

CyberEthics

Morality and Law in Cyberspace

Third Edition

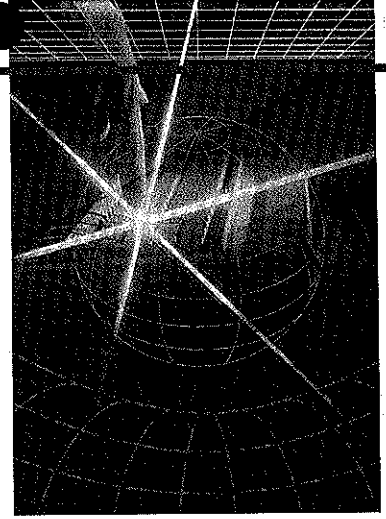
(2006)

Richard A. Spinello
Carroll School of Management
Boston College

JONES AND BARTLETT PUBLISHERS
Sudbury, Massachusetts
BOSTON TORONTO LONDON SINGAPORE

CHAPTER ONE

The Internet and Ethical Values



The end [of ethics] is action, not knowledge.
—Aristotle[1]

More than three decades have passed since the first communications were transmitted over a fledgling global network, which would later be called the *Internet*. At the time, few would have predicted the Internet's explosive growth and persistent encroachment on our personal and professional lives. This radically decentralized network has been described in lofty terms as empowering and democratizing. It has lived up to this ideal by creating opportunity for many new voices with extraordinary reach. Although the claim that the Internet will revolutionize communications may be hyperbole, there is no doubt that the Internet has the potential to magnify the power of the individual and fortify democratic processes.

Many governments, however, are clearly threatened by some of this decentralized power and they have sought to impose some centralized controls on this anarchic network. The United States has attempted to regulate speech through the ill-fated Communications Decency Act and to restrict the use of encryption technology through its key recovery scheme. More draconian regulations have been imposed by countries like Iran and Saudi Arabia. The Net and its stakeholders have steadfastly resisted the imposition of such controls, and this has led to many of the tensions and controversies we consider throughout this book.

Although the control of technology through law and regulation has often been a futile effort, "correcting" technology with other technology has been

more effective. The regime of law has had a hard time suppressing the dissemination of pornography on the Internet, but blocking software systems that filter out indecent material have been much more successful. This reflects the net's paradoxical nature—it empowers individuals and allows them to exercise their rights such as free speech more vigorously, but it also makes possible effective technical controls that can undermine those rights.

Although the primary axis of discussion in this book is the ethical issues that surface on the Internet, we must devote attention to these related matters of cyber-governance and public policy. Thus, we explore in some detail the tensions between the radical empowerment that the Net allows and the impulse to tame this technology through laws and other mechanisms.

Because this is a book about ethics, about *acting* well in this new realm of cyberspace, we begin by reviewing some basic concepts that will enrich our moral assessment of these issues. Hence, in this introductory chapter our purpose is to provide a concise overview of the traditional ethical frameworks that can guide our analysis of the moral dilemmas and social problems that arise in cyberspace.

More important, we also elaborate here on the two underlying assumptions of this work: (a) the *directive* and architectonic role of moral ideals and principles in determining responsible behavior in cyberspace and (b) the capacity of free and responsible human beings to exercise some control over the forces of technology (technological realism). Let us begin with the initial premise concerning the proper role of cyberethics.

Cyberethics and the "Law of the Horse"

An ethical norm such as the imperative to be truthful is just one example of a constraint on our behavior. In the real world, there are other constraints including the laws of civil society or even the social pressures of the communities in which we live and work. There are many forces at work limiting our behavior, but where does ethics fit in?

This same question can be posed about cyberspace and to help us reflect on this question we turn to the framework of Larry Lessig. In his highly influential book, *Code and Other Laws of Cyberspace*, Lessig first describes the four constraints that regulate our behavior in real space: law, norms, the market, and code.

Laws, according to Lessig, are rules imposed by the government which are enforced through *ex post* sanctions. There is, for example, the complicated IRS tax code, a set of laws that dictates how much taxes we owe the Federal government. If we break these laws, we can be subjected to fines or other penalties levied by the government. Thanks to law's coercive pedagogy, those who get caught violating tax laws are usually quick to reform.

Social norms, on the other hand, are expressions of the community. Most communities have a well-defined sense of normalcy, which is reflected in their norms or standards of behavior. Cigar smokers are not usually welcome at most community functions. There may be no laws against cigar smoking in a particular setting, but those who try to smoke cigars will most likely be stigmatized and ostracized by others. When we deviate from these norms, we are behaving in a way that is socially "abnormal."

The third regulative force is the market. The market regulates through the price it sets for goods and services or for labor. Unlike norms and laws, market forces are not an expression of a community and they are imposed immediately (not in *ex post* fashion). Unless you hand over $2 at the local Starbucks you cannot walk away with a cup of their coffee.

The final modality of regulation is known as architecture. The world consists of many physical constraints on our behavior—some of these are natural (such as the Rocky Mountains) whereas others are human constructs (such as buildings and bridges). A room without windows imposes certain constraints because no one can see outside. Once again "enforcement" is not *ex post* but at the same time the constraint is imposed. Moreover, this architectural constraint is "self-enforcing"—it does not require the intermediation of an agent who makes an arrest or who chastises a member of the community. According to Lessig, "the constraints of architecture are self-executing in a way that the constraints of law, norms, and the market are not."[2]

In cyberspace we are subject to the same four constraints. Laws, such as the ones that provide copyright and patent protection, regulate behavior by proscribing certain activities and by imposing *ex post* sanctions for violators. It may be commonplace to download and upload copyrighted digital music, but this activity breaks the law. There is a lively debate about whether cyberspace requires a unique set of laws or whether the laws that apply to real space will apply here as well with some adjustments and fine tuning. Judge Frank Easterbrook has said that just as there is no need for a "law of the horse," there is no need for a "law of cyberspace."[3]

Markets regulate behavior in various ways—advertisers gravitate to more popular Web sites, which enables those sites to enhance services; the pricing policies of the Internet Service Providers determine access to the Internet; and so forth. It should be noted that the constraints of the market are often different in cyberspace than they are in real space. For instance, pornography is much easier and less expensive to distribute in cyberspace than in real space, and this increases its available supply.

The counterpart of architectural constraint in the physical world is software "code," that is, programs and protocols that make up the Internet. They too constrain and control our activities. These programs are often referred to as the "architectures of cyberspace." Code, for example, limits access to certain Web sites by demanding a username and password. Cookie

technology enables e-commerce, but compromises the consumer's privacy. Sophisticated software is deployed to filter out unsolicited commercial e-mail (or spam). In the long run, code may be more effective than law in containing spam, which rankles many users.

Finally, there are norms that regulate cyberspace behavior, including Internet etiquette and social customs. For example, flaming is considered "bad form" on the Internet and those who do it will most likely be disciplined by other members of the Internet community. Those who misrepresent themselves in a chat room also violate those norms and they too will be reproached if their true identity is revealed. Just as in real space, in cyberspace communities rely on shame and social stigma to enforce cultural norms.

But what role does ethics play in this neat regulatory framework? Lessig apparently includes ethical standards in the broad category he calls "norms," but in our view cultural norms should be segregated from ethical ideals and principles. Cultural norms are nothing more than variable social action guides, completely relative and dependent upon a given social or cultural environment. Their validity depends to some extent on custom, prevalent attitudes, public opinion, and a myriad of other factors. Just as customs differ from country to country the social customs of cyberspace could be quite different from the customs found in real space. Also, these customs will likely undergo some transformation over time as the Internet continues to evolve.

The fundamental principles of ethics, however, are metanorms; they have universal validity. They remain the same whether we are doing business in Venezuela or interacting in cyberspace. Like cultural norms they are prescriptive, but unlike these norms, they have lasting and durable value because they transcend space and time. Ethics is about (or should be about) intrinsic human goods and the moral choices that realize those goods. Hence the continuity of ethical principles despite the diversity of cultures.

Our assumption that ethics and customs (or cultural norms) must be kept distinct defies the popular notion of ethical relativism, which often equates the two. A full refutation of that viewpoint is beyond the scope of our discussion here. But consider this reflection of the contemporary philosopher Phillippa Foot:

> Granted that it may be wrong to assume identity of aim between people of different cultures; nevertheless there is a great deal all men have in common. All need affection, the cooperation of others, a place in community, and help in trouble. It isn't true to suppose that human beings can flourish without these things—being isolated, despised or embattled, or without courage or hope. We are not, therefore, simply expressing values that we happen to have if we think of some moral systems as good moral systems and others as bad.[4]

None of this by any means invalidates Lessig's framework. His chief insight is that "code and market and norms and law together regulate in cyberspace as architecture and market and norms and law regulate in real space."[5] Also, according to Lessig, "Laws affect the pace of technological change, but the structures of software can do even more to curtail freedom. In the long run the shackles built by programmers could well constrain us more."[6] This notion that private code can be a more potent constraining force than public law has significant implications. The use of code as a surrogate for law may mean that certain public goods or moral values once protected by law will now be ignored or compromised by those who develop or utilize this code. Moreover, there is a danger that government itself will regulate the architectures of cyberspace to make it more controllable. It could, for instance, mandate the traceability of all Internet transactions, and thereby increase its capacity for surveillance or oversight of all interactions in cyberspace. In the hands of the private or public sector the architectures of cyberspace can have extraordinary regulatory power.

Thus, Lessig's model is quite instructive and we rely on it extensively in the pages to come. However, I would argue that the model would be more useful for our purposes if greater attention were given to the role of fixed ethical values as a constraining force. But how do these values fit with the other regulatory forces?

Before we can answer this question we must say something about the nature of those values. The notion that there are transcendent moral values grounded in our common human nature has a deep tradition in the history of philosophy. It is intuitively obvious that there are basic human goods that contribute to human well-being or human flourishing. Although there are several different versions of what these goods might be, they do not necessarily contradict each other. Some versions of the human good are "thin," while others are "thick." James Moor's list of core human goods includes life, happiness, and autonomy. According to Moor, *happiness* is "pleasure and the absence of pain," and *autonomy* includes those goods that we need to complete our projects (ability, security, knowledge, freedom, opportunity, reason). Individuals may rank these values differently but all human beings attribute value to these goods or "they would not survive very long."[7]

Oxford philosopher, John Finnis, offers a thicker version of the human good. He argues persuasively for the following list of intrinsic goods: life, knowledge, play (and skillful work), aesthetic experience, sociability, religion, and practical reasonableness (which includes autonomy). According to Finnis, participation in these goods allows us to achieve genuine human flourishing. They are opportunities for realizing our full potential as human beings, for being all that we can be. Hence the master principle of morality: one's choices should always be open to *integral human fulfillment*, the fulfillment of all persons and communities. None of our projects or objectives provides sufficient reason for setting aside or ignoring that responsibility.

For both Moor and Finnis, then, the ulitmate source of moral normativity is these intelligible, authentically *human* goods, which adequately explain the reasons for our choices and actions, and overcome the presumption of subjectivism. Morality can begin to claim objectivity because this collection of basic human goods is not subjective, that is, subject to cultural differences or individual whims.

The ultimate good, the human flourishing of ourselves and of others, should function as a prescriptive guidepost of enduring value, serving as a basis for crafting laws, developing social institutions, or regulating the Internet. Because this moral ideal is rather lofty, its application to policy making can be difficult. As a result, we are also guided by intermediate ethical principles, such as the Golden Rule, which states that "whatever you wish that men would do to you, do so to them" (Matthew 7:12). Similarly one could be guided by Kant's second version of the categorical imperative: "Act so that you treat humanity always as an end and never as a means."[8] From these principles one can derive more *specific core moral values* about murder, theft, or lying. These principles can function as more practical guidelines for moral decision making and enable us to pursue the basic human goods in a way that respects our fellow humanity. According to Finnis, our fundamental responsibility is to respect each of these human goods "in each person whose well-being we choose to affect."[9]

We contend, therefore, that these intelligible goods, intrinsic to human persons and essential for human flourishing, along with basic moral principles (such as the Golden Rule) should play an architectonic or *directive role* in the regulation of cyberspace. They should guide and direct the ways in which code, laws, the market, and social norms exercise their regulatory power. The value of human flourishing is the ultimate constraint on our behavior in real space and in cyberspace. Accordingly, we have enhanced Lessig's model as depicted in Figure 1-1.

To illustrate our point about the role of these supreme ethical values and how they can be translated into the actual world of our experience, let us consider the regulatory impact of code. There are responsible and irresponsible ways of developing code that constrains behavior. As we will see in Chapter 3, blocking software systems have become a common way of protecting young children from pornography. Those who write this code have developed proprietary blocking criteria and as a rule they do not reveal these criteria or the specific sites that are blocked. In some cases sex education or health-related sites are filtered out along with the pornography. If this is done inadvertently, the software should be fixed; if it is done deliberately, parents should be informed that the scope of the filtering mechanism is broader than just pornography. One could certainly make the case that parents should know what the blocking criteria are in order to make an informed judgement about the suitability of this software. Failure to reveal this information is tantamount to disrespecting parental autonomy.

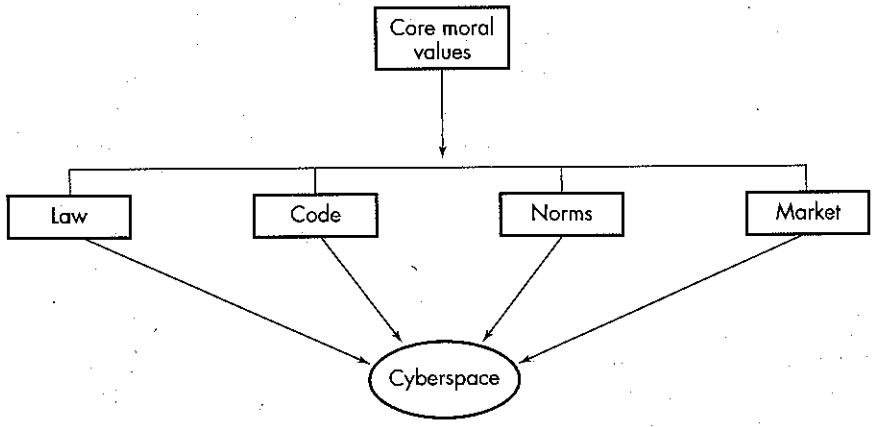


FIGURE 1-1 Constraints on cyberspace activities. (Adopted from Professor Lessig's Framework.)

As a result, one could argue that when the criteria are obscured for some ulterior agenda, the code is not being deployed in a responsible manner that is consistent with the core good of autonomy.

I am not suggesting that this is a clear-cut matter or that moral principles can provide all the answers to proper cyberspace regulations. And I am not making a judgment about whether law or code is the more effective constraint for cyberporn. I am simply claiming that those who write these programs or formulate laws to regulate cyberspace should rely on ethics as a guide. Code writers must be responsible and prudent enough to incorporate into the new architectures of cyberspace structures that preserve basic moral values such as autonomy and privacy. Further, government regulations of cyberspace must not yield to temptation to impose excessive controls. Regulators too must be guided by high moral standards and respect for basic human values such as freedom and privacy. The code itself is a powerful sovereign force and unless it is developed and regulated appropriately it will surely threaten the preservation of those values.

The role of morality should now be quite evident: it must be the ultimate regulator of cyberspace that sets the boundaries for activities and policies. It should direct and harmonize the forces of law, code, the market, and social norms so that interactions and dealings there will be measured, fair, and just.

Iron Cage or Gateway to Utopia?

Although most of us agree that some constraints will need to be imposed on the technologies of networking and computing that have come to

pervade the home and workplace, there is legitimate skepticism about anyone's ability to control the ultimate evolution and effects of these technologies. Are our attempts to regulate cyberspace merely a chimera? Are we too trammeled by the forces of technology, or are we still capable of exercising sovereignty over the code that constitutes the inner workings of the Internet?

As we observed in the preface, some philosophers have long regarded technology as a dark and oppressive force that menaces our individuality and authenticity. These technology determinists see technology as an independent and dehumanizing force beyond humanity's capacity to control it. The French philosopher Jacques Ellul presents a disturbing vision of technology in his seminal work, *The Technological Society*. His central argument is that *technique* has become a dominant and untranscendable human value. He defines technique as "*the totality of methods rationally arrived at and having absolute efficiency* (for a given stage of development) in *every* field of human activity."[10] According to Ellul, technique is beyond our control; it has become autonomous and "fashioned an omnivorous world which obeys its own laws and which has renounced all tradition."[11] For Ellul, modern technology has irreversibly shaped the way we live, work, and interact in this world.

Ellul was not alone in advancing such a pessimistic outlook on technology. Max Weber coined the term *iron cage* to connote how technology locks us in to certain ways of being or patterns of behavior. And Martin Heidegger saw technology not merely as a tool that we can manipulate but as a way of "being in the world" that deeply affects how we relate to that world. But is it really so that technology forces us into this "iron cage," and into a more fragmented, narrow-minded society dominated by a crude instrumental rationality?

In contrast to the bleak outlook of Ellul and Heidegger we find technology neutralists who argue that technology is a neutral force, completely dependent on human aims and objectives. According to this viewpoint, technologies are free of bias and do not promote one type of behavior over another. Technology *is* only a tool, and it does not compromise our human freedom or determine our destiny in any appreciable way—it is up to us whether this powerful force is used for good or ill purposes.

Some go even further and embrace a sort of "technological utopianism" that regards certain technologies as making possible an ideal world with improved lifestyles and workplaces. This optimistic philosophy assumes that humanity can eradicate many of technology's adverse effects and manipulate this tool effectively to improve the human condition.

The philosophy of technological neutralism (or, for that matter, utopianism) seems problematic for several reasons. Technology does condition our choices with certain "givens" that are virtually impossible to fully overcome. Langdon Winner describes this as a process of reverse adaptation or "the adjustment of human ends to match the character of the available means."[12]

However, in our view, it is also an exaggeration to claim that computer and network technology locks us into a virtual but inescapable iron cage. The middle ground between these extreme positions is *technological realism*, which holds that "although technology has a force of its own, it is not independent of political and social forces."[13] Technological realism acknowledges that technology has reconfigured our political and social reality and that it does influence human behavior in particular ways. To some extent, this notion is echoed in Lessig's work. He argues that we fail to see sometimes how code is an instrument of social and political control. Code is not neutral. Most often, embedded within code are certain value decisions that define the set of options for policy problems.

Nonetheless, although technology determines to some degree how we live and work, we still have the capacity to redirect or subdue it when necessary. In effect, we can still shape and dictate how certain technological innovations will be deployed and restrained, particularly when there is a conflict with the common good or core human goods. Our human freedom is undoubtedly attenuated by technology's might and its atomizing tendencies, but it is not completely effaced. We can still choose to implement systems and develop code in ways that protect fundamental human rights such as autonomy or privacy. We can be liberated from the thralldom of privacy-invading code by developing new code that enhances privacy.

In this postmodern age such a position may also seem simplistic and outdated. Although social psychologists talk about the "social construction of the self" and French psychoanalysts like Jacques Lacan refer to the unconscious as the controlling center of the self, we still presume that beneath it all is a conscious, thinking self or moral agent responsible for its actions *and* responsible for making choices about the deployment of various technologies.

Beyond any doubt, technology and its counterpart instrumental rationality are dominant forces in this society that exert enormous pressures on us to make choices and behave in certain ways. But as Charles Taylor points out, one can find throughout history pockets of concerted opposition to oppressive technologies. Further, the chances for such successful resistance are greatly enhanced when there is some common understanding about a particular threat or imperilment, such as the threat to our ecology that occupied us during the 1970s. Perhaps the same common consciousness will emerge about the threat to personal privacy, and this will provide yet another impetus for human choice to trump the dominating forces of information technology. Although we should not be overly optimistic about our freedom and our capacity for resisting infatuation with new technology, we must recognize that we still have *some* degree of freedom in this world. Thus, we agree with Taylor's assessment: "We are not, indeed, locked in. But there is a slope, an incline in things that is all too easy to slide down."[14]

How then do we avoid this fatal slide? This brings us to our next topic of discussion in this introduction—the importance of cultivating and sustaining a moral point of view as one deliberates about how to constrain behavior on the Internet through market forces, code, norms, or law.

Ethical Values and the Digital Frontier

We avoid this slide and its accompanying perils only if we conscientiously adopt the moral point of view as we evaluate technological capabilities and make decisions about the ground rules of the digital frontier. How can we characterize this moral point of view? According to Kenneth Goodpaster, it can be seen "as a mental and emotional standpoint from which all persons have a special dignity or worth, from which the Golden Rule derives its worth, and from which words like *ought* and *duty* derive their meaning."[15] This is quite consistent with our earlier claim that the fundamental moral imperative is the promotion of human flourishing both in ourselves and in others.

Several distinct types of ethical reasoning have been associated with the moral point of view, and they provide us with the basic principles that serve as a moral yardstick or "compass" that can assist us in making normative judgements. Our discussion here is concise; for the interested reader it can certainly be amplified by many other books on ethical theory or on applied ethics.[16] We consider several models of ethical reasoning based on moral frameworks emphasizing the maximization of social utility, natural rights, contract rights, and moral duties.

The fact that there are several different theories embodying the moral point of view does not contradict our assumption regarding the core human goods that form the basis of a unifying moral framework. All of these theories recognize such goods in one form or another. Kant embraces the principle that we must respect humanity in all our choices and actions, although he might define *humanity* differently from Finnis. And rights-based theories discuss core human goods in terms of protection of human rights such as the rights to life, liberty, and the pursuit of happiness. The utilitarian approach emphasizes happiness and, although it may have a hard time standing on its own, it can be complemented by other theories to form a more comprehensive framework.

All of these theories are worth our careful consideration. Each represents a valuable perspective from which complex moral issues can be assessed and reflected upon. They help us to engage in the critical moral analysis necessitated by the thorny dilemmas that are beginning to surface all over the Internet.

Before we discuss these theories, it is worth pointing out that modern ethical frameworks fall under two broad categories: teleological or deontological. *Teleological* derives from the Greek *telos*, which means *goal* or *end*. These theories argue that the rightness or wrongness of an action depends on whether or not they bring about the end in question (such as happiness). *Deontological* theories, on the other hand, consider actions to be intrinsically right or wrong—their rightness or wrongness does not depend in any way on the consequences which they effect. These frameworks emphasize duty and obligation (*deon* is the Greek word for *duty*).

Utilitarianism

Utilitarianism is a teleological theory, and it is by far the most popular version of consequentialism. Classic utilitarianism was developed by two British philosophers, Jeremy Bentham 1748–1832 and John Stuart Mill 1806–1873. According to this theory, the right course of action is to promote the general good. This general good can also be described in terms of "utility," and this principle of utility is the foundation of morality and the ultimate criterion of right and wrong. *Utility* refers to the net benefits (or good) created by an action. According to Frankena, utilitarianism is the view that "the sole ultimate standard of right, wrong and obligation is the *principle of utility* or *beneficence*, which says quite strictly that the moral end to be sought in all that we do is *the greatest possible balance of good over evil* (or the least possible balance of evil over good)."[17] Thus, an action or policy is right if it produces the greatest net benefits or the lowest net costs (assuming that all of the alternatives impose some net cost).

It should be emphasized that utilitarianism is quite different from ethical egoism. An action is right not if it produces utility for the person performing that action but for all parties affected by the action. With this in mind we might reformulate the moral principle of utilitarianism as follows: persons ought to act in a way that promotes the maximum net expectable utility, that is, the greatest net benefits or the lowest net costs, for the broadest community affected by their actions.

On a practical level, utilitarianism requires us to make moral decisions by means of a rational, objective cost/benefit analysis. In most ethical dilemmas there are several possible alternatives or courses of action. Once one has sorted out the most viable and sensible alternatives, each one is evaluated in terms of its costs and benefits (both direct and indirect). Based on this analysis, one chooses the alternative that produces the greatest net expectable utility, that is, the one with the greatest net benefits (or the lowest net costs) for the widest community affected by that alternative.

A concrete example illustrates how cost/benefit analysis might work. Let us assume that a corporation has to make a policy decision about random inspection of employee e-mail. This might be done as a routine part of a performance review as a means of checking to make sure that workers are using e-mail only for work-related purposes and are not involved in any

untoward activities. This practice is perfectly legal, but some managers wonder if it's really the right thing to do; it seems to violate the privacy rights of employees. Rightness in the utilitarian ethical model is determined by consequences that become transparent in a cost-benefit analysis. In this case, the managers might face three options: e-mail messages are not inspected on a routine basis and are kept confidential (unless some sort of malfeasance or criminal activity is suspected); e-mail messages are inspected regularly by managers, but employees are informed of this policy and reminded of it every time they log in to the e-mail system, so that there is no expectation of privacy; or e-mail is regularly but surreptitiously perused by managers with employees uninformed of the company policy. Which of these alternatives promotes the general good, that is, produces the greatest net expectable utility?

Table 1-1 provides an idea of how this analysis might work out. It becomes clear from this exercise that it's difficult to objectively calculate the diffuse consequences of our actions or policies and to weight them appropriately. And herein lies a major obstacle in using this approach. Nonetheless, there is value in performing this type of analysis; it induces us to consider the broad consequences of our actions and to take into account the human as well as the economic costs of implementing various technologies.

Although this theory does have certain strengths it is also seriously flawed in some ways. Depending on the context, utilitarianism could be used to justify the infliction of pain on a small number of individuals for the sake of the happiness or benefits of the majority. There are no intrinsically unjust or immoral acts for the utilitarian, and this poses a problem. What happens when human rights conflict with utility? Can those rights be suppressed on occasion for the general good? There is nothing in utilitarianism to prevent this from happening, as long as a cogent and objective case is made that the benefits of doing so exceeds the costs. The primary problem then is that this theory lacks the proper sensitivity to the vital ideals of justice and human rights.

Contract Rights (Contractarianism)

Another mode of reasoning that exemplifies the moral point of view is rights-based analysis, which is sometimes called *contractarianism*. Unlike utilitarianism, contractarianism is a deontologic theory. It looks at moral issues from the viewpoint of the human rights that may be at stake. A *right* is an entitlement or a claim to something. For instance, thanks to the Fourth Amendment, American citizens are entitled to protection from unwarranted search and seizures in the privacy of their homes. In contrast to the utilitarian view, the consequences of an action are morally irrelevant for those who support contractarianism. Rights are unequivocally enjoyed by

	Costs	Benefits
1. Confidential e-mail	Lack of control over employees; difficult to prevent misuses of e-mail; e-mail could be used for various personal reasons without company knowledge.	Maintains morale and an environment of trust and respect for workers; protects personal privacy rights.
2. Inspect e-mail with employees informed of policy	Violates privacy rights; diminishes trust and impairs morale; workers less likely to use e-mail if communications are not confidential—instead they will rely on less efficient modes of communication.	Prevent misuse along with inappropriate comments about superiors and fellow workers via e-mail; workers know the risks of using e-mail; they are less likely to use e-mail for personal purposes.
3. Inspect surreptitiously	Same as option 2, but even more loss of trust and morale if company policy is uncovered.	Better chance to catch employees doing something wrong such as transmitting trade secrets; perfectly legal.

TABLE 1-1 Illustrative Cost/Benefit Analysis

all citizens, and the rights of the minority cannot be suspended or abolished even if that abolition will maximize social welfare.

An important distinction needs to be made between positive and negative rights. Possession of a *negative right* implies that one is free from external interference in one's affairs. Examples of negative rights include the right to free speech, the right to property, and the right to privacy. Because all citizens have a right to privacy in their homes, the state cannot interfere in their affairs by tapping their phone calls unless it has demonstrated a strong probability that laws are being broken.

A *positive right*, on the other hand, implies a requirement that the holder of this right be provided with whatever one needs to pursue one's legitimate interests. The rights to medical care and education are examples of positive rights. In the United States the right to universal health care is rather dubious but the right to education is unequivocal. Therefore the state has a duty to educate children through the twelfth grade. If everyone had a "right" to Internet access, there would be a correlative duty on the part of the government to provide that access for those who could not afford it.

Rights can be philosophically grounded in several ways. Some traditional philosophers such as Locke and Rousseau and the contemporary social philosopher John Rawls claim that we have basic rights by virtue of an implicit social contract between the individual and civil society. Individuals agree to a contract outside of the organized civil society that stipulates

the fundamental principles of their association including their rights and duties. Rights are one side of a quid pro quo—we are guaranteed certain rights (e.g., life, liberty, and the pursuit of happiness) as long as we obey the laws and regulations of civil society. This contract is not real but hypothetical. According to Kelbley, "we are not discussing facts but an ideal which rational individuals can embrace as a standard to measure the moral nature of social institutions and efforts at reform."[18]

According to this perspective, moral reasoning should be governed by respect for these individual rights and by a philosophy of fairness. As Ken Goodpaster observes, "fairness is explained as a condition that prevails when all individuals are accorded equal respect as participants in social arrangements."[19] In short, then, this rights-based approach to ethics focuses on the need to respect an individual's legal, moral, and contractual rights as the basis of justice and fairness.

The problem with most rights-based theories is that they do not provide adequate criteria for resolving practical disputes when rights are in conflict. For example, those who send spam (unsolicited commercial e-mail) over the Internet claim that they are exercising their right to free speech, but many recipients argue that spam is intrusive, maybe even a form of trespass. Hence they claim that the transmission of spam is an invasion of their property rights. The real difficulty is how we adjudicate this conflict and determine which right takes priority. Rights-based theories are not always helpful in making this determination.

Moral Duty (Pluralism)

The next framework for consideration is not based on rights, but on duty. The moral philosophy of Immanuel Kant (1724–1804), which can be found in his short but difficult masterpiece on ethics, *Fundamental Principles of the Metaphysics of Morals,* is representative of this approach. It assumes that the moral point of view is best expressed by discerning and carrying out one's moral duty. This duty-based, deontological ethical framework is sometimes referred to as *pluralism.*

Kant believed that consequences of an action are morally irrelevant: "an action performed from duty does not have its moral worth in the purpose which is to be achieved through it but in the maxim by which it is determined."[20] According to Kant, actions only have moral worth when they are done for the sake of duty. But what is our duty and how is it derived? In Kant's systematic philosophy our moral duty is simple: to follow the moral law which, like the laws of science or physics, must be rational. Also, like all rational laws, the moral law must be universal, because, universality represents the common character of rationality and law. And this universal moral law is expressed as the categorical imperative: "I should never act except in such a way that I can also will that my maxim should become a universal law."[21] The imperative is "categorical" because it does not allow for any exceptions.

A *maxim* as referred to in Kant's categorical imperative is an implied general principle or rule underlying a particular action. If, for example, I usually break my promises, then I act according to the private maxim that promise breaking is morally acceptable when it is in my best interests to do so. But can one take this maxim and transform it into a universal moral law? As a universal law this particular maxim would be expressed as follows: "It is permissible for everyone to break promises when it is in their best interests to do so." Such a law, however, is invalid because it entails both a pragmatic and a logical contradiction. There is a pragmatic (or practical) contradiction because the maxim is self-defeating if it is universalized. According to Korsgaard, "your action would become ineffectual for the achievement of your purpose if everyone (tried to) use it for that purpose."[22] Consider this example. An individual borrows some money from a friend and he promises to pay her back. However, he has no intention of keeping that promise. But this objective, that is, getting some money from her without repaying it, cannot be achieved by making a false promise in a world where this maxim has been universalized. As Korsgaard puts it, "The efficacy of the false promise as a means of securing money depends on the fact that not everyone uses promises that way."[23]

Universal promise breaking also implies a logical contradiction (like a square circle); if everyone broke promises, the entire institution of promising would collapse; there would be no such thing as a "promise" because in such a climate anyone making a promise would lack credibility. A world of universalized promise breaking is inconceivable. Thus, in view of the contradictions involved in universalizing promise breaking, we have a perfect duty to keep all of our promises.

Kant strongly implies that *perfect duties,* that is, duties that we are always obliged to follow such as telling the truth or keeping a promise, entail both a logical and pragmatic contradiction. Violations of imperfect duties, however, are pragmatic contradictions. Korsgaard explains that "perfect duties of virtue arise because we must refrain from particular actions *against* humanity in our own person or that of another."[24] *Imperfect duties,* on the other hand, are duties to develop one's talents where the individual has the latitude to fulfill this duty using many different means.

Kant's categorical imperative is his ultimate ethical principle. It is the acid test of whether an action is right or wrong. According to Kant, then, any self-contradictory universalized maxims are morally forbidden. The categorical imperative functions as a guide, a "moral compass" that gives us a reliable way of determining a correct and consistent course of action. According to Norman Bowie, "the test of the categorical imperative becomes a principle of fair play—one of the essential features of fair play is that one should not make an exception of oneself."[25]

Also, from the categorical imperative we can derive other duties such as the duty to keep contracts, to tell the truth, to avoid injury to others, and so forth. Kant would maintain that each of these duties is also categorical, admitting of no exceptions, because the maxim underlying such an exception cannot be universalized.

How might we apply Kant's theory to the mundane ethical problems that arise in cyberspace? Consider the issue of intellectual property. As Korsgaard observes, "property is a practice,"[26] and this practice arguably makes sense for both physical property as well as intellectual property. But a maxim that permitted stealing of such property would be self-defeating. That maxim would say, "It's acceptable for me to steal the intellectual property validly owned by the creators or producers of that property." Such a universalized maxim, permitting everyone to take this intellectual property, is self-defeating precisely because it leads to the destruction of the entire "practice" of intellectual property protection. Because the maxim allowing an individual to freely appropriate another's intellectual property does not pass the universalization test, a moral agent is acting immorally when he or she engages in acts such as the unauthorized copying of a digital movie or music file.[27]

At the heart of Kant's ethical system is the notion that there are rational constraints on what we can do. We may want to engage in some action (such as downloading copyrighted files) but we are inconsistent and hence unethical unless we accept the implications of everyone doing the same thing. According to Kant, it is unethical to make arbitrary exceptions for ourselves. In the simplest terms, the categorical imperative suggests the following question: What if everybody did what you are doing?

Before concluding this discussion on Kant it is worth restating his second formulation of the categorical imperative: "Act in such a way that you treat humanity, whether in your own person or in the person of another, always at the same time as an end and never simply as a means."[28] For Kant as well as for other moralists (such as Finnis), the principle of humanity as an end in itself serves as a limiting condition of every person's freedom of action. We cannot exploit other human beings and treat them exclusively as a means to our ends or purposes. This could happen, for example, through actions that deceive one's fellow human beings or actions that force them to do things against their will. According to Korsgaard,

> According to [Kant's] Formula of Humanity, coercion and deception are the most fundamental forms of wrongdoing to others—the roots of all evil. Coercion and deception violate the conditions of possible assent, and all actions which depend for their nature and efficacy on their coercive or deceptive character are ones that others cannot assent to. . . . Physical coercion treats someone's person as a tool; lying treats someone's reason as a tool.[29]

If we follow this categorical imperative we will make sure that our projects and objectives do not supercede the worth of other human beings. This principle can also be summed up in the notion of *respect*. One way to express universal morality is in terms of the general principle of respect for other human beings, who deserve that respect because of their dignity as free and rational persons.

One of the problems with Kant's moral philosophy is its rigidity. There are no exceptions to the moral laws derived from the absolute categorical imperative. Hence lying is *always* wrong even though we can envision situations where telling a lie (e.g., to save a human life) is a reasonable and proper course of action. In cases like this there is a conflict of moral laws: the law to tell the truth and the law to save a life in jeopardy, and we have no alternative but to admit an exception to one of them. As A. C. Ewing points out,

> In cases where two laws conflict it is hard to see how we can rationally decide between them except by considering the goodness or badness of the consequences. However important it is to tell the truth and however evil to lie, there are surely cases where much greater evils can still be averted by a lie, and is lying wrong then?[30]

Ewing's argument that it is difficult to avoid an appeal to consequences when two laws conflict poses problems for Kant's moral philosophy, despite its powerful appeal.

An alternative duty-based philosophy proposed by William D. Ross (1877–1940), a contemporary English philosopher, attempts to obviate the difficulties posed by Kant's inflexibility. Ross argues in his book *The Right and the Good*[31] that we are obliged to follow several basic *prima facie* duties that each of us can intuit through simple reflection. These duties are *prima facie* in the sense that they are conditional and not absolute. This means that under normal circumstances we must follow a particular duty, but in those unusual situations where duties conflict with one another, one duty may be overridden by another duty that is judged to be superior, at least under these specific circumstances. According to Ross, moral rules or principles are not categorical as they are for Kant, so they can have exceptions. Thus, a moral principle can be sacrificed or overridden, but only for another moral principle, not just for arbitrary, selfish, or even utilitarian reasons.

According to Ross, the seven *prima facie* moral duties that are binding on all moral agents are the following:

1. One ought to keep promises and tell the truth (*fidelity*);
2. One ought to right the wrongs that one has inflicted on others (*reparation*);
3. One ought to distribute goods justly (*justice*)
4. One ought to improve the lot of others with respect to virtue, intelligence, and happiness (*beneficence*);

5. One ought to improve oneself with respect to virtue and intelligence (*self-improvement*);
6. One ought to exhibit gratitude when appropriate (*gratitude*); and
7. One ought to avoid injury to others (*noninjury*).

Ross makes little effort to provide any substantial rationalization or theoretical grounding of these duties. We might just say that they are common rules of morality, obvious to all rational humans because they have the general effect of reducing harm or evil to others.

The Achilles' heel of Ross's theory can be isolated by examining two specific problems: (1) his list of duties seems arbitrary because it is not metaphysically or even philosophically grounded, and (2) the list seems incomplete—where, for example, is the duty not to steal property from another? It may be included under the duty to avoid injury to others, but that is not altogether clear. Moreover, is it really true that all human beings (even those in different cultures) simply "intuit" these same principles? Finally, *The Right and the Good* provides little help for resolving situations where two prima facie duties do conflict. Ross offers few concrete criteria for determining when one obligation is more stringent and compelling than another.

Despite these shortcomings, however, Ross's framework, like the others we have considered, is not without some merit. A focus on one's moral duty (or even conflicting duties) in a particular situation is a worthy starting point for moral reasoning about some dilemma or quandry. Further, for many moral conundrums, a sincere and rational person can develop sound, objective reasons for determining which duty should take priority.

New Natural Law

The natural law tradition has been neglected in most books on business and computer ethics. Detractors claim that it's too "impractical" and too closely associated with the theistic philosophy of St. Thomas Aquinas. MacIntyre, however, makes the case that the natural law ethic is superior to the "theories of those imprisoned within modernity [that] can provide only ideological rationalizations [such as] modern consequentialism and modern contractarianism."[32]

The new natural law, developed by John Finnis and Germain Grisez, remains faithful to the broad lines of natural law theory found in the philosophy of Aquinas. But it also attempts to make some necessary modifications demanded by the complexity of contemporary moral problems. Like Aquinas, Finnis and Grisez claim that the starting point of moral reflection is the first practical principle: "Good should be done and evil avoided," where *good* means what is intelligibly worthwhile. For the most part, human beings behave rationally and pursue what is good for them, what perfects their nature and makes them better off. But what is the good? Recall Finnis' argument that there are seven basic human goods that are the key to human flourishing: life and health, knowledge of the truth, play (and some forms of work), aesthetic experience, sociability (including friendship and marriage), religion, and practical reasonableness. All of our choices ultimately point to one of these intelligible goods. For example, if someone asks Paul why he plays golf so much he could answer that he enjoys the game or that he likes the exercise. The first answer points to the basic human good of play and the second to the good of health.

Each one of us participates in these basic goods, though we may participate in some goods more than others, and we do so to achieve "fullness of life." Practical reasonableness, which includes the value of authenticity, shapes one's participation in the other basic goods. And one requirement of practical reasonableness is that it is unreasonable to choose directly against any basic value, "whether in oneself or in one's fellow human beings."[33]

But how do we get from these basic human goods to specific moral norms and human rights? Our practical reason grasps that each of these basic human goods is an aspect of human flourishing and that a good in which any person shares also fulfills other persons. Whenever one intentionally destroys, impedes, or damages one of these goods which should be allowed to be, there is moral evil. Thus, we can stipulate the First Principle of Morality: *keep one's choices open to integral human fulfillment*, the fulfillment of all persons and communities.[34]

This principle, however, is too general and so we also need intermediate principles to specify the primary moral principle. Grisez calls these *modes of responsibility*, which include the Golden Rule (or the universalizability principle), "for a will marked by egoism or partiality cannot be open to integral human fulfillment."[35] These modes also include the imperative to avoid acting out of hostility or vengeance and never to choose evil as the means to a good end. The good or the end of my actions does not justify the use of unjust means that damage a basic good. According to this principle, for example, one could not justify telling a lie that damages the truth to advance a friendship. In this case, one is exercising favoritism with regard to these goods, which are incommensurable and all deserving of the same respect.

Specific moral norms can be deduced from those basic human goods with the help of the intermediate principles such as the Golden Rule. For example, because human life is a basic human good, certain acts such as the taking of innocent life are forbidden as a matter of natural law. Finnis states this natural law (or absolute moral norm) as follows: "every act which is intended, whether as end or means, to kill an innocent human being and every act done by a private person which is intended to kill any human being" is prohibited.[36] This precludes necessary acts of self-defense. And

from the basic good of knowledge of the truth, we can we can deduce the moral imperative of veracity and "the right not to be positively lied to in any situation in which factual communication is reasonably expected."[37]

The new natural law provides a different vantage point from which to judge ethical conundrums in cyberspace. The value of this approach is its unwavering fidelity to the role of basic human goods such as life, health, and knowledge of the truth. It compels us to consider whether certain policies or actions are consistent with human flourishing, that is, with the realization of these basic human goods identified by Finnis and Grisez. It is difficult to argue, for instance, that deceptive spamming has any moral legitimacy; by undermining the truth in factual Internet communications, this form of spam deserves to be classified as morally reprehensible. The natural law framework allows us to appreciate why this is so wrong by focusing on its true negative impact.

Although Finnis and Grisez have tried to disengage the natural law framework from the metaphysics of Aquinas, critics claim that they do not succeed. According to Lisska, "One intuits the basic goods and it just happens that set of goods correspond to human well being. But what establishes the causal relationship?"[38] Nonetheless, according to Grisez, this theory attempts to combine the strengths of teleology and deontology. It grounds morality in human goods, "the goods of real people living in the world of experience," and it protects each person's dignity with intermediate principles and moral absolutes.[39]

Postscript on Moral Theory

As we have seen, none of these theories are without flaws or contradictions, but they do represent viable avenues for reasoning about moral issues, especially when those issues go beyond the level of moral common sense. They also have certain elements in common, particularly an orientation to "the other"—along with the need to consider the interests and perspectives of the affected parties in assessing alternative action plans, the other's moral and legal rights, and our duty to treat the other as an end and not as a means. And they all stand in opposition to the dangerous and myopic philosophy of ethical egoism, which is blind to the rights and aspirations of others.

Before concluding this material on ethical theory we can summarize how they can be applied to some of the moral quandaries that arise in the electronic frontier of cyberspace. Table 1-2 provides a concise framework for putting these four basic theories into action.

In some cases these four frameworks converge on the same solution to an ethical quandry. At other times, they suggest different solutions to the

Theory Type	Operative Questions
Consequentialism/utilitarianism	Which action or policy generates the best overall consequences or the greatest net expectable utility for all affected parties?
Duty-based morality	Can the maxim underlying the course of action being considered be universalized? Is the principle of fair play being violated? If there appears to be conflicting duties which is the stronger duty?
Rights-based morality	Which action or policy best protects the human and legal rights of the individuals involved?
New natural law	Does the proposed action or policy promote the basic requirements of human flourishing? Does it impede, damage, or destroy basic human goods?

TABLE 1-2 Summary of Ethical Frameworks

problem and one must decide which framework should "trump" or override the others. Should one respect the rights of some group or individual, even though following that alternative will be less beneficial to all affected parties than other alternatives? Resolving such questions requires careful and objective reasoning, but responsible behavior sometimes requires that this extra step be taken. To be sure, the Internet presents unique ethical challenges that could never have been envisioned by Aquinas, Kant, or Mill, but these frameworks still provide a general way of coming to terms with these tough questions.

Normative Principles

For those who find ethical theory too abstract, they can turn to an approach known as *principilism*. It is commonly used in biomedical ethics and has become popularized through the work of Beauchamp and Childress.[40] These moral principles are derived from and compatible with all of the moral theories articulated here. They constitute *prima facie* duties that are always in force but may conflict on occasion. The four principles proposed by Beauchamp and Childress are autonomy, nonmaleficence, beneficence, and justice. Those who advocate this approach also prescribe certain "prudential requirements" that determine when one prima facie principle should be given more weight than another. These include "being sure that there is a realistic prospect of achieving the moral objective one has chosen to honor; no alternative course of action is possible that would honor both conflicting obligations; and we minimize the effects of infringing on the prima facie duty."[41] A brief sketch of these four principles follows.

Autonomy

Kant and other philosophers have consistently argued that a defining element of personhood is one's capacity to be autonomous or self-determining. According to Gary Doppelt, "the Kantian conception of personhood ties the moral identity of persons to the supreme value of their rational capacities for normative self-determination."[42] All rational persons have two key moral powers or capacities: they possess the ability to develop and revise a rational plan to pursue their conception of the good life, and they possess the capacity to respect this same capacity of self-determination in others. Thus, autonomy is not only a necessary condition of moral responsibility, it is also through the exercise of autonomy that individuals shape their destiny according to their notion of the best sort of life worth living. When someone is deprived of their autonomy, their plans are interfered with and they are not treated with the respect they deserve. Of course, respect for autonomy must be balanced against other moral considerations and claims.

Nonmaleficence

The principle of nonmaleficence can best be summarized in the moral injunction: "above all, do no harm." According to this core principle, one ought to avoid unnecessary harm or injury to others whenever possible. This negative injunction against doing injury to others is sometimes called the "moral minimum." However, one may choose to develop a moral code of conduct, this injunction must be given a preeminent status. Most moral systems go well beyond this minimum requirement, as we have seen in the theories already discussed, but that does not detract from the central importance of this principle. According to Jon Gunneman and his co-authors,

> We know of no societies, from the literature of anthropology or comparative ethics, whose moral codes do not contain some injunction against harming others. The specific notion of *harm* or *social injury* may vary, as well as the mode of correction and restitution but the injunctions are present.[43]

Beneficence

This is a positive duty and has been formulated in may ways. In the simplest terms it means that we should act in such a way that we advance the welfare of other people when we are able to do so. In other words, we have a duty to help others. But what does this really mean? When am I duty bound to help another person or even an institution? It is obvious that we cannot help everyone or intervene in every situation when someone is in need. Hence some criteria are necessary for determining when such a moral obligation arises. In general, it can be argued that we have a duty to help others under the following conditions:

1. the need is serious or urgent;
2. we have knowledge or awareness of the situation; and
3. we have the capability to provide assistance ("ought assumes can" is the operative principle).

If, for instance, one is an Olympic swimmer and sees someone drowning at the beach one has an obligation to attempt a rescue of that person, especially if this is the only recourse and there is little risk to one's own life. This principle has some relevance when we evaluate society's questionable duty of beneficence to provide universal Internet service.

Justice

Although theories of justice have their differences most have in common adherence to this basic formal principle: "Similar cases ought to be treated in similar ways." Above all else justice requires fair treatment and impartiality. This is a formal procedural principle of justice and needs to be supplemented by the criteria for determining "similar" cases. This leads into theories of distributive justice, which attempt to formulate an underlying principle for how we should distribute the benefits and burdens of social life. Some theories emphasize equality, that is, all goods should be distributed equally. John Rawls, for example, adopts an egalitarian approach, though he does argue that an unequal distribution of goods is acceptable when it works for the advantage of everyone especially the least advantaged (the difference principle).[44] Other theories emphasize contribution and effort as formulated in this maxim: Benefits or resources should be distributed according to the contribution each individual makes to the furtherance of society's goals. And still another theory of justice that has typically been associated with socialism argues for justice based on need: "From each according to his ability, to each according to his needs."[45]

Our purpose here is not to defend one of these theories against the other, but to illustrate that moral judgements should be based in part on the formal principle of justice and take into account some standard regarding how the benefits and burdens should be fairly distributed within a group or society at large.

There is no reason that these formal moral principles cannot be applied to some of the controversial problems that we consider in this book. They are certainly general enough to have applicability in the field of computer and Internet ethics as well as bioethics. A person who makes choices and develops policies attentive to the core human goods and to these more practical principles which generally promote those goods would surely be acting with the care and prudence that is consistent with the moral point of view.

Discussion Questions

1. Do you agree with the philosophy of technological realism?
2. Explain the basic elements of Lessig's framework. What does he mean when he says that in cyberspace "the code is the law"?
3. Explain and critically analyze the essentials of Kant's moral theory.
4. In your estimation, which of the moral frameworks presented in this chapter has the most promise for dealing with the moral dilemmas that arise in cyberspace?

References

1. Aristotle, *Nicomachean Ethics*, I, 3: 1095a6.
2. Larry Lessig, *Code and Other Values of Cyberspace* (New York: Basic Books, 1999), p. 236.
3. See Frank Easterbrook, "Cyberspace and the Law of the Horse," *University of Chicago Law Forum* 207 (1996).
4. Phillippa Foot, "Moral Relativism" (1979 Lindley Lecture, Department of Philosophy, University of Kansas).
5. Larry Lessig, "The Laws of Cyberspace;" available at: http://cyberlaw.stanford.edu/lessig
6. Larry Lessig, "Tyranny in the infrastructure," *Wired* 5.07 (1997): 96.
7. Jim Moor, "Just Consequentialism and Computing," in *Readings in Cyberethics* ed. R. Spinello and H. Tavani (Sudbury, MA: Jones and Bartlett, 2001), p. 100.
8. Immanuel Kant, *Foundations of the Metaphysics of Morals* (Indianapolis: Bobbs Merrill, 1959), p. 33.
9. John Finnis, *Fundamentals of Ethics* (Washington, D.C.: Georgetown University Press, 1983), p. 125.
10. Jacques Ellul, *The Technological Society*, trans. John Wilkinson (New York: Vintage Books, 1964), p. xxv.
11. Ibid., p. 14.
12. Langdon Winner, *Autonomous Technology: Technics-out-of Control as a Theme of Political Thought* (Cambridge: MIT Press, 1977), p. 229.
13. Priscilla Regan, *Legislating Privacy* (Chapel Hill: University of North Carolina Press, 1995), p. 12.
14. Charles Taylor, *The Ethics of Authenticity* (Cambridge: Harvard University Press, 1991), p. 101.
15. Kenneth Goodpaster, "Some Avenues for Ethical Analysis in Management," in *Policies and Persons* ed. John Matthews, et al. (New York: McGraw-Hill, 1985), p. 495.
16. See, for example, James Rachels, *The Elements of Moral Philosophy* (New York: Random House, 1986); and William K. Frankena, *Ethics* (Englewood Cliffs, NJ: Prentice-Hall, 1963).
17. Frankena, *Ethics*, p. 29.
18. Charles Kelbley, "Freedom from the Good," in *Freedom and Value*, ed. R. Johann (New York: Fordham University Press, 1975), p. 173.
19. Goodpaster, p. 497.
20. Kant, p. 16.
21. Kant, p. 18.
22. Christine Korsgaard, *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996), p. 78.
23. Ibid, p. 92.
24. Ibid, p. 21.
25. Norman Bowie, *Business Ethics: A Kantian Perspective* (Oxford: Blackwell Publishers, 1999), p. 17.

26. See Korsgaard, p. 97.
27. R. A. Spinello, "Beyond Copyright: A Moral Investigation of Intellectual Property Protection in Cyberspace," in *The Impact of the Internet on our Moral Lives* ed. R. J. Cavalier (Albany, NY: SUNY Press, 2005), pp. 27–48.
28. Kant, p. 36.
29. Korsgaard, p. 194.
30. A. C. Ewing, *Ethics* (New York: Free Press, 1965), p. 58.
31. William D. Ross, *The Right and the Good* (Oxford: Oxford University Press, 1930).
32. Alasdair MacIntyre, *Three Rival Versions of Moral Enquiry: Encyclopaedia, Genealogy, Tradition* (Notre Dame, IN: University of Notre Dame Press, 1990), p. 194.
33. John Finnis, *Natural Law and Natural Rights* (New York: Oxford University Press, 1980), p. 225.
34. Germain Grisez, "A Contemporary Natural Law Ethic," in *Normative Ethics and Objective Reason*, ed. G. McLean (2001). http://216.25.45.103/book/Series01/1-11/chapter_xi.htm.
35. Ibid.
36. John Finnis, *Aquinas* (Oxford: Oxford University Press, 1998), p. 141.
37. Finnis, *Natural Law and Natural Rights*, p. 197.
38. Anthony Lisska, *Aquinas' Theory of Natural Law* (New York: Oxford University Press, 1996), p. 161.
39. Grisez.
40. Thomas Beauchamp and J. F. Childress, *Principles of Biomedical Ethics*, 4th ed. (New York: Oxford University Press, 1994).
41. Mark Kaczeski, "Casuistry and the Four Principles Approach," in *Encyclopedia of Applied Ethics* ed. Ruth Chadwick (San Diego, CA: Academic Press, 1998), vol. 1, p. 430.
42. Gary Doppelt, "Beyond Liberalism and Communitarianism: A Critical Theory of Social Justice," *Philosophy and Social Criticism* 14 (1988): 278.
43. Jon Gunneman, et al., *The Ethical Investor* (New Haven, CT: Yale University Press, 1972), p. 20.
44. John Rawls, *A Theory of Justice* (Cambridge: Harvard University Press, 1971), pp. 85–90.
45. Karl Marx, *Critique of the Gotha Program* (London: Lawrence and Werhart, Ltd., 1938), p. 14.

Welton on the philosophy of body, with Ken Baynes on the Frankfurt School ideas about community, and in particular from discussions with Norah Martin on social theories of the self and with Emily Zakin on the construction of gender.

In thanking Stony Brook colleagues, I would be remiss if I forgot to thank Patrick Heelan, whose actions as Dean of Humanities and Fine Arts back in 1991-92 first piqued my interest in the legal issues of cyberspace, and led me to make my initial contacts with the Electronic Frontier Foundation. I might have found the EFF in any case, but I doubt that I would have had the same empathy for its causes had I not myself been the target of extreme technophobia.

Even with all the help just mentioned, this project absolutely would not have happened were it not for the encouragement (i.e., arm twisting) by Teri Mendelsohn and Amy Pierce of the MIT Press. There were times (most of the times, in fact) when I didn't think I could afford to expend the effort necessary to complete this project. As it turns out, I was right. I couldn't afford it. Still, somehow I did it, and I'm glad I did.

Much credit is also due my students in phil 285, who were assigned an earlier incarnation of this material, and who helped me fine-tune the reading list. Although that course ended with the last class of Fall 1994, it lives on through the course e-list which continues to be active as of this writing. (School is never out in cyberspace.)

Finally, I want to thank Lori Repetti, who has discussed virtually every reading in this volume with me, and who has read and commented on all of the section introductions. Her contribution to the paper on the Italian crackdown was also invaluable, as she not only provided helpful comments, but also corrected errors in my translations from the Italian. I won't say the usual words here about how she was so understanding while I devoted my attention to this project. Rather, I want to thank her for taking time to help me, when those energies should have been devoted to her much more important research on the (rapidly dying) dialects of Northern Italy. For these, and many other reasons, this book is dedicated to her.

I

Property Rights, Piracy, etc.: Does Information "Want to Be Free"?

Peter Ludlow

In the industrialized Western nations we have a fairly well developed notion of property and property rights. We are inclined to think, for example, that we can own pieces of land if we pay for them and that we can likewise own the minerals in that land. More generally, we are inclined to think that if we have legally purchased something like a car or a bicycle, it is ours. No one may take it from us without fair compensation. There are other views, or course. A Marxist might argue that all property belongs to the state (or the workers, after the state has withered away). Others might argue that no one (neither persons nor nations) has sole claim to land. This is a view often attributed to hunter-gatherer cultures, for example.

It would probably be useful for philosophers to give more serious consideration to the nature of property rights and how they are grounded, but, even if our assumptions about property rights are poorly grounded, in the case of physical property they at least have the virtue of being clear. The same cannot be said about our grasp of intellectual property rights.

To see the problem raised by intellectual property rights, simply contrast the case of the car that I own with the case of a program I have written. You can steal my car, and in doing so you deprive me of my property (at least until the insurance company pays up). But if you copy my computer program (in effect, if you pirate it) I still have the program. In a certain sense, intellectual property cannot be "stolen." Indeed, the law distinguishes "theft" (which applies to physical property), from "infringement" (which applies to intellectual property). You can infringe on my intellectual property rights by copying my program or song or patent, and in so doing you apparently deprive me of income, but it isn't exactly the same thing as stealing.

A number of individuals have gone so far as to question the very idea of intellectual property rights (see, for example, the readings in this section by Barlow and by Garfinkel, Stallman, and Kapor). It is one thing, they say, for someone to claim ownership of land or the means of production, but it is quite another for someone to claim ownership of ideas or information. Suppose, for example, that I discover a property of natural numbers that turns out to be useful in encryption technology. Is

it really right that I should claim ownership of this idea? Isn't it the height of hubris for me to claim ownership of a mathematical law that, for all we know, some clever extraterrestrial programmer came up with centuries ago?

This issue has all the markings of a classical ethical dispute. On the one hand there are those who claim that it is a basic right to enjoy the fruits of one's labors, while on the other hand there are those who claim that no labors can entitle one to ownership of information. How are such disputes to be settled? A number of participants in the debate appear to make "consequentialist" arguments for their respective positions. In the great tradition of utilitarian philosophers like Jeremy Bentham and John Stuart Mill, they argue that protecting (or, contrarily, ignoring) intellectual property rights will contribute to the greatest good for the greatest number of people.

For example, some argue that if intellectual property rights are not protected no one will take the trouble to develop programs and patents and so forth (see, for example, the reading from Heckel). Perhaps potential programmers will go into law or accounting or (even worse) philosophy, and we shall all be poorer as a consequence. On the other hand there are those (e.g., Garfinkel, Stallman, and Kapor) who argue that eliminating software copyrights would actually contribute to the development and distribution of ideas and that it would benefit the programmer as well. Which side is right? For the utilitarian it may come down to a debate over the economic consequences of enforced property rights.

Of course, utilitarianism is not the only brand of ethics in town, and other folks might be inclined to argue for property rights on the basis of other considerations (e.g., natural law). Such arguments are difficult to find, or at least difficult to find in any sort of articulate form. Still, one can imagine arguments of this character, and I suppose they need to be considered. For example, one might think that quite apart from the economic consequences of protecting property there is the deeper ethical principle that one has the right to the fruits of one's labors. Of course, for the most part, principles of this nature tend to be defended on economic grounds (people produce more when they keep the fruits of their labors, etc.), but if this economic argument were somehow discredited it is likely that the general ethical principle would be retained—at least in some quarters. But why would it be retained? That is difficult to say.

It is one thing to ask, in the abstract, whether intellectual property rights are a good thing. It is quite another to ask whether, given our current laws, intellectual property rights should be flouted—that is, whether software piracy should be (illegally) practiced. The editorial from the electronic 'zine *Pirate* argues that it should. The stated justification is that pirate boards contribute to the widespread distribution of software, and that eventually corporations will want to buy that software for service, upgrades, and so forth.

Before we consider the merits of such a position we should perhaps get clear on what piracy is. For example, we need to distinguish software "piracy" from software "bootlegging." Roughly, the bootlegger copies software for profit (e.g., making multiple copies and selling them), while the pirate merely makes illicit copies for personal use, and perhaps for swapping pirated software with friends. A "pirate board," then, is a BBS set up to be a location where individuals can swap pirated software, and where money is not changing hands.

Strictly speaking, most of the readers of this book are software pirates. Nearly all of you have at least one piece of software that is unregistered or that was copied by a friend. Some of you may even have swapped software on a pirate board. Technically, this is copyright infringement and against the law; but just how wrong is it? For example, it is technically against the law to drive 56 miles an hour in a 55 mph zone, but we would hardly consider someone who drove over 55 a menace to society (in fact, we might reserve that opinion for those who obstruct traffic by puttering along at 55). Is software piracy like breaking the speed limit a little? Or is it even more justifiable, because in the long run it leads to good ends? Or is software piracy a form of civil disobedience? That may seem like a stretch, but if you really believed that information "wants to be free" (i.e., that it is wrong for individuals to hoard information), you might be tempted by this argument.

Even if we agree that software piracy is wrong, a number of conceptual issues surround that of how we should respond to piracy when we encounter it. First of all, what is the value of the pirated intellectual

property? With normal property theft—for example, when someone steals my toaster—it is clear enough how to measure the loss. But how does one measure the loss of a piece of pirated software? By the retail value of the software? Such measures seem more apt for evaluating the loss from the theft of toasters. If someone steals a toaster from the store, the store will be unable to recoup its investment in that piece of inventory, but if someone downloads a program from a pirate board, no one's inventory is diminished. Nor is it reasonable to suppose that everyone who pulls free software off a pirate board would have otherwise purchased that piece of software. Precisely how do we quantify the losses (if, indeed, there are any)?

The situation becomes even more complex when the piece of intellectual property is not a program but a trade secret or a piece of proprietary information. The reading by Godwin discusses one of the most celebrated cases of such piracy (for background, see also the discussion in Barlow's "Crime and Puzzlement" in appendix 1). After a hacker downloaded a description of AT&T's 911 system, AT&T calculated the value of the loss at several hundred thousand dollars. The case was eventually dismissed when it turned out that another branch of AT&T was selling the same document for just a few dollars. Where did AT&T go wrong in evaluating its losses? That remains to be seen. This case, however, shows just how easy it is to be off by several orders of magnitude in evaluating such losses.

It is also worth keeping in mind that whatever the evils of pirated intellectual property, attempts to thwart such piracy can be disastrously counterproductive. Several recent "crackdowns" have highlighted the dangers. One example surrounded the case of the pilfered 911 document and subsequent events, such as the bust of Steve Jackson Games discussed by Barlow in appendix 1. Another example is the Italian crackdown code-named "Hardware 1" (see appendix 2) in which nearly one third of the electronic bulletin boards in Italy were busted because of the pretense of software piracy. In each case the rights of a number of innocent people were trampled by government zeal to crack down on alleged piracy of some form or another. The moral is that whatever the evils of piracy might be, it does not immediately follow that any arbitrary ham-handed

governmental response is appropriate. To the contrary, responses will have to be measured, and government institutions will need to be cognizant of the rights of innocent bystanders. Is effective government action even possible within such constraints? That is very much an open question. It may well be that government institutions (American, Italian, or whatever) are just not the appropriate organizations for halting the activities of amateur pirates.

• Encryption will be the technical basis for most intellectual property protection. (And should, for this and other reasons, be made more widely available.)
• The economy of the future will be based on relationship rather than possession. It will be continuous rather than sequential.

And finally, in the years to come, most human exchange will be virtual rather than physical, consisting not of stuff but the stuff of which dreams are made. Our future business will be conducted in a world made more of verbs than nouns.

Ojo Caliente, New Mexico, October 1, 1992
New York, New York, November 6, 1992
Brookline, Massachusetts, November 8, 1992
New York, New York, November 15, 1993
San Francisco, California, November 20, 1993
Pinedale, Wyoming, November 24–30, 1993
New York, New York, December 13–14, 1993

This expression has lived and grown to this point over the time period and in the places detailed above. Despite its print publication here, I expect it will continue to evolve in liquid form, possibly for years.

The thoughts in it have not been "mine" alone but have assembled themselves in a field of interaction that has existed between myself and numerous others, to whom I am grateful. They particularly include: Pamela Samuelson, Kevin Kelly, Mitch Kapor, Mike Godwin, Stewart Brand, Mike Holderness, Miriam Barlow, Danny Hillis, Trip Hawkins, and Alvin Toffler.

2

Why Patents Are Bad for Software

Simson L. Garfinkel, Richard M. Stallman,
and Mitchell Kapor

In September 1990, users of the popular XyWrite word processing program got a disturbing letter in the mail from XyQuest, Inc., the program's publisher:

> In June of 1987, we introduced an automatic correction and abbreviation expansion feature in XyWrite III Plus. Unbeknownst to us, a patent application for a related capability had been filed in 1984 and was subsequently granted in 1988. The company holding the patent contacted us in late 1989 and apprised us of the existence of their patent.
>
> We have decided to modify XyWrite III Plus so that it cannot be construed as infringing. The newest version of XyWrite III Plus (3.56) incorporates two significant changes that address this issue: You will no longer be able to automatically correct common spelling errors by pressing the space bar after the misspelled word. In addition, to expand abbreviations stored in your personal dictionary, you will have to press control-R or another designated hot key.

XyQuest had been bitten by a software patent—one of the more than two thousand patents on computer algorithms and software techniques that have been granted by the U.S. Patent and Trademark Office since the mid-1980s. The owner of the patent, Productivity Software, had given XyQuest a choice: license the patent or take a popular feature out of XyWrite, XyQuest's flagship product. If XyQuest refused, a costly patent-infringement lawsuit was sure to follow.

Some choice.

XyQuest tried to license the patent, says Jim Adelson, vice president for marketing, but Productivity Software kept changing its terms. First Productivity said that XyQuest could keep the feature in some versions of XyWrite, but not in others. Then the company said that XyQuest could use one part of the "invention," but not other parts. And Productivity

Software kept increasing the amount of money it wanted. XyQuest finally gave up and took the feature out.

XyQuest was lucky it had that option. Other firms—including some of the nation's largest and most profitable software publishers—have been served with notice of patents that strike to the heart of their corporate vitality. In one of the most publicized cases, a company called Refac International—whose sole business is acquiring and litigating patents—sued Lotus, Microsoft, Ashton-Tate, and three other spreadsheet publishers, claiming they had all infringed on patent number 4,398,249, which spells out the order in which to recalculate the values in a complicated model when one parameter in the model changes. (Refac has since dropped its claims against all the companies except Lotus, but only because company lawyers anticipated a better chance of success if they faced just one opponent.)

Patent 4,398,249 does not have anything to do with spreadsheets in particular; the technique also appears in some graphics drawing and artificial intelligence programs. And the idea that values in a spreadsheet should be recalculated in the order specified by the patent is so obvious that it has probably occurred to nearly everyone who has written a spreadsheet program. But the Patent Office's standard for obviousness is extremely low; patents have been granted for ideas so elementary that they could have been answers to problems in a first-year programming course.

Practically once a month, the nation's computer networks are abuzz with news of another patent issued on a fundamental concept that is widely used. Although the Patent Office isn't supposed to grant patents on ideas, that's essentially what it's doing with software patents, carving up the intellectual domain of computer science and handing little pieces to virtually any company that files an application. And the practice is devastating America's software industry.

If Congress does not act quickly to redefine the applicability of patent law to computer programs, the legal minefield confronting the introduction of new computer programs will be so intimidating—and potentially so costly—that small companies will effectively be barred from the marketplace, while large, established firms will become embroiled in litigation that will have a stultifying effect on the entire industry.

What's Being Patented?

Software patents do not cover entire programs; instead, they cover algorithms and techniques—the instructions that tell a computer how to carry out a specific task in a program. Thousands of instructions make up any one computer program. But whereas the unique combination of algorithms and techniques in a program is considered an "expression" (like a book or a song) and is covered by copyright law, the algorithms and techniques themselves are treated as procedures eligible for patenting.

The judicial basis for this eligibility is tenuous at best. U.S. law does not allow inventors, no matter how brilliant they are, to patent the laws of nature, and in two Supreme Court cases (*Gottschalk v. Benson*, 1972, and *Parker v. Flook*, 1978) the Court extended this principle to computer algorithms and software techniques. But in the 1981 case *Diamond v. Diehr*, the Court said that a patent could be granted for an industrial process that was controlled by certain computer algorithms, and the Patent Office seems to have taken that decision as a green light on the patentability of algorithms and techniques in general.

Software patents are now being granted at an alarming rate—by some counts, more than a thousand are issued each year. Unfortunately, most of the patents have about as much cleverness and originality as a recipe for boiled rice—simple in itself but a vital part of many sophisticated dishes. Many cover very small and specific algorithms or techniques that are used in a wide variety of programs. Frequently the "inventions" mentioned in a patent application have been independently formulated and are already in use by other programmers when the application is filed.

When the Patent Office grants a patent on an algorithm or technique, it is telling programmers that they may not use a particular method for solving a problem without the permission of the idea's "owner." To them, patenting an algorithm or technique is like patenting a series of musical notes or a chord progression, then forcing composers to purchase a "musical sequence license."

Systems at Odds

The traditional rationale for patents is that protection of inventions will spur innovation and aid in the dissemination of information about technical advances. By prohibiting others from copying an invention, patents allow inventors to recoup their investment in development while at the same time revealing the workings of the new invention to the public.

But there's evidence that the patent system is backfiring in the computer industry; indeed, the system itself seems unsuited to the nature of software development. Today's computer programs are so complex that they contain literally thousands of algorithms and techniques, each considered patentable by the Patent Office's standards. Is it reasonable to expect a software company to license each of those patents, or even to bring such a legally risky product into the marketplace? To make things even more complicated, the Patent Office has also granted patents on combinations of algorithms and techniques that produce a particular feature. For example, Apple was sued because its Hypercard program allegedly violates patent number 4,736,308, which covers a specific technique that, in simplified terms, entails scrolling through a database displaying selected parts of each line of text. Separately, the scrolling and display functions are ubiquitous fixtures of computer programming, but combining them without a license from the holder of patent 4,736,308 is now apparently illegal.

Another problem with patenting software is the amount of time it takes to do so. The two to five years required to file for and obtain a patent are acceptable if a company is patenting, say, the formula for Valium, which hasn't changed in more than 20 years. But in the software industry, companies that don't continually bring out new versions of their programs go out of business. Success for them depends on spotting needs and developing solutions as quickly as possible.

Unfortunately, conducting a patent search is a slow, deliberative process that, when harnessed to software development, could stop innovation in its tracks. And because patent applications are confidential, there is simply no way for computer programmers to ensure that what they write will not violate some patent that is yet to be issued. Thus XyQuest "reinvented" its automatic spelling-error correction system and brought

the product to market between the time that Productivity Software had filed for its application and been awarded the patent.

Such examples are becoming increasingly common. In another case, the journal *IEEE Computer* in June 1984 published a highly efficient algorithm for performing data compression; unbeknownst to the journal's editors or readers, the authors of the article had simultaneously applied for a patent on their invention. In the following year, numerous programs were written and widely distributed for performing the so-called "LZW data compression." The compression system was even adopted as a national standard and proposed as an international one. Then, in 1985, the Patent Office awarded patent number 4,558,302 to one of the authors of the article. Now Unisys, the holder of the patent, is demanding royalties for the use of the algorithm. Although programs incorporating the algorithm are still in the public domain, using these programs means risking a lawsuit.

Not only is the patent approval process slow, but the search for "prior art"—the criterion the Patent Office uses to determine whether an invention already exists at the time of a patent application—is all but impossible to conduct in the realm of computer software. After more than 25 years, the Patent Office has not developed a system for classifying patents on algorithms and techniques, and no such system may be workable. Just as mathematicians are sometimes unaware that essentially identical mental processes are being used in separate areas of mathematics under different terminology, different parts of computer science frequently reinvent the same algorithm to serve different purposes. It is unreasonable to expect that a patent examiner, pressed for time, would recognize all such duplication. For example, IBM was issued a patent on the same data-compression algorithm that Unisys supposedly owns. The Patent Office was probably not aware of granting two patents for the same algorithm because the descriptions in the patents themselves are quite different even though the formulas are mathematically equivalent.

The search for prior art is complicated by the fact that the literature of computer science is unbelievably large. It contains not only academic journals, but also users' manuals, published source code, and popular accounts in magazines for computer enthusiasts. Whereas a team of chemists working at a major university might produce 20 or 30 pages of

published material per year, a single programmer might easily produce a hundred times that much. The situation becomes even more complex in the case of patented combinations of algorithms and techniques. Programmers often publish new algorithms and techniques, but they almost never publish new ways of combining old ones. Although individual algorithms and techniques have been combined in many different ways in the past, there's no good way to establish that history.

The inability to search the literature thoroughly for prior art is crucial, because unless an examiner can find prior art, he or she is all but obligated to issue the patent. As a result, many patents have been granted—and successfully defended in court—that are not "original," even by the Patent Office's definition. It was simply the case that neither the patent examiner nor the defendants in the lawsuit knew of the prior art's existence.

Some members of the commercial software community are now proposing the creation of a "Software Patent Institute" to identify software's prior art that existed before 1980. But even if such an institute could catalogue every discovery made by every programmer in the United States, it makes no sense to arbitrarily declare that only pre-1980 work is in the public domain. Besides, what would be the purpose? To allow the patenting of nature's mathematical laws?

Bad for Business

Even when patents *are* known in advance, software publishers have generally not licensed the algorithms or techniques; instead, they try to rewrite their programs to avoid using the particular procedure that the patent describes. Sometimes this isn't possible, in which case companies have often chosen to avoid implementing new features altogether. It seems clear from the evidence of the last few years that software patents are actually *preventing* the adoption of new technology, rather than encouraging it.

And they don't seem to be encouraging innovation, either. Software patents pose a special danger to small companies, which often form the vanguard of software development but can't afford the cost of patent searches or litigation. The programming of a new product can cost a few hundred thousand dollars; the cost of a patent search for each technique

and combination of techniques that the new program uses could easily equal or even exceed that. And the cost of a single patent suit can be more than a million dollars.

"I'm not familiar with any type of litigation that is any more costly than patent litigation," says R. Duff Thompson, vice president and general counsel of the WordPerfect Corporation. But Thompson's greatest fear is that software patents will wipe out young, independent programmers, who until now have been the software industry's source of inspiration. Imagine what happens, says Thompson, when "some 23-year-old kid who has a terrific idea in a piece of software is hammered by a demand letter from someone holding a patent."

As for aiding the exchange of information, the expansion of software patents could mean instead the end of software developed at universities and distributed without charge—software that has been a mainstay of computer users in universities, corporations, and government for years. Many such programs—the X Window system, the EMACS text editor, the "compress" file-compression utility, and others—appear to be in violation of existing patents. Patents could also mean an end to public-domain software, which has played an important part in making computers affordable to public schools. There is obviously no way that an author who distributes a program for free could arrange to pay for royalties if one of the hundreds of techniques that were combined to create the program happens to be patented.

Few programmers and entrepreneurs believe that patents are necessary for their profession. Instead, the impetus for patents on algorithms and techniques comes from two outside sources: managers of large companies, who see patents as a means for triumphing over their competitors without having to develop superior products, and patent attorneys, who see the potential for greatly expanding their business.

Today, most patenting by companies is done to have something to trade or as a defense against other patent-infringement suits. Attorneys advise that patenting software may strengthen competitive position. Although this approach will work for large companies such as Microsoft, Apple, and IBM, small and even mid-sized companies can't play in their league. A future startup will be forced to pay whatever price the giants choose to impose.

Copyright and Trade Secrecy

The best argument against the wisdom of software patents may be history itself. Lotus, Microsoft, WordPerfect, and Novell all became world leaders in the software publishing industry on the strength of their products. None of these companies needed patents to secure funding or maintain their market position. Indeed, all made their fortunes before the current explosion of software patents began. Clearly patents are not necessary to ensure the development of computer programs. And for those who want more control over what they see as their property, the computer industry has already adopted two other systems: copyright and trade secrecy.

Today, nearly all programs are copyrighted. Copyright prohibits the users of a software program from making copies of it (for example, to give to their friends) without the permission of the individual or company that licenses the program. It prevents one company from appropriating another company's work and selling it as its own. But the existence of a copyright doesn't prevent other programmers from using algorithms or techniques contained in the program in their own work. A single software technique can be implemented in different ways to do totally different jobs; copyright only prohibits appropriating the actual code that a particular programmer wrote.

In general, copyrighting and patenting are thought to apply to very different kinds of material: the former to the expression of ideas, and the latter to a process that achieves a certain result. Until just a few years ago, computer algorithms and techniques were widely seen as unpatentable. And as Harvard University policy analyst Brian Kahin notes, this is the first time in history that an industry in which copyright was widely established was suddenly subjected to patenting.

Indeed, without conscious action by Congress or the Supreme Court, the most fundamental rule of software publishing—if you write a program, you own it—will change. The new rule will be that you might own what you write—if it is so revolutionary that it owes nothing to any previous work. No author in areas other than software is held to such an unrealistically high standard.

The U.S. patent system was created because the framers of the Constitution hoped that patents would discourage trade secrecy. When techniques are kept secret for commercial advantage, they may never become available for others to use and may even be lost. But although trade secrecy is a problem for software, as it is for other fields, it is not a problem that patents help to correct.

Many of the useful developments in the field of software consist of new features such as the automatic correction and abbreviation expansion feature in XyWrite III Plus. Since it is impossible to keep a program's features secret from the users of the program, there is no possibility of trade secrecy and thus no need for measures to discourage it. Techniques used internally in a software system can be kept secret; but in the past, the important ones rarely were. It was normal for computer scientists in the commercial as well as the academic world to publish their discoveries. Once again, since secrecy about techniques was not a significant problem, there is little to be gained by adopting the patent system to discourage it.

The place where trade secrecy *is* used extensively in software is in the "source code" for programs. In computer programming, trade secrets are kept by distributing programs in "machine code," the virtually indecipherable translation of programming languages that computers read. It is extremely difficult for another programmer to glean from a machine-code program the original steps written by the program's author. But software patents haven't done anything to limit this form of trade secrecy. By withholding the source code, companies keep secret not a particular technique, but the way that they have combined dozens of techniques to produce a design for a complete system. Patenting the whole design is impractical and ineffective. Even companies that have software patents still distribute programs in machine code only. Thus, in no area do software patents significantly reduce trade secrecy.

Reversing Direction

Many policymakers assume that any increase in intellectual property protection must be good for whoever works in the field. As we've tried to show, this is assuredly not the case in the field of computer programming. Nearly all programmers view patents as an unwelcome intrusion, limiting both their ability to do their work and their freedom of expression.

At this point, so many patents have been issued by the Patent and Trademark Office that the prospect of overturning them by finding prior art, one at a time, is almost unthinkable. Even if the Patent Office learns to understand software better in the future, the mistakes that are being made now will follow the industry into the next century unless there is a dramatic turnaround in policy.

The U.S. Patent and Trademark Office recently established an Advisory Commission on Patent Law Reform that is charged with examining a number of issues, including software patents—or what it prefers to call patents on "computer-program-related inventions." Unfortunately, the commission's subcommittee on software does not include any prominent software industry representatives who have expressed doubts about software patents. But the subcommittee is required to consider public comment.

Although influencing the Patent Office might produce some benefits, the really necessary reforms are likely to come only through intervention by the Supreme Court or Congress. Waiting for Court action is not the answer: No one can force the Supreme Court to rule on a relevant case, and there is no guarantee that the Court would decide to change Patent Office practice or to do anything about existing patents. The most effective course of action, therefore, is to encourage Congress to amend the patent law to disallow software patents and, if possible, invalidate those that have already been awarded. The House Subcommittee on Intellectual Property and the Administration of Justice should take the lead by scheduling hearings on the subject and calling for a congressionally sponsored economic analysis of the effect of software patents on the industry.

The computer industry grew to be vibrant and healthy without patents. Unless those who want software patents can demonstrate that they are necessary to the health of the industry, Congress should feel justified in eliminating this barrier to innovation.

Recommended Reading

Brian Kahin, "The Software Patent Crisis," *Technology Review* (April 1990): 53–58.

Mitchell Kapor, Testimony at Hearings before U.S. House of Representatives, Subcommittee on Courts, Intellectual Property and the Administration of Justice, of the Committee on the Judiciary (March 5, 1990).

Pamela Samuelson, "Benson Revisited: Should Patent Protection Be Available for Algorithms and Other Computer Program-Related Inventions?" *Emory Law Journal* (Fall 1990): 1025–1154.

Pamela Samuelson, "Should Program Algorithms Be Patented?" *Communications of the ACM* (August 1990): 23–27.

1

Selling Wine without Bottles: The Economy of Mind on the Global Net

John Perry Barlow

If nature has made any one thing less susceptible than all others of exclusive property, it is the action of the thinking power called an idea, which an individual may exclusively possess as long as he keeps it to himself; but the moment it is divulged, it forces itself into the possession of everyone, and the receiver cannot dispossess himself of it. Its peculiar character, too, is that no one possesses the less, because every other possesses the whole of it. He who receives an idea from me, receives instruction himself without lessening mine; as he who lights his taper at mine, receives light without darkening me. That ideas should freely spread from one to another over the globe, for the moral and mutual instruction of man, and improvement of his condition, seems to have been peculiarly and benevolently designed by nature, when she made them, like fire, expansible over all space, without lessening their density at any point, and like the air in which we breathe, move, and have our physical being, incapable of confinement or exclusive appropriation. Inventions then cannot, in nature, be a subject of property.

—Thomas Jefferson

Throughout the time I've been groping around Cyberspace, there has remained unsolved an immense conundrum that seems to be at the root of nearly every legal, ethical, governmental, and social vexation to be found in the Virtual World. I refer to the problem of digitized property.

The riddle is this: if our property can be infinitely reproduced and instantaneously distributed all over the planet without cost, without our knowledge, without its even leaving our possession, how can we protect it? How are we going to get paid for the work we do with our minds? And, if we can't get paid, what will assure the continued creation and distribution of such work?

Since we don't have a solution to what is a profoundly new kind of challenge, and are apparently unable to delay the galloping digitization

of everything not obstinately physical, we are sailing into the future on a sinking ship.

This vessel, the accumulated canon of copyright and patent law, was developed to convey forms and methods of expression entirely different from the vaporous cargo it is now being asked to carry. It is leaking as much from within as without.

Legal efforts to keep the old boat floating are taking three forms: a frenzy of deck chair rearrangement, stern warnings to the passengers that if she goes down, they will face harsh criminal penalties, and serene, glassy-eyed denial.

Intellectual property law cannot be patched, retrofitted, or expanded to contain the gasses of digitized expression any more than real estate law might be revised to cover the allocation of broadcasting spectrum. (Which, in fact, rather resembles what is being attempted here.) We will need to develop an entirely new set of methods as befits this entirely new set of circumstances.

Most of the people who actually create soft property—the programmers, hackers, and Net surfers—already know this. Unfortunately, neither the companies they work for nor the lawyers these companies hire have enough direct experience with immaterial goods to understand why they are so problematic. They are proceeding as though the old laws can somehow be made to work, either by grotesque expansion or by force. They are wrong.

The source of this conundrum is as simple as its solution is complex. Digital technology is detaching information from the physical plane, where property law of all sorts has always found definition.

Throughout the history of copyrights and patents, the proprietary assertions of thinkers have been focused not on their ideas but on the expression of those ideas. The ideas themselves, as well as facts about the phenomena of the world, were considered to be the collective property of humanity. One could claim franchise, in the case of copyright, on the precise turn of phrase used to convey a particular idea or the order in which facts were presented.

The point at which this franchise was imposed was that moment when the "word became flesh" by departing the mind of its originator and entering some physical object, whether book or widget. The subsequent

arrival of other commercial media besides books didn't alter the legal importance of this moment. Law protected expression and, with few (and recent) exceptions, to express was to make physical.

Protecting physical expression had the force of convenience on its side. Copyright worked well because, Gutenberg notwithstanding, it was hard to make a book. Furthermore, books froze their contents into a condition that was as challenging to alter as it was to reproduce. Counterfeiting or distributing counterfeit volumes were obvious and visible activities, easy enough to catch somebody in the act of doing. Finally, unlike unbounded words or images, books had material surfaces to which one could attach copyright notices, publisher's marques, and price tags.

Mental to physical conversion was even more central to patent. A patent, until recently, was either a description of the form into which materials were to be rendered in the service of some purpose or a description of the process by which rendition occurred. In either case, the conceptual heart of patent was the material result. If no purposeful object could be rendered due to some material limitation, the patent was rejected. Neither a Klein bottle nor a shovel made of silk could be patented. It had to be a thing and the thing had to work.

Thus the rights of invention and authorship adhered to activities in the physical world. One didn't get paid for ideas but for the ability to deliver them into reality. For all practical purposes, the value was in the conveyance and not the thought conveyed.

In other words, the bottle was protected, not the wine.

Now, as information enters Cyberspace, the native home of Mind, these bottles are vanishing. With the advent of digitization, it is now possible to replace all previous information storage forms with one meta-bottle: complex—and highly liquid—patterns of ones and zeros.

Even the physical/digital bottles to which we've become accustomed, floppy disks, CD-ROM's, and other discrete, shrink-wrappable bit-packages, will disappear as all computers jack in to the global Net. While the Internet may never include every single CPU on the planet, it is more than doubling every year and can be expected to become the principal medium of information conveyance, and perhaps eventually, the only one.

Once that has happened, all the goods of the Information Age—all of expressions once contained in books or film strips or records or

newsletters—will exist either as pure thought or something very much like thought: voltage conditions darting around the Net at the speed of light, in conditions which one might behold in effect, as glowing pixels or transmitted sounds, but never touch or claim to "own" in the old sense of the word.

Some might argue that information will still require some physical manifestation, such as its magnetic existence on the titanic hard disks of distant servers, but these are bottles that have no macroscopically discrete or personally meaningful form.

Some will also argue that we have been dealing with unbottled expression since the advent of radio, and they would be right. But for most of the history of broadcast, there was no convenient way to capture soft goods from the electromagnetic ether and reproduce them in anything like the quality available in commercial packages. Only recently has this changed and little has been done legally or technically to address the change.

Generally, the issue of consumer payment for broadcast products was irrelevant. The consumers themselves were the product. Broadcast media were supported either by selling the attention of their audience to advertisers, using government to assess payment through taxes, or the whining mendicancy of annual donor drives.

All of broadcast support models are flawed. Support either by advertisers or government has almost invariably tainted the purity of the goods delivered. Besides, direct marketing is gradually killing the advertiser support model anyway.

Broadcast media gave us another payment method for a virtual product in the royalties which broadcasters pay songwriters through such organizations as ASCAP and BMI. But, as a member of ASCAP, I can assure you this is not a model that we should emulate. The monitoring methods are wildly approximate. There is no parallel system of accounting in the revenue stream. It doesn't really work. Honest.

In any case, without our old methods of physically defining the expression of ideas, and in the absence of successful new models for non-physical transaction, we simply don't know how to assure reliable payment for mental works. To make matters worse, this comes at a time when the human mind is replacing sunlight and mineral deposits as the principal source of new wealth.

Furthermore, the increasing difficulty of enforcing existing copyright and patent laws is already placing in peril the ultimate source of intellectual property, the free exchange of ideas.

That is, when the primary articles of commerce in a society look so much like speech as to be indistinguishable from it, and when the traditional methods of protecting their ownership have become ineffectual, attempting to fix the problem with broader and more vigorous enforcement will inevitably threaten freedom of speech.

The greatest constraint on your future liberties may come not from government but from corporate legal departments laboring to protect by force what can no longer be protected by practical efficiency or general social consent.

Furthermore, when Jefferson and his fellow creatures of The Enlightenment designed the system that became American copyright law, their primary objective was assuring the widespread distribution of thought, not profit. Profit was the fuel that would carry ideas into the libraries and minds of their new republic. Libraries would purchase books, thus rewarding the authors for their work in assembling ideas, which otherwise "incapable of confinement" would then become freely available to the public. But what is the role of libraries if there are no books? How does society now pay for the distribution of ideas if not by charging for the ideas themselves?

Additionally complicating the matter is the fact that along with the physical bottles in which intellectual property protection has resided, digital technology is also erasing the legal jurisdictions of the physical world, and replacing them with the unbounded and perhaps permanently lawless seas of Cyberspace.

In Cyberspace, there are not only no national or local boundaries to contain the scene of a crime and determine the method of its prosecution, there are no clear cultural agreements on what a crime might be. Unresolved and basic differences between European and Asian cultural assumptions about intellectual property can only be exacerbated in a region where many transactions are taking place in both hemispheres and yet, somehow, in neither.

Even in the most local of digital conditions, jurisdiction and responsibility are hard to assess. A group of music publishers filed suit against

Compuserve this fall for it having allowed its users to upload musical compositions into areas where other users might get them. But since Compuserve cannot practically exercise much control over the flood of bits that pass between its subscribers, it probably shouldn't be held responsible for unlawfully "publishing" these works.

Notions of property, value, ownership, and the nature of wealth itself are changing more fundamentally than at any time since the Sumerians first poked cuneiform into wet clay and called it stored grain. Only a very few people are aware of the enormity of this shift and fewer of them are lawyers or public officials.

Those who do see these changes must prepare responses for the legal and social confusion that will erupt as efforts to protect new forms of property with old methods become more obviously futile, and, as a consequence, more adamant.

From Swords to Writs to Bits

Humanity now seems bent on creating a world economy primarily based on goods that take no material form. In doing so, we may be eliminating any predictable connection between creators and a fair reward for the utility or pleasure others may find in their works.

Without that connection, and without a fundamental change in consciousness to accommodate its loss, we are building our future on furor, litigation, and institutionalized evasion of payment except in response to raw force. We may return to the Bad Old Days of property.

Throughout the darker parts of human history, the possession and distribution of property was a largely military matter. "Ownership" was assured those with the nastiest tools, whether fists or armies, and the most resolute will to use them. Property was the divine right of thugs.

By the turn of the first millennium AD, the emergence of merchant classes and landed gentry forced the development of ethical understandings for the resolution of property disputes. In the late Middle Ages, enlightened rulers like England's Henry II began to codify this unwritten "common law" into recorded canons. These laws were local, but this didn't matter much as they were primarily directed at real estate, a form of property that is local by definition. And which, as the name implied, was very real.

This continued to be the case as long as the origin of wealth was agricultural, but with the dawning of the Industrial Revolution, humanity began to focus as much on means as ends. Tools acquired a new social value and, thanks to their own development, it became possible to duplicate and distribute them in quantity.

To encourage their invention, copyright and patent law were developed in most western countries. These laws were devoted to the delicate task of getting mental creations into the world where they could be used—and enter the minds of others—while assuring their inventors compensation for the value of their use. And, as previously stated, the systems of both law and practice that grew up around that task were based on physical expression.

Since it is now possible to convey ideas from one mind to another without ever making them physical, we are now claiming to own ideas themselves and not merely their expression. And since it is likewise now possible to create useful tools that never take physical form, we have taken to patenting abstractions, sequences of virtual events, and mathematical formulae—the most un-real estate imaginable.

In certain areas, this leaves rights of ownership in such an ambiguous condition that once again property adheres to those who can muster the largest armies. The only difference is that this time the armies consist of lawyers.

Threatening their opponents with the endless Purgatory of litigation, over which some might prefer death itself, they assert claim to any thought that might have entered another cranium within the collective body of the corporations they serve. They act as though these ideas appeared in splendid detachment from all previous human thought. And they pretend that thinking about a product is somehow as good as manufacturing, distributing, and selling it.

What was previously considered a common human resource, distributed among the minds and libraries of the world, as well as the phenomena of nature herself, is now being fenced and deeded. It is as though a new class of enterprise had arisen which claimed to own air and water.

What is to be done? While there is a certain grim fun to be had in it, dancing on the grave of copyright and patent will solve little, especially when so few are willing to admit that the occupant of this grave is even

deceased and are trying to uphold by force what can no longer be upheld by popular consent.

The legalists, desperate over their slipping grip, are vigorously trying to extend it. Indeed, the United States and other proponents of GATT are making adherence to our moribund systems of intellectual property protection a condition of membership in the marketplace of nations. For example, China will be denied Most Favored Nation trading status unless they agree to uphold a set of culturally alien principles that are no longer even sensibly applicable in their country of origin.

In a more perfect world, we'd be wise to declare a moratorium on litigation, legislation, and international treaties in this area until we had a clearer sense of the terms and conditions of enterprise in Cyberspace. Ideally, laws ratify already developed social consensus. They are less the Social Contract itself than a series of memoranda expressing a collective intent that has emerged out of many millions of human interactions.

Humans have not inhabited Cyberspace long enough or in sufficient diversity to have developed a Social Contract that conforms to the strange new conditions of that world. Laws developed prior to consensus usually serve the already established few who can get them passed and not society as a whole.

To the extent that either law or established social practice exists in this area, they are already in dangerous disagreement. The laws regarding unlicensed reproduction of commercial software are clear and stern—and rarely observed. Software piracy laws are so practically unenforceable and breaking them has become so socially acceptable that only a thin minority appears compelled, either by fear or conscience, to obey them.

I sometimes give speeches on this subject, and I always ask how many people in the audience can honestly claim to have no unauthorized software on their hard disks. I've never seen more than ten percent of the hands go up.

Whenever there is such profound divergence between the law and social practice, it is not society that adapts. And, against the swift tide of custom, the Software Publishers' current practice of hanging a few visible scapegoats is so obviously capricious as to only further diminish respect for the law.

Part of the widespread popular disregard for commercial software copyrights stems from a legislative failure to understand the conditions into which it was inserted. To assume that systems of law based in the physical world will serve in an environment that is as fundamentally different as Cyberspace is a folly for which everyone doing business in the future will pay.

As I will discuss in the next segment, unbounded intellectual property is very different from physical property and can no longer be protected as though these differences did not exist. For example, if we continue to assume that value is based on scarcity, as it is with regard to physical objects, we will create laws that are precisely contrary to the nature of information, which may, in many cases, increase in value with distribution.

The large, legally risk-averse institutions most likely to play by the old rules will suffer for their compliance. The more lawyers, guns, and money they invest in either protecting their rights or subverting those of their opponents, the more commercial competition will resemble the Kwakiutl Potlatch Ceremony, in which adversaries competed by destroying their own possessions. Their ability to produce new technology will simply grind to a halt as every move they make drives them deeper into a tar pit of courtroom warfare.

Faith in law will not be an effective strategy for high tech companies. Law adapts by continuous increments and at a pace second only to geology in its stateliness. Technology advances in the lunging jerks, like the punctuation of biological evolution grotesquely accelerated. Real world conditions will continue to change at a blinding pace, and the law will get further behind, more profoundly confused. This mismatch is permanent.

Promising economies based on purely digital products will either be born in a state of paralysis, as appears to be the case with multimedia, or continue in a brave and willful refusal by their owners to play the ownership game at all.

In the United States one can already see a parallel economy developing, mostly among small, fast-moving enterprises who protect their ideas by getting into the marketplace quicker than their larger competitors who base their protection on fear and litigation.

Perhaps those who are part of the problem will simply quarantine themselves in court while those who are part of the solution will create a new society based, at first, on piracy and freebooting. It may be that when

the current system of intellectual property law has collapsed, as seems inevitable, that no new legal structure will arise in its place.

But something will happen. After all, people do business. When a currency becomes meaningless, business is done in barter. When societies develop outside the law, they develop their own unwritten codes, practices, and ethical systems. While technology may undo law, technology offers methods for restoring creative rights.

A Taxonomy of Information

It seems to me that the most productive thing to do now is to look hard into the true nature of what we're trying to protect. How much do we really know about information and its natural behaviors?

What are the essential characteristics of unbounded creation? How does it differ from previous forms of property? How many of our assumptions about it have actually been about its containers rather than their mysterious contents? What are its different species and how does each of them lend itself to control? What technologies will be useful in creating new virtual bottles to replace the old physical ones?

Of course, information is, by its nature, intangible and hard to define. Like other such deep phenomena as light or matter, it is a natural host to paradox. And as it is most helpful to understand light as being both a particle and a wave, an understanding of information may emerge in the abstract congruence of its several different properties that might be described by the following three statements:

- Information is an activity.
- Information is a life form.
- Information is a relationship.

In the following section, I will examine each of these.

Information Is an Activity

Information Is a Verb, Not a Noun

Freed of its containers, information is obviously not a thing. In fact, it is something that happens in the field of interaction between minds or objects or other pieces of information.

Gregory Bateson, expanding on the information theory of Claude Shannon, said, "Information is a difference which makes a difference." Thus, information only really exists in the Δ. The making of that difference is an activity within a relationship. Information is an action that occupies time rather than a state of being which occupies physical space, as is the case with hard goods. It is the pitch, not the baseball, the dance, not the dancer.

Information Is Experienced, Not Possessed

Even when it has been encapsulated in some static form like a book or a hard disk, information is still something that happens to you as you mentally decompress it from its storage code. But, whether it's running at gigabits per second or words per minute, the actual decoding is a process that must be performed by and upon a mind, a process that must take place in time.

There was a cartoon in the *Bulletin of Atomic Scientists* a few years ago which illustrated this point beautifully. In the drawing, a holdup man trains his gun on the sort of bespectacled fellow you'd figure might have a lot of information stored in his head. "Quick," orders the bandit, "Give me all your ideas."

Information Has to Move

Sharks are said to die of suffocation if they stop swimming, and the same is nearly true of information. Information that isn't moving ceases to exist as anything but potential—at least until it is allowed to move again. For this reason, the practice of information hoarding, common in bureaucracies, is an especially wrong-headed artifact of physically based value systems.

Information Is Conveyed by Propagation, Not Distribution

The way in which information spreads is also very different from the distribution of physical goods. It moves more like something from nature than from a factory. It can concatenate like falling dominos or grow in the usual fractal lattice, like frost spreading on a window; but it cannot be shipped around like widgets, except to the extent that it can be contained in them. It doesn't simply move on. It leaves a trail of itself everywhere it's been.

The central economic distinction between information and physical property is the ability of information to be transferred without leaving the possession of the original owner. If I sell you my horse, I can't ride him after that. If I sell you what I know, we both know it.

Information Is a Life Form

Information Wants to Be Free

Stewart Brand is generally credited with this elegant statement of the obvious, recognizing both the natural desire of secrets to be told and the fact that they might be capable of possessing something like a "desire" in the first place.

English biologist and philosopher Richard Dawkins proposed the idea of "memes," self-replicating patterns of information that propagate themselves across the ecologies of mind, saying they were like life forms.

I believe they are life forms in every respect but a basis in the carbon atom. They self-reproduce, they interact with their surroundings and adapt to them, they mutate, they persist. Like any other life form they evolve to fill the possibility spaces of their local environments, which are, in this case, the surrounding belief systems and cultures of their hosts, namely, us.

Indeed, the sociobiologists, like Dawkins, make a plausible case that carbon-based life forms are information as well, and that, as the chicken is an egg's way of making another egg, the entire biological spectacle is just the DNA molecule's means of copying out more information strings exactly like itself.

Information Replicates into the Cracks of Possibility

Like DNA helices, ideas are relentless expansionists, always seeking new opportunities for lebensraum. And, as in carbon-based nature, the more robust organisms are extremely adept at finding new places to live. Thus, just as the common housefly has insinuated itself into practically every ecosystem on the planet, so has the meme of "life after death" found a niche in most minds, or psycho-ecologies.

The more universally resonant an idea or image or song, the more minds it will enter and remain within. Trying to stop the spread of a really robust piece of information is about as easy as keeping killer bees south of the border. The stuff just leaks.

Information Wants to Change

If ideas and other interactive patterns of information are indeed life forms, they can be expected to evolve constantly into forms that will be more perfectly adapted to their surroundings. And, as we see, they are doing this all the time.

But for a long time, our static media, whether carvings in stone, ink on paper, or dye on celluloid, have strongly resisted the evolutionary impulse, exalting as a consequence the author's ability to determine the finished product. But, as in an oral tradition, digitized information has no "final cut."

Digital information, unconstrained by packaging, is a continuing process more like the metamorphosing tales of prehistory than anything that will fit in shrink wrap. From the Neolithic to Gutenberg, information was passed on, mouth to ear, changing with every re-telling (or re-singing). The stories that once shaped our sense of the world didn't have authoritative versions. They adapted to each culture in which they found themselves being told.

Because there was never a moment when the story was frozen in print, the so-called "moral" right of storytellers to keep the tale their own was neither protected nor recognized. The story simply passed through each of them on its way to the next, where it would assume a different form. As we return to continuous information, we can expect the importance of authorship to diminish. Creative people may have to renew their acquaintance with humility.

But our system of copyright makes no accommodation whatever for expressions that don't at some point become "fixed" nor for cultural expressions which lack a specific author or inventor.

Jazz improvisations, standup comedy routines, mime performances, developing monologues, and unrecorded broadcast transmissions all lack the Constitutional requirement of fixation as a "writing." Without being fixed by a point of publication the liquid works of the future will all look more like these continuously adapting and changing forms and will therefore exist beyond the reach of copyright.

Copyright expert Pamela Samuelson tells of having attended a conference last year convened around the fact that Western countries may legally appropriate the music, designs, and biomedical lore of aboriginal people without compensation to their tribe of origin since that tribe is not an "author" or "inventor."

But soon most information will be generated collaboratively by the cyber-tribal hunter-gatherers of Cyberspace. Our arrogant legal dismissal of the rights of "primitives" will be back to haunt us soon.

Information Is Perishable

With the exception of the rare classic, most information is like farm produce. Its quality degrades rapidly both over time and in distance from the source of production. But even here, value is highly subjective and conditional. Yesterday's papers are quite valuable to the historian. In fact, the older they are, the more valuable they become. On the other hand, a commodities broker might consider news of an event that is more than an hour old to have lost any relevance.

Information Is a Relationship

Meaning Has Value and Is Unique to Each Case

In most cases, we assign value to information based on its meaningfulness. The place where information dwells, the holy moment where transmission becomes reception, is a region that has many shifting characteristics and flavors depending on the relationship of sender and receiver, the depth of their interactivity.

Each such relationship is unique. Even in cases where the sender is a broadcast medium, and no response is returned, the receiver is hardly passive. Receiving information is often as creative an act as generating it.

The value of what is sent depends entirely on the extent to which each individual receiver has the receptors—shared terminology, attention, interest, language, paradigm—necessary to render what is received meaningful.

Understanding is a critical element increasingly overlooked in the effort to turn information into a commodity. Data may be any set of facts, useful or not, intelligible or inscrutable, germane or irrelevant. Computers can crank out new data all night long without human help, and the

results may be offered for sale as information. They may or may not actually be so. Only a human being can recognize the meaning that separates information from data.

In fact, information, in the economic sense of the word, consists of data that have been passed through a particular human mind and found meaningful within that mental context. One fella's information is all just data to someone else. If you're an anthropologist, my detailed charts of Tasaday kinship patterns might be critical information to you. If you're a banker from Hong Kong, they might barely seem to be data.

Familiarity Has More Value Than Scarcity

With physical goods, there is a direct correlation between scarcity and value. Gold is more valuable than wheat, even though you can't eat it. While this is not always the case, the situation with information is usually precisely the reverse. Most soft goods increase in value as they become more common. Familiarity is an important asset in the world of information. It may often be the case that the best thing you can do to raise the demand for your product is to give it away.

While this has not always worked with shareware, it could be argued that there is a connection between the extent to which commercial software is pirated and the amount that gets sold. Broadly pirated software, such as Lotus 1-2-3 or WordPerfect, becomes a standard and benefits from the Law of Increasing Returns based on familiarity.

Regarding my own soft product, rock and roll songs, there is no question that the band I write them for, the Grateful Dead, has increased its popularity enormously by giving them away. We have been letting people tape our concerts since the early seventies, but instead of reducing the demand for our product, we are now the largest concert draw in America, a fact that is at least in part attributable to the popularity generated by those tapes.

True, I don't get any royalties on the millions of copies of my songs that have been extracted from concerts, but I see no reason to complain. The fact is, no one but the Grateful Dead can perform a Grateful Dead song, so if you want the experience and not its thin projection, you have to buy a ticket from us. In other words, our intellectual property protection derives from our being the only real-time source of it.

Exclusivity Has Value

The problem with a model that turns the physical scarcity/value ratio on its head is that sometimes the value of information is very much based on its scarcity. Exclusive possession of certain facts makes them more useful. If everyone knows about conditions that might drive a stock price up, the information is valueless.

But again, the critical factor is usually time. It doesn't matter if this kind of information eventually becomes ubiquitous. What matters is being among the first who possess it and act on it. While potent secrets usually don't stay secret, they may remain so long enough to advance the cause of their original holders.

Point of View and Authority Have Value

In a world of floating realities and contradictory maps, rewards will accrue to those commentators whose maps seem to fit their territory snugly, based on their ability to yield predictable results for those who use them.

In aesthetic information, whether poetry or rock 'n' roll, people are willing to buy the new product of an artist, sight-unseen, based on their having been delivered a pleasurable experience by previous work.

Reality is an edit. People are willing to pay for the authority of those editors whose filtering point of view seems to fit best. And again, point of view is an asset that cannot be stolen or duplicated. No one but Esther Dyson sees the world as she does and the handsome fee she charges for her newsletter is actually for the privilege of looking at the world through her unique eyes.

Time Replaces Space

In the physical world, value depends heavily on possession, or proximity in space. One owns that material that falls inside certain dimensional boundaries and the ability to act directly, exclusively, and as one wishes upon what falls inside those boundaries is the principal right of ownership. And of course there is the relationship between value and scarcity, a limitation in space.

In the virtual world, proximity in time is a value determinant. An informational product is generally more valuable the closer the purchaser can place himself to the moment of its expression, a limitation in time.

Many kinds of information degrade rapidly with either time or reproduction. Relevance fades as the territory they map changes. Noise is introduced and bandwidth lost with passage away from the point where the information is first produced.

Thus, listening to a Grateful Dead tape is hardly the same experience as attending a Grateful Dead concert. The closer one can get to the headwaters of an informational stream, the better his chances of finding an accurate picture of reality in it. In an era of easy reproduction, the informational abstractions of popular experiences will propagate out from their source moments to reach anyone who's interested. But it's easy enough to restrict the real experience of the desirable event, whether knock-out punch or guitar lick, to those willing to pay for being there.

The Protection of Execution

In the hick town I come from, they don't give you much credit for just having ideas. You are judged by what you can make of them. As things continue to speed up, I think we see that execution is the best protection for those designs that become physical products. Or, as Steve Jobs once put it, "Real artists ship." The big winner is usually the one who gets to the market first (and with enough organizational force to keep the lead).

But, as we become fixated upon information commerce, many of us seem to think that originality alone is sufficient to convey value, deserving, with the right legal assurances, of a steady wage. In fact, the best way to protect intellectual property is to act on it. It's not enough to invent and patent, one has to innovate as well. Someone claims to have patented the microprocessor before Intel. Maybe so. If he'd actually started shipping microprocessors before Intel, his claim would seem far less spurious.

Information as Its Own Reward

It is now a commonplace to say that money is information. With the exception of Krugerands, crumpled cab-fare, and the contents of those suit-cases which drug lords are reputed to carry, most of the money in the informatized world is in ones and zeros. The global money supply sloshes around the Net, as fluid as weather. It is also obvious, as I have discussed, that information has become as fundamental to the creation of modern wealth as land and sunlight once were.

What is less obvious is the extent to which information is acquiring intrinsic value, not as a means to acquisition but as the object to be acquired. I suppose this has always been less explicitly the case. In politics and academia, potency and information have always been closely related.

However, as we increasingly buy information with money, we begin to see that buying information with other information is simple economic exchange without the necessity of converting the product into and out of currency. This is somewhat challenging for those who like clean accounting, since, information theory aside, informational exchange rates are too squishy to quantify to the decimal point.

Nevertheless, most of what a middle class American purchases has little to do with survival. We buy beauty, prestige, experience, education, and all the obscure pleasures of owning. Many of these things can not only be expressed in non-material terms, they can be acquired by non-material means.

And then there are the inexplicable pleasures of information itself, the joys of learning, knowing, and teaching. The strange good feeling of information coming into and out of oneself. Playing with ideas is a recreation which people must be willing to pay a lot for, given the market for books and elective seminars. We'd likely spend even more money for such pleasures if there weren't so many opportunities to pay for ideas with other ideas.

This explains much of the collective "volunteer" work that fills the archives, newsgroups, and databases of the Internet. Its denizens are not working for nothing, as is widely believed. Rather they are getting paid in something besides money. It is an economy that consists almost entirely of information.

This may become the dominant form of human trade, and if we persist in modeling economics on a strictly monetary basis, we may be gravely misled.

Getting Paid in Cyberspace

How all the foregoing relates to solutions to the crisis in intellectual property is something I've barely started to wrap my mind around. It's fairly paradigm-warping to look at information through fresh eyes—to see how very little it is like pig iron or pork bellies, to imagine the tottering travesties of case law we will stack up if we go on treating it legally as though it were.

As I've said, I believe these towers of outmoded boilerplate will be a smoking heap sometime in the next decade and we mind miners will have no choice but to cast our lot with new systems that work.

I'm not really so gloomy about our prospects as readers of this jeremiad so far might conclude. Solutions will emerge. Nature abhors a vacuum and so does commerce.

Indeed, one of the aspects of the Electronic Frontier that I have always found most appealing—and the reason Mitch Kapor and I used that phrase in naming our foundation—is the degree to which it resembles the 19th century American West in its natural preference for social devices which emerge from it conditions rather than those which are imposed from the outside.

Until the West was fully settled and "civilized" in this century, order was established according to an unwritten Code of the West that had the fluidity of etiquette rather than the rigidity of law. Ethics were more important than rules. Understandings were preferred over laws, which were, in any event, largely unenforceable.

I believe that law, as we understand it, was developed to protect the interests that arose in the two economic "waves" which Alvin Toffler accurately identified in *The Third Wave*. The First Wave was agriculturally based and required law to order ownership of the principal source of production, land. In the Second Wave, manufacturing became the economic mainspring, and the structure of modern law grew around the centralized institutions that needed protection for their reserves of capital, manpower, and hardware.

Both of these economic systems required stability. Their laws were designed to resist change and to assure some equability of distribution within a fairly static social framework. The possibility spaces had to be constrained to preserve the predictability necessary to either land stewardship or capital formation.

In the Third Wave we have now entered, information to a large extent replaces land, capital, and hardware, and as I have detailed in the preceding section, information is most at home in a much more fluid and

adaptable environment. The Third Wave is likely to bring a fundamental shift in the purposes and methods of law that will affect far more than simply those statutes that govern intellectual property.

The "terrain" itself—the architecture of the Net—may come to serve many of the purposes that could only be maintained in the past by legal imposition. For example, it may be unnecessary to constitutionally assure freedom of expression in an environment that, in the words of my fellow EFF co-founder John Gilmore, "treats censorship as a malfunction" and re-routes proscribed ideas around it.

Similar natural balancing mechanisms may arise to smooth over the social discontinuities that previously required legal intercession to set right. On the Net, these differences are more likely to be spanned by a continuous spectrum that connects as much as it separates.

And, despite their fierce grip on the old legal structure, companies which trade in information are likely to find that in their increasing inability to deal sensibly with technological issues, the courts will not produce results that are predictable enough to be supportive of long-term enterprise. Every litigation becomes like a game of Russian roulette, depending on the depth of the presiding judge's clue-impairment.

Uncodified or adaptive "law," while as "fast, loose, and out of control" as other emergent forms, is probably more likely to yield something like justice at this point. In fact, one can already see in development new practices to suit the conditions of virtual commerce. The life forms of information are evolving methods to protect their continued reproduction.

For example, while all the tiny print on a commercial diskette envelope punctiliously requires much of that who would open it, there are, as I say, few who read those provisos, let alone follow them to the letter. And yet, the software business remains a very healthy sector of the American economy.

Why is this? Because people seem to eventually buy the software they really use. Once a program becomes central to your work, you want the latest version of it, the best support, the actual manuals, all privileges that are attached to ownership. Such practical considerations will, in the absence of working law, become more and more important in getting paid for what might easily be obtained for nothing.

I do think that some software is being purchased in the service of ethics or the abstract awareness that the failure to buy it will result in its not being produced any longer, but I'm going to leave those motivators aside. While I believe that the failure of law will almost certainly result in a compensating re-emergence of ethics as the ordering template of society, this is a belief I don't have room to support here.

Instead, I think that, as in the case cited above, compensation for soft products will be driven primarily by practical considerations, all of them consistent with the true properties of digital information, where the value lies in it, and how it can be both manipulated and protected by technology.

While the conundrum remains a conundrum, I can begin to see the directions from which solutions may emerge, based in part on broadening those practical solutions that are already in practice.

Relationship and Its Tools

I believe one idea is central to understanding liquid commerce: Information economics, in the absence of objects, will be based more on relationship than possession.

One existing model for the future conveyance of intellectual property is real time performance, a medium currently used only in theater, music, lectures, stand-up comedy and pedagogy. I believe the concept of performance will expand to include most of the information economy from multi-casted soap operas to stock analysis. In these instances, commercial exchange will be more like ticket sales to a continuous show than the purchase of discrete bundles of that which is being shown.

The other model, of course, is service. The entire professional class—doctors, lawyers, consultants, architects, etc.—are already being paid directly for their intellectual property. Who needs copyright when you're on a retainer?

In fact, this model was applied to much of what is now copyrighted until the late 18th century. Before the industrialization of creation, writers, composers, artists, and the like produced their products in the private service of patrons. Without objects to distribute in a mass market, creative people will return to a condition somewhat like this, except that they will serve many patrons, rather than one.

We can already see the emergence of companies that base their existence on supporting and enhancing the soft property they create rather than selling it by the shrink-wrapped piece or embedding it in widgets.

Trip Hawkins' new company for creating and licensing multimedia tools, 3DO, is an example of what I'm talking about. 3DO doesn't intend to produce any commercial software or consumer devices. Instead, they will act as a kind of private standards setting body, mediating among software and device creators who will be their licensees. They will provide a point of commonality for relationships between a broad spectrum of entities.

In any case, whether you think of yourself as a service provider or a performer, the future protection of your intellectual property will depend on your ability to control your relationship to the market—a relationship that will most likely live and grow over time.

The value of that relationship will reside in the quality of performance, the uniqueness of your point of view, the validity of your expertise, its relevance to your market, and, underlying everything, the ability of that market to access your creative services swiftly, conveniently, and interactively.

Interaction and Protection

Direct interaction will provide a lot of intellectual property protection in the future, and, indeed, it already has. No one knows how many software pirates have bought legitimate copies of a program after calling its publisher for technical support and being asked for some proof of purchase, but I would guess the number is very high.

The same kind of controls will be applicable to "question and answer" relationships between authorities (or artists) and those who seek their expertise. Newsletters, magazines, and books will be supplemented by the ability of their subscribers to ask direct questions of authors.

Interactivity will be a billable commodity even without authorship. As people move into the Net and increasingly get their information directly from its point of production, unfiltered by centralized media, they will attempt to develop the same interactive ability to probe reality which only experience has provided them in the past. Live access to these distant "eyes and ears" will be much easier to cordon than access to static bundles of stored but easily reproducible information.

In most cases, control will be based on restricting access to the freshest, highest bandwidth information. It will be a matter of defining the ticket, the venue, the performer, and the identity of the ticket holder, definitions that I believe will take their forms from technology, not law.

In most cases, the defining technology will be cryptography.

Crypto Bottling

Cryptography, as I've said perhaps too many times, is the "material" from which the walls, boundaries—and bottles—of Cyberspace will be fashioned.

Of course there are problems with cryptography or any other purely technical method of property protection. It has always appeared to me that the more security you hide your goods behind, the more likely you are to turn your sanctuary into a target. Having come from a place where people leave their keys in their cars and don't even have keys to their houses, I remain convinced that the best obstacle to crime is a society with its ethics intact.

While I admit that this is not the kind of society most of us live in, I also believe that a social over-reliance on protection by barricades rather than conscience will eventually wither the latter by turning intrusion and theft into a sport, rather than a crime. This is already occurring in the digital domain as is evident in the activities of computer crackers.

Furthermore, I would argue that initial efforts to protect digital copyright by copy protection contributed to the current condition in which most otherwise ethical computer users seem morally untroubled by their possession of pirated software.

Instead of cultivating among the newly computerized a sense of respect for the work of their fellows, early reliance on copy protection led to the subliminal notion that cracking into a software package somehow "earned" one the right to use it. Limited not by conscience but by technical skill, many soon felt free to do whatever they could get away with. This will continue to be a potential liability of the encryption of digitized commerce.

Furthermore, it's cautionary to remember that copy protection was rejected by the market in most areas. Many of the upcoming efforts to use cryptography-based protection schemes will probably suffer the same

fate. People are not going to tolerate much which makes computers harder to use than they already are without any benefit to the user.

Nevertheless, encryption has already demonstrated a certain blunt utility. New subscriptions to various commercial satellite TV services sky-rocketed recently after their deployment of more robust encryption of their feeds. This, despite a booming backwoods trade in black decoder chips conducted by folks who'd look more at home running moonshine than cracking code.

Another obvious problem with encryption as a global solution is that once something has been unscrambled by a legitimate licensee, it may be openly available to massive reproduction.

In some instances, reproduction following decryption may not be a problem. Many soft products degrade sharply in value with time. It may be that the only real interest in some such products will be among those who have purchased the keys to immediacy.

Furthermore, as software becomes more modular and distribution moves online, it will begin to metamorphose in direct interaction with its user base. Discontinuous upgrades will smooth into a constant process of incremental improvement and adaptation, some of it man-made and some of it arising through genetic algorithms. Pirated copies of software may become too static to have much value to anyone.

Even in cases such as images, where the information is expected to remain fixed, the unencrypted file could still be interwoven with code which could continue to protect it by a wide variety of means.

In most of the schemes I can project, the file would be "alive" with permanently embedded software that could "sense" the surrounding conditions and interact with them. For example, it might contain code that could detect the process of duplication and cause it to self-destruct.

Other methods might give the file the ability to "phone home" through the Net to its original owner. The continued integrity of some files might require periodic "feeding" with digital cash from their host, which they would then relay back to their authors.

Of course files that possess the independent ability to communicate upstream sound uncomfortably like the Morris Internet Worm. "Live" files do have a certain viral quality. And serious privacy issues would arise if everyone's computer were packed with digital spies.

The point is that cryptography will enable a lot of protection technologies that will develop rapidly in the obsessive competition that has always existed between lock-makers and lock-breakers.

But cryptography will not be used simply for making locks. It is also at the heart of both digital signatures and the aforementioned digital cash, both of which I believe will be central to the future protection of intellectual property.

I believe that the generally acknowledged failure of the shareware model in software had less to do with dishonesty than with the simple inconvenience of paying for shareware. If the payment process can be automated, as digital cash and signature will make possible, I believe that soft product creators will reap a much higher return from the bread they cast upon the waters of Cyberspace.

Moreover, they will be spared much of the overhead that presently adheres to the marketing, manufacture, sales, and distribution of information products, whether those products are computer programs, books, CD's, or motion pictures. This will reduce prices and further increase the likelihood of non-compulsory payment.

But of course there is a fundamental problem with a system that requires, through technology, payment for every access to a particular expression. It defeats the original Jeffersonian purpose of seeing that ideas were available to everyone regardless of their economic station. I am not comfortable with a model that will restrict inquiry to the wealthy.

An Economy of Verbs

The future forms and protections of intellectual property are densely obscured from the entrance to the Virtual Age. Nevertheless, I can make (or reiterate) a few flat statements that I earnestly believe won't look too silly in fifty years.

- In the absence of the old containers, almost everything we think we know about intellectual property is wrong. We are going to have to unlearn it. We are going to have to look at information as though we'd never seen the stuff before.
- The protections that we will develop will rely far more on ethics and technology than on law.

• Encryption will be the technical basis for most intellectual property protection. (And should, for this and other reasons, be made more widely available.)
• The economy of the future will be based on relationship rather than possession. It will be continuous rather than sequential.

And finally, in the years to come, most human exchange will be virtual rather than physical, consisting not of stuff but the stuff of which dreams are made. Our future business will be conducted in a world made more of verbs than nouns.

Ojo Caliente, New Mexico, October 1, 1992
New York, New York, November 6, 1992
Brookline, Massachusetts, November 8, 1992
New York, New York, November 15, 1993
San Francisco, California, November 20, 1993
Pinedale, Wyoming, November 24–30, 1993
New York, New York, December 13–14, 1993

This expression has lived and grown to this point over the time period and in the places detailed above. Despite its print publication here, I expect it will continue to evolve in liquid form, possibly for years.

The thoughts in it have not been "mine" alone but have assembled themselves in a field of interaction that has existed between myself and numerous others, to whom I am grateful. They particularly include: Pamela Samuelson, Kevin Kelly, Mitch Kapor, Mike Godwin, Stewart Brand, Mike Holderness, Miriam Barlow, Danny Hillis, Trip Hawkins, and Alvin Toffler.

2

Why Patents Are Bad for Software

Simson L. Garfinkel, Richard M. Stallman,
and Mitchell Kapor

In September 1990, users of the popular XyWrite word processing program got a disturbing letter in the mail from XyQuest, Inc., the program's publisher:

> In June of 1987, we introduced an automatic correction and abbreviation expansion feature in XyWrite III Plus. Unbeknownst to us, a patent application for a related capability had been filed in 1984 and was subsequently granted in 1988. The company holding the patent contacted us in late 1989 and apprised us of the existence of their patent.
>
> We have decided to modify XyWrite III Plus so that it cannot be construed as infringing. The newest version of XyWrite III Plus (3.56) incorporates two significant changes that address this issue: You will no longer be able to automatically correct common spelling errors by pressing the space bar after the misspelled word. In addition, to expand abbreviations stored in your personal dictionary, you will have to press control-R or another designated hot key.

XyQuest had been bitten by a software patent—one of the more than two thousand patents on computer algorithms and software techniques that have been granted by the U.S. Patent and Trademark Office since the mid-1980s. The owner of the patent, Productivity Software, had given XyQuest a choice: license the patent or take a popular feature out of XyWrite, XyQuest's flagship product. If XyQuest refused, a costly patent-infringement lawsuit was sure to follow.

Some choice.

XyQuest tried to license the patent, says Jim Adelson, vice president for marketing, but Productivity Software kept changing its terms. First Productivity said that XyQuest could keep the feature in some versions of XyWrite, but not in others. Then the company said that XyQuest could use one part of the "invention," but not other parts. And Productivity

3

Against Software Patents

The League for Programming Freedom

Software patents threaten to devastate America's computer industry. Patents granted in the past decade are now being used to attack companies such as the Lotus Development Corporation for selling programs they have independently developed. Soon new companies will often be barred from the software arena—most major programs will require licenses for dozens of patents, making them infeasible. This problem has only one solution: software patents must be eliminated.

The Patent System and Computer Programs

The framers of the United States Constitution established the patent system to provide inventors with an incentive to share their inventions with the general public. In exchange for divulging an invention, the patent grants the inventor a 17-year monopoly on its use. The patent holder can license others to use the invention, but may also refuse to do so. Independent reinvention of the same technique by another person does not give that person the right to use it.

Patents do not cover systems. Instead, they cover particular techniques that can be used to build systems, or particular features that systems can offer. Once a technique or feature is patented, it may not be used in a system without the permission of the patent holder—even if it is implemented in a different way. Since a computer program typically uses many techniques and provides many features, it can infringe many patents at once.

Until recently, patents were not used in the software field. Software developers copyrighted individual programs or made them trade secrets.

Copyright was traditionally understood to cover the implementation details of a particular program. It did not cover the features of the program, or the general methods used. And trade secrecy, by definition, could not prohibit any development work by someone who did not know the secret.

On this basis, software development was extremely profitable, and received considerable investment, without any prohibition on independent software development. But it no longer works this way. A change in U.S. government policy in the early 1980s stimulated a flood of applications. Now many have been approved, and the rate is accelerating. Many programmers are unaware of the change and do not appreciate the magnitude of its effects. Today the lawsuits are just beginning.

Absurd Patents

The Patent Office and the courts have had a difficult time with computer software. Until recently the Patent Office refused to hire computer science graduates as examiners, and even now does not offer competitive salaries for the field. Patent examiners are often ill-prepared to evaluate software patent applications to determine if they represent techniques that are widely known or obvious—both of which are grounds for rejection.

Their task is made more difficult because many commonly used software techniques do not appear in the scientific literature of computer science. Some seemed too obvious to publish while others seemed insufficiently general; some were open secrets.

Computer scientists know many techniques that can be generalized to widely varying circumstances. But the Patent Office seems to believe each separate use of a technique is a candidate for a new patent. For example, Apple was sued because the Hypercard program allegedly violates patent number 4,736,308, a patent that covers displaying portions of two or more strings together on the screen—effectively, scrolling with multiple subwindows. Scrolling and subwindows are well-known techniques, but combining them is now apparently illegal.

The granting of a patent by the Patent Office carries a presumption in law that the patent is valid. Patents for well-known techniques that were in use many years before the patent application have been upheld by federal courts. It can be difficult to prove a technique was well known at the time in question.

For example, the technique of using exclusive-or to write a cursor onto a screen is both well known and obvious. (Its advantage is that another identical exclusive-or operation can be used to erase the cursor without damaging the other data on the screen.) This technique can be implemented in a few lines of a program, and a clever high school student might well reinvent it. But it is covered by patent number 4,197,590, which has been upheld twice in court even though the technique was used at least five years before the patent application. Cadtrak, the company that owns this patent, collects millions of dollars from large computer manufacturers.

English patents covering customary graphics techniques, including airbrushing, stenciling, and combining two images under control of a third one, were recently upheld in court, despite the testimony of the pioneers of the field that they had developed these techniques years before. (The corresponding U.S. patents, including 4,633,416 and 4,602,286, have not yet been tested in court, but they probably will be soon.)

All the major developers of spreadsheet programs have been threatened on the basis of patent 4,398,249, covering "natural order recalc"—the recalculation of all the spreadsheet entries that are affected by changes the user makes, rather than recalculation in a fixed order. Currently Lotus alone is being sued, but a victory for the plaintiff in this case would leave the other developers little hope. The League has found prior art that may defeat this patent, but this is not assured.

Nothing protects programmers from accidentally using a technique that is patented, and then being sued for it. Taking an existing program and making it run faster may also make it violate half a dozen patents that have been granted, or are about to be granted.

Even if the Patent Office learns to understand software better, the mistakes it is making now will follow us into the next century, unless Congress or the Supreme Court intervenes to declare these patents void.

However, this is not the entire problem. Computer programming is fundamentally different from the fields the patent system previously covered. Even if the patent system were to operate "as intended" for software, it would still obstruct the industry it is supposed to promote.

What Is Obvious?

The patent system will not grant or uphold patents that are judged to be obvious. However, the system interprets the word "obvious" in a way that might surprise computer programmers. The standard of obviousness developed in other fields is inappropriate for software.

Patent examiners and judges are accustomed to considering even small, incremental changes as deserving new patents. For example, the famous *Polaroid vs. Kodak* case hinged on differences in the number and order of layers of chemicals in a film—differences between the technique Kodak was using and those described by previous, expired patents. The court ruled that these differences were unobvious.

Computer scientists solve problems quickly because the medium of programming is tractable. They are trained to generalize solution principles from one problem to another. One such generalization is that a procedure can be repeated or subdivided. Programmers consider this obvious—but the Patent Office did not think it was obvious when it granted the patent on scrolling multiple strings, described earlier.

Cases such as this cannot be considered errors. The patent system is functioning as it was designed to—but with software, it produces outrageous results.

Patenting What Is Too Obvious to Publish

Sometimes it is possible to patent a technique that is not new precisely because it is obvious—so obvious that no one would have published a paper about it.

For example, computer companies distributing the free X Window System developed by MIT are now being threatened with lawsuits by AT&T over patent number 4,555,775, covering the use of "backing store" in a window system that lets multiple programs have windows. Backing store means that the contents of a window which is temporarily partly hidden are saved in off-screen memory, so they can be restored quickly if the obscuring window disappears.

Early window systems were developed on computers that could not run two programs at once. Since computers had small memories, saving window contents was obviously a waste of scarce memory space. Later, larger multiprocessing computers led to the use of backing store, and to permitting each program to have its own windows. The combination was inevitable.

The technique of backing store was used at MIT in the Lisp Machine System before AT&T applied for a patent. (By coincidence, the Lisp Machine also supported multiprocessing.) The Lisp Machine developers published nothing about backing store at the time, considering it too obvious. It was mentioned when a programmers' manual explained how to turn it on and off.

But this manual was published one week after the AT&T patent application—too late to count as prior art to defeat the patent. So the AT&T patent may stand, and MIT may be forbidden to continue using a method that MIT used before AT&T.

The result is that the dozens of companies and hundreds of thousands of users who accepted the software from MIT on the understanding that it was free are now faced with possible lawsuits. (They are also being threatened with Cadtrak's exclusive-or patent.) The X Window System project was intended to develop a window system that all developers could use freely. This public service goal seems to have been thwarted by patents.

Why Software Is Different

Software systems are much easier to design than hardware systems of the same number of components. For example, a program of 100,000 components might be 50,000 lines long and could be written by two good programmers in a year. The equipment needed for this costs less than $10,000; the only other cost would be the programmer's own living expenses while doing the job. The total investment would be less than $100,000. If done commercially in a large company, it might cost twice that amount. By contrast, an automobile typically contains under 100,000 components; it requires a large team and costs tens of millions of dollars to design.

And software is also much cheaper to manufacture: copies can be made easily on an ordinary workstation costing under $10,000. Producing a complex hardware system often requires a factory costing tens of millions of dollars.

What is the reason for these differences in cost? A hardware system must be designed using real components. They have varying costs; they have limits of operation; they may be sensitive to temperature, vibration or humidity; they may generate noise; they drain power; they may fail either momentarily or permanently. They must be physically assembled in their proper places, and they must be accessible for replacement in case they fail.

Moreover, each of the components in a hardware design is likely to affect the behavior of many others. This greatly complicates the task of determining what a hardware design will do: mathematical modeling may prove wrong when the design is built.

By contrast, a computer program is built from ideal mathematical objects whose behavior is defined, not modeled approximately, by abstract rules. When an if-statement follows a while-statement there is no need to study whether the if-statement will draw power from the while-statement and thereby distort its output, or whether it could overstress the while-statement and make it fail.

Despite being built from simple parts, computer programs are incredibly complex. The program with 100,000 parts is as complex as an automobile, though far easier to design.

While programs cost substantially less to write, market and sell than automobiles, the cost of dealing with the patent system will not be less. The same number of components will, on the average, involve the same number techniques that might be patented.

The Danger of a Lawsuit

Under the current patent system, a software developer who wishes to follow the law must determine which patents a program violates and negotiate with each patent holder a license to use that patent. Licensing may be prohibitively expensive, or even unavailable if the patent is held by a competitor. Even "reasonable" license fees for several patents can add up to make a project infeasible. Alternatively, the developer may wish to avoid using the patent altogether, but there may be no way around it.

License negotiations may be a problem in themselves, as the developers of XyWrite recently learned. This summer they sent the users of

XyWrite a "downgrade," removing a popular feature: the space bar served as a command to correct spelling errors and expand abbreviations. Threatened by the holder of a patent covering this feature, they tried to negotiate a license, but found that the patent holder kept increasing his demands. Eventually they felt compelled to remove the feature of the program.

The worst danger of the patent system is that a developer might find, after releasing a product, that it infringes one or many patents. The resulting lawsuit and legal fees could force even a medium-sized company out of business.

Worst of all, there is no practical way for a software developer to avoid this danger since there is no effective way to find out what patents a system will infringe. There is a way to try to find out—a patent search—but searches are unreliable and in any case too expensive to use for software projects.

Patent Searches Are Prohibitively Expensive

A system with a hundred thousand components can use hundreds of techniques that might already be patented. Since each patent search costs thousands of dollars, searching for all the possible points of danger could easily cost over a million. This is far more than the cost of writing the program.

The costs do not stop there. Patent applications are written by lawyers for lawyers. A programmer reading a patent may not believe that his or her program violates the patent, but a federal court may rule otherwise. It is thus now necessary to involve patent attorneys at every phase of program development.

Yet this only reduces the risk of being sued later—it does not eliminate the risk. Therefore, it is necessary to have a reserve of cash for the eventuality of a lawsuit.

When a company spends millions to design a hardware system, and plans to invest tens of millions to manufacture it, an extra million or two to pay for dealing with the patent system might be bearable. However, for the inexpensive programming project, the same extra cost is prohibitive. Individuals and small companies especially cannot afford these costs. Software patents will put an end to software entrepreneurs.

Patent Searches Are Unreliable

Even if developers could afford patent searches, these are not a reliable method of avoiding the use of patented techniques. This is because patent searches do not reveal pending patent applications (which are kept confidential by the Patent Office). Since it takes several years on the average for a software patent to be granted, this is a serious problem: A developer could begin designing a large program after a patent has been applied for, and release the program before the patent is approved. Only later will the developer learn that distribution of the program is prohibited.

For example, the implementors of the widely used public domain data compression program Compress followed an algorithm obtained from the journal *IEEE Computer*. (This algorithm is also used in several popular programs for microcomputers, including PKZIP.) They and the user community were surprised to learn later that patent number 4,558,302 had been issued to one of the authors of the article. Now Unisys is demanding royalties for using this algorithm. Although the program Compress is still in the public domain, using it means risking a lawsuit.

The Patent Office does not have a workable scheme for classifying software patents. Although patents are most frequently classified by end results, such as "converting iron to steel," many patents cover algorithms whose use in a program is entirely independent of the purpose of the program. For example, a program to analyze human speech might infringe the patent on a speedup in the Fast Fourier Transform; so might a program to perform symbolic algebra (in multiplying large numbers). But the category to search for such a patent would be difficult to predict.

You might think it would be easy to keep a list of the patented software techniques, or even simply remember them. However, managing such a list is nearly impossible. A list compiled in 1989 by lawyers specializing in the field omitted some of the patents mentioned in this column.

Obscure Patents

When you imagine an invention, you probably think of something that could be described in a few words, such as "a flying machine with fixed, curved wings" or "an electrical communicator with a microphone and a speaker." But most patents cover complex detailed processes that have

no simple descriptions—often they are speedups or variants of well-known processes that are themselves complex.

Most of these patents are neither obvious nor brilliant; they are obscure. A capable software designer will "invent" several such improvements in the course of a project. However, there are many avenues for improving a technique, so no single project is likely to find any given one.

For example, IBM has several patents (including patent number 4,656,583) on workmanlike, albeit complex, speedups for well-known computations performed by optimizing compilers, such as register coloring and computing the available expressions.

Patents are also granted on combinations of techniques that are already widely used. One example is IBM patent 4,742,450, which covers "shared copy-on-write segments." This technique allows several programs to share the same piece of memory that represents information in a file. If any program writes a page in the file, that page is replaced by a copy in all of the programs, which continue to share that page with one another but no longer share with the file.

Shared segments and copy-on-write have been used since the 1960s; this particular combination may be new as a specific feature, but is hardly an invention. Nevertheless, the Patent Office thought it merited a patent, which must now be taken into account by the developer of any new operating system.

Obscure patents are like land mines: other developers are more likely to reinvent these techniques than to find out about the patents, and will then be sued. The chance of running into any one of these patents is small, but they are so numerous that you cannot go far without hitting one. Every basic technique has many variations, and a small set of basic techniques can be combined in many ways. The Patent Office has now granted at least 2,000 software patents—no less than 700 in 1989 alone, according to a list compiled by EDS. We can expect the pace to accelerate. In 10 years, programmers will have no choice but to march on blindly and hope they are lucky.

Problems of Patent Licensing

Most large software companies are trying to solve the problem of patents by getting patents of their own. Then they hope to cross-license with the

other large companies that own most of the patents, freeing them to go on as before.

While this approach will allow companies like Microsoft, Apple and IBM to continue in business, it will shut new companies out of the field. A future start-up, with no patents of its own, will be forced to pay whatever price the giants choose to impose. That price might be high: established companies have an interest in excluding future competitors. The recent Lotus lawsuits against Borland and the Santa Cruz Operation (although involving an extended idea of copyright rather than patents) show how this can work.

Even the giants cannot protect themselves with cross-licensing from companies whose only business is to obtain exclusive rights to patents and then threaten to sue. For example, consider the New York-based Refac Technology Development Corporation, representing the owner of the "natural order recalc" patent. Contrary to its name, Refac does not develop anything except lawsuits—it has no business reason to join a cross-licensing compact. Cadtrak, the owner of the exclusive-or patent, is also a litigation company.

Refac is demanding 5% of sales of all major spreadsheet programs. If a future program infringes on 20 such patents—and this is not unlikely, given the complexity of computer programs and the broad applicability of many patents—the combined royalties could exceed 100% of the sales price. (In practice, just a few patents can make a program unprofitable.)

The Fundamental Question

According to the U.S. Constitution, the purpose of patents is to "promote the progress of science and the useful arts." Thus, the basic question at issue is whether software patents, supposedly a method of encouraging software progress, will truly do so, or will retard progress instead.

So far, we have explained the ways in which patents will make ordinary software development difficult. But what of the intended benefits of patents: more invention, and more public disclosure of inventions? To what extent will these actually occur in the field of software?

There will be little benefit to society from software patents because invention in software was already flourishing before such patents existed,

and inventions were normally published in journals for everyone to use. Invention flourished so strongly, in fact, that the same inventions were often found again and again.

In Software, Independent Reinvention Is Commonplace

A patent is an absolute monopoly. Everyone is forbidden to use the patented process, even those who reinvent it independently. This policy implicitly assumes inventions are rare and precious, since only in those circumstances is it beneficial.

The software field is one of constant reinvention. It is sometimes said that programmers throw away more "inventions" each week than other people develop in a year. And the comparative ease of designing large software systems makes it easy for many people to do work in the field. A programmer solves many problems in developing each program. These solutions are likely to be reinvented frequently as other programmers tackle similar problems.

The prevalence of independent reinvention negates the usual purpose of patents. Patents are intended to encourage inventions and, above all, the disclosure of inventions. If a technique will be reinvented frequently, there is no need to encourage more people to invent it. Since some developers will choose to publish it (if publication is merited), there is no point in encouraging a particular inventor to publish it—at the cost of inhibiting use of the technique.

Overemphasis of Inventions

Many analysts of American and Japanese industry have attributed Japanese success in producing quality products to their emphasis on incremental improvements, convenient features and quality rather than noteworthy inventions.

It is especially true in software that success depends primarily on getting the details right. And that is most of the work in developing any useful software system. Inventions are a comparatively unimportant part of the job.

The idea of software patents is thus an example of the mistaken American preoccupation with inventions rather than products. And patents will encourage this mistaken focus, even as they impede the development work that actually produces better software.

Impeding Innovation

By reducing the number of programmers engaged in software development, software patents will actually impede innovation. Much software innovation comes from programmers solving problems while developing software, not from projects whose specific purpose is to make inventions and obtain patents. In other words, these innovations are byproducts of software development.

When patents make development more difficult, and cut down on development projects, they will also cut down on the byproducts of development—new techniques.

Could Patents Ever Be Beneficial?

Although software patents in general are harmful to society as a whole, we do not claim that every software patent is necessarily harmful. Careful study might show that under certain specific and narrow conditions (necessarily excluding the vast majority of cases) it is beneficial to grant software patents.

Nonetheless, the right thing to do now is to eliminate all software patents as soon as possible, before more damage is done. The careful study can come afterward.

Clearly, software patents are not urgently needed by anyone except patent lawyers. Patents did not solve any problems of the prepatent software industry. There was no shortage of invention, and no shortage of investment.

Complete elimination of software patents may not be the ideal solution, but it is close and is a great improvement. Its very simplicity helps avoid a long delay while people argue about details.

If it is ever shown that software patents are beneficial in certain exceptional cases, the law can be changed again at that time—if it is important enough. There is no reason to continue the present catastrophic situation until that day.

Software Patents Are Legally Questionable

It may come as a surprise that the extension of patent law to software is still legally questionable. It rests on an extreme interpretation of a particular 1981 Supreme Court decision, *Diamond vs. Diehr*.[1]

Traditionally, the only kinds of processes that could be patented were those for transforming matter (such as, for transforming iron into steel). Many other activities which we would consider processes were entirely excluded from patents, including business methods, data analysis, and "mental steps." This was called the "subject matter" doctrine.

Diamond vs. Diehr has been interpreted by the Patent Office as a reversal of this doctrine, but the Court did not explicitly reject it. The case concerned a process for curing rubber—a transformation of matter. The issue at hand was whether the use of a computer program in the process was enough to render it unpatentable, and the Court ruled that it was not. The Patent Office took this narrow decision as a green light for unlimited patenting of software techniques, and even for the use of software to perform specific well-known and customary activities.

Most patent lawyers have embraced the change, saying the new boundaries of patents should be defined over decades by a series of expensive court cases. Such a course of action will certainly be good for patent lawyers, but it is unlikely to be good for software developers and users.

One Way to Eliminate Software Patents

We recommend the passage of a law to exclude software from the domain of patents. No matter what patents might exist, they would not cover implementations in software; only implementations in the form of hard-to-design hardware would be covered. An advantage of this method is it would not be necessary to classify patent applications into hardware and software when examining them.

Many have asked how to define software for this purpose—where the line should be drawn. For the purpose of this legislation, software should be defined by the characteristics that make software patents especially harmful:

- Software is built from ideal infallible mathematical components, whose outputs are not affected by the components into which they feed.
- Ideal mathematical components are defined by abstract rules, so that failure of a component is by definition impossible. The behavior of any system built of these components is likewise defined by the consequences of applying the rules step by step to the components.

- Software can be easily and cheaply copied.

Following this criterion, a program to compute prime numbers is a piece of software. A mechanical device designed specifically to perform the same computation is not software, since mechanical components have friction, can interfere with one another's motion, can fail, and must be assembled physically to form a working machine.

Any piece of software needs a hardware platform in order to run. The software operates the features of the hardware in some combination, under a plan. We propose that combining the features in this way can never create infringement. If the hardware alone does not infringe a patent, then using it in a particular fashion under control of a program should not infringe either. In effect, a program is an extension of the programmer's mind, acting as a proxy for the programmer to control the hardware.

Usually the hardware is a general-purpose computer, which implies no particular application. Such hardware cannot infringe any patents except those covering the construction of computers. Our proposal means that, when a user runs such a program on a general-purpose computer, no patents other than those should apply.

The traditional distinction between hardware and software involves a complex of characteristics that used to go hand in hand. Some newer technologies, such as gate arrays and silicon compilers, blur the distinction because they combine characteristics associated with hardware with others associated with software. However, most of these technologies can be classified unambiguously for patent purposes, either as software or as hardware using the preceding criteria. A few gray areas may remain, but these are comparatively small, and need not be an obstacle to solving the problems patents pose for ordinary software development. They will eventually be treated as hardware, as software, or as something in between.

What You Can Do

One way to help eliminate software patents is to join the League for Programming Freedom. The League is a grass-roots organization of programmers and users opposing software patents and interface copyrights.

(The League is not opposed to copyright on individual programs.) Annual dues for individual members are $42.00 for employed professionals, $10.50 for students, and $21.00 for others. We appreciate activists, but members who cannot contribute their time are also welcome. Contact the League at:·

League for Programming Freedom
1 Kendall Square #143
PO Box 9171
Cambridge, MA 02139
or tel. (617) 243-4091;
Email: {league@prep.ai.mit.edu},

In the United States, you may also help by writing to Congress. You can write to your own representatives, but it may be even more effective to write to the subcommittees that consider such issues:·

House Subcommittee on Intellectual Property
2137 Rayburn Bldg.
Wash., DC 20515

Senate Subcommittee on Patents, Trademarks and Copyrights
United States Senate
Wash., DC 20510

You can phone your representatives at (202) 225-3121, or write to them using the following addresses:·

United States Senate
Wash., DC 20510

House of Representatives
Wash., DC 20510

Fighting Patents One by One

Until we succeed in eliminating all patenting of software, we must try to overturn individual software patents. This is very expensive and can solve only a small part of the problem, but that is better than nothing.

Overturning patents in court requires prior art, which may not be easy to find. The League for Programming Freedom will try to serve as a

clearing house for this information, to assist the defendants in software patent suits. This depends on your help. If you know about prior art for any software patent, please send the information to the League.

If you work on software, you can help prevent software patents by refusing to cooperate in applying for them. The details of this may depend on the situation.

Conclusion

Exempting software from the scope of patents will protect software developers from the insupportable cost of patent searches, the wasteful struggle to find a way clear of known patents, and the unavoidable danger of lawsuits.

If nothing is changed, what is now an efficient creative activity will become prohibitively expensive. To picture the effects, imagine if each square of pavement on the sidewalk had an owner, and pedestrians required a license to step on it. Imagine the negotiations necessary to walk an entire block under this system. That is what writing a program will be like if software patents continue. The sparks of creativity and individualism that have driven the computer revolution will be snuffed out.—Prepared by Richard Stallman and Simson Garfinkle.

Note

1. See Samuelson, P. "Legally Speaking." *Commun. ACM* (Aug. 1990).

4

Debunking the Software Patent Myths

Paul Heckel

Jealousy and Envy deny the merit or the novelty of your invention; but vanity, when the novelty and merit are established, claims it for its own. . . . One would not therefore, of all faculties, or qualities of the mind, wish for a friend, or a child, that he should have that of invention. For his attempts to benefit mankind in that way, however well imagined, if they do not succeed, expose him, though very unjustly, to general ridicule and contempt; and if they do succeed, to envy, robbery, and abuse.

—Ben Franklin, 1775 [6]

The issue of software patentability is an important topic because it affects the environment in which programmers and designers work software innovation, the health of the software industry, and U.S. competitiveness. While the writing of this article was motivated by "Against Software Patents," by the League for Programming Freedom in the Jan. 1992 issue of *Communications* it is an overall defense of software patents.

An Absurd Patent

U.S. Patent 4,736,308, the first patent under the heading "Absurd Patents" in "Against Software Patents," is described: "For example, Apple was sued because the HyperCard program allegedly violates patent number 4,736,308, a patent that covers displaying portions of two or more strings together on the screen, effectively scrolling with multiple subwindows. Scrolling and subwindows are well-known techniques, but combining them is apparently illegal." The League calls this an "outrageous result." Based on this description alone, any reasonable person would have to agree.

But I am that inventor and Apple was actually sued on a prior related patent, 4,486,857. Because my patents were misrepresented, I researched the other patents described in the League's article and am reporting my results.

There is much the League did not say about my patent and the circumstances surrounding it. First, it did not describe my background. In 1963 I worked on the software for the first computer designed to be a timesharing computer. I was at Xerox PARC in its early days, wrote two articles for *Communications* [17,20] and a book on user interface design [16]. My patent covers a commercial product called Zoomracks [19], which introduced a new computer metaphor called the card and rack metaphor. Zoomracks was marketed primarily on the Atari ST. Zoomracks developed a strong base of users who used it for a very broad range of applications, but it was a financial struggle largely because Atari did poorly. In Aug. 1987, Apple Computer introduced HyperCard, which is based on a similar, but more limited card and stack version of the metaphor.

I was then faced with having invested six years of raising money, developing a product, marketing it, and proving its value in the market, only to find I was in debt, my customer base was on a dying computer and Apple was giving away *free* a more polished and featured, although less elegant, version of the metaphor. While Apple may not have set out to rip off Zoomracks, it was aware of Zoomracks (having seen it under nondisclosure), of HyperCard's similarity to Zoomracks, and that Zoomracks was protected by patents.

HyperCard created expectations that Zoomracks could not meet, and other companies began to develop HyperCard clones. Meanwhile, I asserted my rights, sued and settled with Apple, licensing the patents. Apple is to be applauded for respecting my patents.

IBM was less respectful: we had twice brought our patent to its attention with respect to products like HyperCard and we had visibly asserted our patents and sued and settled with Apple by the time IBM decided to bundle what many consider to be a HyperCard clone. If this article has an anti-IBM patina to it, it is because I spent six months patiently trying to deal with IBM. Finally, IBM representatives flew to San Francisco to show us prior art—earlier technology—invalidating our patents that they claimed

to have. When they arrived, they refused to show us the prior art, "for fear the patent office would recertify our patents in error." Even if IBM had been straightforward with me during the six months, to accept such an assertion without evidence would have been naive.

Faced with a choice of accepting IBM's offer of 0.2% of the $5 million IBM is said to have paid to license the token ring patent, or to accept its challenge to "sue us" if we wanted to see the prior art, IBM left me no choice but to fight. But I have chosen to fight in the court of public opinion where possible, rather than the civil courts where, because of its financial strength, IBM has the detailed advantage. I added a description of my dealings with IBM to my book [16] and later sent copies of my book to the members of the Commission on Patent Reform when they asked for comments.

Based on my experience I formulated Heckel's Principle of Dealing with Big Companies: *There is no such thing as a free lunch; unless you're the lunch.*

With Apple and IBM, I did battle against large companies who were sophisticated about intellectual property, rather than small ones that were not. I felt it was in everyone's interest to force companies and the courts to make decisions about software patents so the rules and the marketplace realities can be clear to all, not just the sophisticated few. This article is written in that same spirit. While it is a personal issue, I write to clarify the software patent issues in general, to raise the level of discussion and because like most good inventors, I am curious about what the truth is.

One can only understand the need for patents in light of the competitive marketplace. We need a heavy to show what the innovator faces, just as Humphrey Bogart needed Sidney Greenstreet in *The Maltese Falcon.* IBM has already presented itself in that role; it will reappear as did Sidney Greenstreet.

The Informed Opinion

We will visit the other eight patents mentioned by the League in its article and show that the patents it selected, on examination, disprove its case. But first, we take the broader view.

Should software be patentable like other technologies? The primary issue is a policy one and so we have been influenced by Neustadt and May and their book on governmental decision-making [33]. We ask: "What analogies (to software) exist?" "What are the similarities and differences?" "What are the assumptions, explicit and hidden?" "What is known?" "What is the history of the issues?" "What are the interests of the various players?" We will follow the Goldberg rule and ask, not "What is the problem?" but "What is the story?" Most important, we should ask, "How did things turn out in the past?"

History and innovation economics, more than law and computer science, must be the foundation on which to make policy. We have framed 10 points which are, we believe, the consensus of informed opinion on software patents. We hope they help you crystallize your thoughts on patents and enable you to better articulate your differences, if any, with the informed opinion.

1. By creating property rights, patents promote innovation in non-software areas. They particularly promote innovation from small and mid-size companies.

Most of the arguments against software patents turn out to be arguments against patents *per se*. These arguments are advanced most credibly on the basis of established technologies where data and research already exist.

Patents have been accepted around the world as promoting innovation. Many giants of U.S. industry such as G.E., AT&T, Polaroid, Xerox and Hewlett-Packard, started as small companies that used patent protection to protect their inventions.

Yet, most of the articles on patents in the trade, business and even academic press read by the computer community [5, 13, 15, 26, 27, 28, 41, 42, 48] have an antisoftware patent bias. The reason is that for every patent there is one patentholder who is reluctant to speak because the issue is complex and what someone says could be used against him in litigation. And there are a dozen who might like to use the patented technology without paying for it and so are willing to malign the patent and patent system and pass on unsubstantiated rumors and misinformation.

Economists have researched innovation in other technologies [24, 30, 31, 41] and found the following: patents encourage innovation; and small entities—individual inventors and small companies—are a very important source of innovation. According to Jewkes et al. [24],

> It is almost impossible to conceive of any existing social institution so faulty in so many ways. It survives only because there seems to be nothing better. And yet for the individual inventor or the small producer struggling to market a new idea, the patent right is crucially important. It is the only resource he possesses and, fragile and precarious as his rights may be, without them he would have nothing by which to establish a claim to a reward for his work. The sale of his ideas directly or the raising of capital for exploiting the ideas would be hopeless without the patent.

While several articles discuss software patents and copyrights [8, 46, 47], few have been written for the software, as opposed to the legal, community [11, 16, 37]. Such studies have only rarely been linked to software [7], and we are unaware of any empirical studies of the effect of software patents on innovation other than this one.

If we are to reject patents in principle, we should argue that case. If we accept patents as promoting innovation elsewhere but not in software, then we should differentiate software from other technologies.

2. Patents have evolved to address concerns raised by those who suspect software patents.

The courts have developed a patent jurisprudence as a unifying mechanism to support many technologies and foster evolutionary improvement while balancing the rights of patentholders and potential infringers.

Patents have a long history (see a Brief History of Patents). Most of the concerns about patents raised by the League have been raised long ago in the context of other technologies and addressed in case law and legislation and have stood the test of time. The patent system, like MS/DOS, is not perfect. MS/DOS has a long history of evolutionary improvement: It is a derivative of CP/M, which is a derivative of TOPS-20, which is a derivative of the SDS-940 timesharing system, which evolved from the first timesharing system developed at BBN about 1960. Patent jurisprudence has a similar history of evolution.

Part of the value of patents is they are a proven, public domain *standard* of intellectual property protection having a history of improvement over 500 years, compared to the 30 or 40 years of experience developing operating systems. TopView and OS/2 demonstrate how developing a new operating system and crystallizing a new infrastructure around it are fraught

A Brief History of Patents

Until recently patents were thought to have originated in England and been used only there prior to America, an error propagated by Jefferson, Lincoln and as late as 1948 by the Supreme Court. Recent scholarship shows their Italian origin and their early use in France, Germany, the Netherlands as well as England. Venice granted 10-year monopolies to inventors of silk-making devices in the 1200s. These early patents were *ad hoc* grants. In 1474 Venice passed its first patent statute. It recognized patents as a matter of right, rather than royal favor, and provided for fines and the destruction of infringing devices. Gallileo was granted a patent [6]. In England, the queen granted so many monopolies to her friends that citizens protested; so England, in 1624, passed the Statute of Monopolies. This prevented the granting of monopolies, but gave people the right to obtain patents on inventions and imports new to the realm. This statute distinguished between monopolies, which it outlawed as taking from the public what it already had, and patents, which it permitted as giving to the public what it did not yet have.

These patent laws were enacted at the end of the dark ages just before the Renaissance in Italy and the Industrial Revolution in England, suggesting that they *stimulated* innovation.

The Founding Fathers also believed that inventions (and writings) belong to their creators inherently—rather than to the state to be granted at its pleasure. This principle was embodied in the Constitution [1, 6], where Article 1, Section 8 says,

The Congress shall have the power to promote the progress of science and the useful arts by securing for a limited time to authors and inventors the exclusive right to their respective writings and discoveries.

Congress has the power, not to grant rights but to secure inherent rights. This is the principle expressed in the Declaration of Independence that "all men are endowed by their creator with certain unalienable rights." In the Federalist Papers, James Madison, in describing the patent powers, observed that "The public good fully coincides . . . with the claims of individuals."

The creators of the Constitution knew history and understood the ways of men and women, and patents—9 of the 13 colonies granted patents. We should give weight to findings of fact embodied in the Constitution. Two concern patents: Patent rights are inherent rights like freedom of speech; and patents promote innovation.

with dangers—known and unknown. An infrastructure has crystallized around MS/DOS. It includes developers and consultants who know it, books explaining its use, and commercial products based on it. Similarly, an infrastructure has crystallized around the patent system. It includes patent lawyers, case law examples of valid and invalid, infringed and not infringed patents, and books and articles explaining patents to both lawyers and nonlawyers.

3. Patents are not perfect.

There are problems with the patent system. Only that which is not real is perfect. The patent community and the Patent and Trademark Office (PTO) are aware of the problems and have been working to address them. A Commission on Patent Reform is considering improvements such as better examination procedures and automatic publication of patent applications after 18 to 24 months.

If lack of perfection were a reason to get rid of something, no one would survive his or her teenage years. Other industries find patents useful in spite of these problems; software will too.

Patents, it is said, inhibit standards. They do not; they inhibit the expropriation of intellectual property without just compensation in violation of the Fifth Amendment. Where patents exist standards are created in two ways:

- Where people want a standard that infringes a patent, the standards body usually negotiates an agreement whereby the patentholder in return for having his or her technology required as part of a standard, agrees to make a standard license and rate available to all.
- Often standards are agreed to which do not infringe any intellectual property. The QWERTY keyboard and the standard automobile controls (steering wheel, brake and accelerator) demonstrate that patents don't inhibit standards creation. Both public domain standards were developed during the working lifetime of Edison who received 1,100 patents.

4. Software is not inherently different from other technologies in the way innovation or patents work.

Arguments that software is different should be treated critically; you can be sure those same arguments will be used by those who do not believe

that the protections of the Bill of Rights extend to areas where computers and software are used.

Fred Brooks, following Aristotle, suggested the distinction between essence and accident [3], and that distinction has guided our analysis in the issues raised by the League and the academics (see Obviousness: Polaroid vs. Kodak). The question is whether the differences between software are essential or accidental in their encouraging innovation. The League says software is different and so should be protected differently. They present two arguments.

A. Programs are complex.
Why, so they are; but so are airplanes, silicon chips, silicon chip fabrication plants, potato chip plants, oil refineries and many things. But people find the patent system beneficial in these other technologies.

B. Software is cheap to develop compared to other technologies because it is a cottage industry.
Other industries have cottage manufacturers and they deal with patents. Outside of software much invention is a cottage industry; about 5,000 independent inventors belong to the 37 organizations that are members

Obviousness: Polaroid vs. Kodak

As an example of how the patent system has evolved to address the concerns the League raises, consider the Polaroid patent which, they say, describes "differences in the number and order of layers of chemical in a film—differences between the technique Kodak was using and those described by previous expired patents." The League says such differences were obvious. The court held otherwise. Kodak could have avoided infringement by using the order described in the earlier, expired, patent the League refers to. Why would Kodak use the new order described in the later patent rather than the earlier one? Why would Polaroid patent it? Could it have been better? This demonstrates three things to those who must deal with patents. First, an active patent is a territorial warning. Second, technology described in expired patents is in the public domain. Third, patents protect the innovator. Polaroid was the innovator in instant photography. Kodak wanted a share of that market. The major obstacle was Polaroid patents. Kodak tried to get too near the fire. Kodak got burned and paid Polaroid over $900 million.

of the National Congress of Independent Inventors. And most cottage industries do not rely on invention.

We should no more optimize an intellectual property system for cottage developers than we should for Fortune 500 companies.

When one talks about marketing and maintaining commercial software products, the costs are much greater than the estimates made by the League. At the other extreme, IBM is reported to have spent 2.5 *billion* dollars to develop OS/2, including applications.

It is expensive to develop software if the task is to design it from scratch and make it a success in the market; it is cheap if the task is to clone something that already exists or is precisely specified.

Indeed, that clone software is so much cheaper to develop argues for the necessity of patent protection if one wants to stimulate the development of products worth cloning.

Making software nonpatentable or subjecting it to a different form of protection creates practical difficulties, rather like a state seceding from the Union and setting up check points on its border. And if one state secedes, they all can. If each technology has its own *sui generis* (unique) form of protection, we would have to set up boundaries between the different technologies and would need rules for what happens at the boundaries.

This situation occurs in software development. Should programmers be able to define their own conventions or should they conform to the system conventions even where they are not optimal? Do new programmers get to define their own conventions just because they were not involved in the original decision? Aren't programmers expected to abide by the conventions so the code will integrate better and others can maintain it later? Of course, as problems surface, it is foolish to resist all change in conventions just because changes have repercussions. Changes are made, but as part of a deliberative process in which the burden of proof is on those who advocate the changes.

The evolution of the law works the same way: computer law is just another subsystem to be integrated into the fabric of jurisprudence.

The problem of having different conventions in different areas is demonstrated by the Cadtrak patent. It is a hardware, rather than a software, patent. It requires a display device but no software to infringe it. A

computer can be designed so it does not infringe, although a simple program running on it can. This demonstrates that simple software programs can infringe almost pure hardware patents and suggests the difficulty of drawing a legal distinction between hardware and software.

Pamela Samuelson laments that "patent lawyers [do not] claim software-related inventions in a straightforward manner" [39]. Patent lawyers are faced with a Catch 22 situation because of the line between what is patentable and what is not: Write a straightforward patent, and get it rejected as pure software because of *Benson*; write one that is patentable subject matter and it will not be straightforward. Attempting to make software unpatentable will no more prevent practical software patents from issuing and being enforced than prohibition will eliminate alcoholism.

The PTO and those patent lawyers who prosecute software patents have much more experience in the nitty-gritty of protecting software than academics. And the PTO and the courts have more experience weaving new technologies into the fabric of the patent system than the software community has in creating forms of intellectual property protection.

The debate in the academic world is described in The Academic Debate: Considered Opinion and Advocacy. The mainstream view taken by the practicing and academic patent bar and most computer lawyers on one side and the contrarians led by Samuelson, argue against software having the same breadth of protection as other technologies. The League for Programming Freedom has launched an offense in the debate. Like the Battle of the Bulge, it might appear formidable when seen up close. But while it must be treated seriously, it is the last gasp of a dying cause.

The pioneers in each new technology see that technology as new, different, and central, and expect the world to accommodate it. To some extent the world does. But each new technology slowly becomes woven into the tapestry of knowledge encompassing other technologies—each distinctive in its picture—but using the same threads and the same weave.

5. A nonprofit Marxist economic system is not optimal in promoting innovation in software.

This is the paradox one must confront if one argues software patents decrease innovation. The essential difference between Marxism and

The Academic Debate: Considered Opinion and Advocacy

The Mainstream View

Donald Chisum has written the standard reference on patent law [9] and is frequently quoted in judicial decisions. His expertise is in integrating patent decisions into a coherent view of the patent law across mechanical, electrical, chemical and other technologies as they are handed down. His reputation rests on the soundness of his analysis in predicting how courts will rule.

Chisum views software as one of many technologies and says software is almost as patentable as anything else and to the degree that it is not, it should be [8]. He has two concerns. First, the current decisions uphold most software as patentable in a way that forces patent lawyers and the technical community to focus on legal technicalities rather than the technical ones. Second, he questions whether lack of patentability will create underinvestment in software innovation as compared to other technologies.

His approach is similar to Judge Schwarzer's in determining how to calibrate the credibility of possible "junk science" in a courtroom. The question is not "Does this make sense in isolation?" but "How does it fit into an organized body of knowledge?" [44]

The Contrarian's View

The League provides no citations for its position in "Against Software Patents." But a similar article by the same authors [15] has four citations: Mitchell Kapor's congressional testimony and articles by lawyers Pamela Samuelson [39, 41] and Brian Kahin [26] let us see how the antisoftware patent position fits in with "an organized body of knowledge."

While patent lawyers conversant with software are in virtual unanimity that software should be patentable, neither of these lawyers is a patent lawyer.

As background we should consider the legal education most lawyers get. First, very few study patent law and thus are not exposed to the fundamental concept that patent rights are property rights. Second, most law students do take antitrust law where they learn that monopolies are illegal. Later, when they discover patents they often mistakenly see monopolies. Third, law students are taught to be advocates. Their job is not to make a considered opinion, but to present their case as best they can; someone else—the judge, Congress, the public—makes decisions.

This author's opinion is that those who took an antipatent point of view did it as a knee jerk intuitive reaction. Away from the mainline of software business and patent law, distrustful of monopolies and inexperienced with

intellectual property, these lawyers and software developers found support in one another. This view gained respectability as a contrarian's view and is always welcome in legal journals and at conferences.

Samuelson argues that the basis of software patentability is weak and says the primary issue is a policy one [39]. She presents two arguments against software patentability and Kahin a third:

A. *The industry has flourished and innovated without patents.*

A more accurate statement would be: "The industry has flourished and innovated *with little realization in the software community* that patent protection was available."

The growth of the software industry was not due to differences between software and other technologies, but to its synergetic relationship with the computer industry and the new business opportunities created by the computer's rapidly decreasing costs. Until recently software has been seen by computer companies as a loss leader. Software created the demand for computers which was, and still is, the dominant industry: But software developers were to be kept fat, dumb and happy; salaries were high, patents were not mentioned, and there was lots of technology to play with.

Accidental Empires shows that the personal computer software successes were achieved by amateurs who were lucky enough to be in the right place at the right time [12]. The successful early PC software companies (a) marketed innovations *pioneered by others* and (b) aggressively pursued their own intellectual property rights: Microsoft's MS/DOS is a derivative of CP/M; and Lotus' 1-2-3 of VisiCalc. Had Digital Research, VisiCorp and Software Arts asserted intellectual property rights as aggressively as Microsoft and Lotus, they might not have been eclipsed by them.

In the early stage of the industrial life cycle, the first person in his garage who acts on the opportunity starts with the biggest and most established company in the business.

When there is no established competition, new companies can compete without patents. As the industry matures it becomes difficult for new companies to enter the market without a sustainable advantage such as patents. The free enterprise system rewarded entrepreneurs whose personal computer software companies were successful, as it should, but that success should not blind us to its nature—bringing the innovations of others to market. We expect that if patents had been more widely used by the software industry, the true innovators would have received a fairer share of the rewards, thus rewarding innovation as well as business savvy. Had patents been more widespread, the software industry would have been more profitable, it would have grown with less of a boom-and-bust cycle, albeit less rapidly and there would have been a greater diversity of software product categories and features. To base a software intellectual property system just on the experience of the last 10 years would be like raising teenagers in the expectation that their childhood will be repeated.

B. *Many in the software industry say software patents will discourage innovation.*

Samuelson says,

> . . . It is primarily from the widespread concerns about the effects of patents from within the industry and the technical community that she has pursued this study questioning the patent protection for computer program-related inventions. [39]

Samuelson addresses this perception by conducting a survey to see how widespread the perception is, at least in the related area of user interface copyrights [42], rather than attempting to determine if the perception reflects reality. In reporting the perception she gives it more credence, reinforcing any error, and creating more concern. It would seem to be more constructive to research the effect of patents on innovation and business formation in other industries to see how it might apply to software, and analyze software patents that exist for their effect on innovation and new business formation as we do.

It makes as much sense to devise theories of innovation by polling the software community as it does to devise laws of physics by polling people who walk. The intuitive answer is not necessarily the correct one. Ask people, "Assume you are walking at a steady pace holding a ball, and you drop the ball. Will the ball hit the ground in front of you, in back of you, or next to you?" most people will say "In back of me" [32]. A physicist will tell you that Newton's laws predict, "next to me," and can perform an experiment to prove it.

In contrast to Chisum and the mainstream of patent law where software is viewed from a broad perspective of many technologies, Samuelson views the law from inside the software community looking out. She has always advocated narrow protection—arguing as late as 1984 that CONTU's recommendation that software in machine readable form be copyrightable was ill considered [40].

Samuelson acknowledges that the conditions that promoted software innovation until now may be different from those that will promote it in the future.

Samuelson proposes to design a special *(sui generis)* form of protection for software, saying, "It is possible to design a law that is appropriate to the kind of subject matter that software is." How can one be certain of one's ability to build a skyscraper, if one is uncertain if one is building on bedrock or marshland?

Writing laws is like designing systems. They will have desired and undesired, intended and unintended consequences. The current law, especially *Benson,* forces patent lawyers into a Catch 22 dilemma: Write a straightforward patent and get it rejected as a mathematical algorithm; write one that is patentable subject matter and it will not be straightforward. Attempting to make only software unpatentable will no more prevent practical software patents from issuing and being enforced than prohibition will eliminate alcoholism. Samuelson laments that "patent lawyers [do not] claim software-related inventions in a straightforward manner" [39]; but this is like blaming a program's users or the market for not behaving as expected—a sure sign of an inexperienced designer. Law, software, and marketing strategies must be designed the same way the Constitution was: based not on how we would like people to behave, but on how self-interested people will actually behave.

Academics such as Samuelson, who pontificate on software patents and want to create a *sui generis* system of protection, seem ready to reinvent the wheel before they understand how a wagon works or the infrastructure of the highway system. Most show no knowledge or understanding of the processes of innovation over hundreds of years in a variety of technologies. Most have little, if any, experience prosecuting (filing), analyzing, or litigating patents. They argue that software is different from other technologies, not from the perspective of the history of innovation and patents over a range of technologies, but from the myopic view of the development of a single technology, largely in a single software marketplace in an atypical decade—the personal computer market in the 1980s.

Ben Franklin described a professor who was so learned he knew the word for horse in five languages: *equus* in Latin, *caballo* in Spanish, *cheval* in French, *cavallo* in Italian, and *Pferd* in German. He then went out to purchase one, and returned with a cow. Given your experience observing complex software systems develop and how the hype manifests itself in reality, when you hear *sui generis* protection systems being proposed as transport into the twenty-first century, ask yourself: "Will I ride a horse or a cow?"

Taking a contrary position and fighting for it, as Samuelson does, is in the highest tradition of the law. In part it is because truth emerges from the debate and if there is no one to debate with, a poor sort of truth will emerge. Yesterday's contrarian view may emerge as the mainstream view. It is like having advocates for a programming language like Forth. It forces identification and discussion of issues and influences the industry—PostScript is Forth-like. But software developers seriously consider programming in Forth only in unusual circumstances.

C. Software helps disseminate information

Brian Kahin, a research fellow in the Science, Technology and Public Policy Program at Harvard University's Kennedy School of Government offers a different reason software should not be patentable [25]:

A deeper, more disturbing problem in patenting programs was barely evident before computers became ubiquitous personal tools. . . . The computer has developed into a medium for human expression and a mediator of human experience. Thus, what is increasingly at stake in software patents is the generation and flow of information.

The "barely evident" "problem" was addressed by Lincoln who, in his *Lecture on Discoveries* [23], said:

certain inventions and discoveries occurred of particular value on account of their efficiency in facilitating all other inventions and discoveries. Of these were the arts of writing and of printing—the discovery of America and the introduction of Patent-laws.

Lincoln not only believed that the patent laws encouraged innovation, but he anticipated Kahin in realizing that inventions which promote dissemination of information are particularly important. Since Lincoln's speech, patents seem to have encouraged many inventions which engender the "generation and flow of information": the telephone (Bell), phonograph (Edison), movie camera (Edison), xerography (Carlson), radio (Armstrong and Marconi), phototypesetting (Scheffer), and TV (Philo T. Farnsworth).

Even at the birth of patenting, Aldus Manutius, the famous Venetian scholar-printer for whom the desktop publishing company is named, received two patents—one on a form of Greek type—and a ten year monopoly to use the Italic font which he invented.

Having discovered that he is treading in the footsteps of Lincoln, albeit in the opposite direction, Kahin might, in planning his future travel, consult the work of his colleagues at the Kennedy School of Government, which has guided our analysis [33].

capitalism is property rights. Patents and copyrights create intellectual property rights that can be bought, sold, rented and licensed like other property rights. Marxism may be better than capitalism in some areas—certainly not in Russia, but capitalism, with all its flaws, has outperformed Marxism.

The paradox of Marxism is not just a theoretical issue. Stallman, the founder of the *League for Programming Freedom*, heads the Free Software Foundation which is developing and planning to distribute a clone

of the Unix™ operating system. AT&T has invested in Unix based on its ownership as manifest in patents and copyrights. AT&T cannot be pleased when Stallman gives away free copies of a clone of a product it invested millions in developing and marketing.

If AT&T had not used patents and user interface copyrights to protect its intellectual property rights, Stallman would have no trouble making and distributing a Unix clone. But AT&T must pay its bills with money it receives from customers and has asserted its rights. If it is acceptable to clone Unix or any program, will anyone invest in new ideas? Should we optimize an intellectual property jurisprudence for, not large entities, not small entities, but companies that distribute free clones of other people's software?

For all his talk about wanting to promote innovation, Stallman seems to get his ideas for technology from AT&T, 1969, and his ideas for intellectual property protection from IBM, 1965.

Many software developers do their work for the fun of it. But the distinction is based, not on the technology, but on amateurism: amateurs flourish in the early stages of a new technology. Professionals have accepted that others work for free, but they bristle if they are expected to work at the same rates.

6. Software, like every technology, has unique problems.

Software patents have unique problems: prior art libraries are limited, the search classification system was designed for hardware patents, few computer scientists are examiners. Still when it gets to specific cases, computer scientists and the PTO see invention similarly (see Document Comparison).

For the last two years the PTO has been improving the situation. It is improving its prior art search facilities in software, has published a new software classification system, and is actively recruiting computer scientists.

The PTO has still not been able to rid itself of the prejudice against software patents, as patent practitioners in the software area will tell you. It still is conservative in its interpretation of what constitutes patentable subject matter, and has rejected several applications that are being appealed.

Document Comparison: An Obvious Patent?

A criticism of the patent system is that computer scientists are qualified to judge invention in software while the PTO is not. In his article [26] Kahin says, the PTO is,

awarding patents merely for automating familiar processes such as . . . comparing documents (Patent No. 4,807,182). But software developers have been routinely automating such [functions] for years.

In fact, the ACM published a refereed paper describing that (hashcoding) technique for comparing two text files albeit for source code, rather than document comparison [17]. It seems that the ACM and the PTO have similar standards of inventiveness.

It so happens I wrote that paper, and I brought it to the attention of the patentholder and (indirectly) to WordPerfect about 10 months before Kahin's article was published. Four companies put on notice about the patent brought the paper to the attention of the patentholder. This suggests that where prior art exists which narrows a patent's scope, it is likely to surface.

Advanced Software was founded to develop and market DocuComp which uses the patented technology. Its inventor, Cary Queen, a Ph.D. in mathematics, has filed over a dozen patents in genetic engineering where he is principle in a startup, Protein Design Labs. He also used the patented hashcoding technique to compare genes to identify similarities.

Cary Queen reports there is more prejudice against patents in software than in biotechnology, where hundreds of startups have been financed, as biotechnology patents are better respected.

7. Legally, software is patentable and it will remain so.

Prior to 1982, about 30 different software-related patent cases went through the Appellate Courts. The range of technologies—seismic, medical, petrochemical, telecommunications, firmware, and software—demonstrate that software is both well grounded in patent law, and basic to the advancement of American industry. Software has become pervasive in industry, which has been basing business decisions on software's being patentable for 10 to 20 years. This has created a sophisticated broad-based constituency for keeping software patentable. Congress has not given in to demands to make less pervasive technologies, such as biotechnology, unpatentable; it is less likely to do so with software. Software has

been clearly patentable longer than it has been copyrightable (see Patents and Copyrights).

Chisum, the leading authority on patents, wrote an article on the patentability of software and concluded:

> The continuing confusion over the patentability of computer programming ideas can be laid on the doorsteps of a single Supreme Court decision, *Gottschalk vs. Benson*, which held that mathematical algorithms cannot be patented, no matter how new and useful. A careful analysis of that decision shows the holding is not supported by any of the authorities on which it relied, that the Court misunderstood the nature of the subject matter before it, and that the Court failed to offer any viable policy justification for excluding by judicial fiat mathematical algorithms from the patent system. The Benson decision is inconsistent with the later Supreme Court ruling in *Diamond vs. Chakabarty* that the patent system applies impartially to new technologies and that any policy issues for excluding new technologies should be addressed to Congress. Policy considerations indicate that patent protection is appropriate for mathematical algorithms that are useful for computer programming as for other technological innovations [8].

Chisum is in the mainstream in saying that the courts made a mistake by making software unpatentable. But courts are reluctant to overturn previous decisions directly, and then only after their scope has been eroded. A similar situation where prejudice had been part of the

Patents and Copyrights

Kahin says, "Never before has an industry in which copyright was widely established suddenly been subjected to patenting" [25].

In fact, patent protection for software was established before copyright protection. *Diamond vs. Diehr* (1981) preceded the two most important software copyright cases, *Apple Computer vs. Franklin Computer Company* (1982) and *SAS Institute Inc. vs. S&H Computer Systems Inc.* (1985). Even today the case law on user interface copyrights is sparse. Software, or "computer-related" patents, were obtained in the 1960s. Martin Goetz of Applied Data Research received U.S. patents 3,380,029 in 1968 on a Sorting System, and 3,533,086 in 1970 on AutoFlow, an automatic flow charting program. (When IBM started giving Roscoe, a flowcharting program, away free, Applied Data Research sued for antitrust and, in 1969, settled for about $2 million in damages.) The recalculation patent filed in 1970 was granted in 1983.

jurisprudence occurred earlier: *Plessy vs. Furgesson* [1896] held that "separate but equal" facilities for whites and blacks were lawful. The courts did not directly overturn it, but eroded its vitality on a case-by-case basis over a period of years in a number of decisions starting with *Murrey vs. Maryland* while appearing to show respect for *Plessy*. Finally, when faced with *Brown vs. Board of Education* [1954], ample precedent had been created for the Supreme Court to overrule *Plessy* directly.

Samuelson, having asked Chisum to write his article, now attempted to refute him, but after arguing for over 100 pages that the basis for software patentability is weak, was forced to conclude that:

> . . . the only principle which seems to have guided the court's decision is one of upholding the patentability of as many program-related inventions as possible while appearing to show respect for the Supreme Court's decisions. [39]

Samuelson's observation seems to be compelling evidence that while she has not been persuaded that *Benson*, like *Plessy*, is a fundamental error in giving prejudice the force of law, the Court has and will, in due course, reverse it and the original intent of Congress will again become law and statutory. Subject matter will "include anything under the sun that is made by man."

8. Whether or not one agrees that software patents are beneficial, patents are here to stay so we should plan to work with them.

The software community will be best served by articles about how to avoid infringement, how to deal with infringement notices, how to find prior art, how to use patents to protect new ideas, how to differentiate products, and how to make the patent system work better for software (based on experience rather than speculation). In brief, we should direct our energies toward making the system work in order to increase innovation and U.S. competitiveness, rather than fighting patents.

9. The practical effect of continuing to spread misinformation on software patents will be to hurt small developers and U.S. competitiveness in software.

Patents, like a cat's claws, function as weapons when necessary. A declawed cat will not survive in the wild; neither can a defenseless startup once it succeeds and attracts substantial competitors. Patents are not the

only defense, but they are vital to innovative startups that must survive. In business, as in the jungle, respect is given only to those who can protect themselves.

Microsoft, IBM and others are applying for patents in quantity. Those who do not understand the situation are not. Many are happy to have software patents attacked. Why let your competitor in on a good thing? Why pioneer new product ideas when it is less risky to copy competing products and incorporate useful features once market success is proven.

From the perspective of large companies, a loud voice, such as the League, yelling against software patents can be useful as a means to destroy one's competition.

The Japanese are aggressively filing for U.S. patents on software. While our strength is innovation, Japan's is in adapting innovations and steady improvement. But if they have the improvement patents and we did not file for the basic patents, we lose. If we arrogantly dismiss the Japanese as incapable of creating good software or cavalierly dismiss patents as undesirable, then 20 years from now we will be trying to get back the software market from Japan just as today we are trying to get back the automobile and semiconductor markets. We are not even trying to get back the consumer electronics market.

Who is responsible for the misperception about the desirability and legality of software patents? In a certain sense, it is the League for Programming Freedom. But it knows not what it does. And its arguments are the ghosts of arguments for IBM's corporate self-interest of a bygone era. Is not the origin of the problem IBM's attempt in the 1960s to declaw a competing technology by depriving its practitioners of their constitutional rights as inventors? (See Software Patents: IBM's Role in History.)

If it were just a question of IBM outfoxing its competitors, we might learn our lesson and let it pass. But we think it useful to ask some questions: Is it in the interest of the United States to have a strong, competitive, innovative software industry? Is it in IBM's interest? Did IBM use its position on the 1966 Patent Commission to put its corporate self-interest ahead of that of the U.S.? Should IBM be held responsible for its role in creating the current software patent mess? Some have proposed making software patents unenforceable. Might a law making IBM patents unenforceable make more sense? Or a law that would prevent IBM from obtaining patents for a period of time, say 5 or 10 years? At a time when competitiveness with Japan is a major concern, what kind of a message should we send about what happens to those who use their positions on government commissions to sacrifice their country's interest to their corporate self-interest?

Similarly, should we eliminate patents to avoid patent litigation as the League suggests? Should we not eliminate all laws so as to avoid all litigation?

Software Patents: IBM's Role in History

In the late 1960s when IBM's internal policy was that software should not be patentable, IBM vice president, J. W. Birkenstock, chaired a presidential commission on the patent system which recommended that software should not be patentable. We expect that the other commission members deferred to IBM's expertise on software, just as members of a commission designing an aviary would defer to its most knowledgeable member on birds: the cat.

Congress rejected this view, but three paragraphs of the Commission's recommendations, e.g., IBM's corporate policy, found their way into *Gottschalk vs. Benson,* the Supreme Court decision that limited the patentability of software. At this time IBM had 70% of the computer market, so it is not surprising that CBEMA, the Computer Business Equipment Manufacturers Association, filed an *amicus curie* brief against software patents in *Benson.*

From this historical perspective we can see that the conventional wisdom that "software has not been patentable," should be more accurately stated as "it was not in the interest of IBM or other computer manufacturers for people to think software is patentable." We have never seen it pointed out in the debate on software patents that the idea that software is not patentable subject matter was formed in the crucible of IBM's self-interest and corporate policies of an earlier time.

IBM and CBEMA have now rejected the Stallman-primal IBM view [7]. But the damage has been done. The PTO and the industry have not taken software patents seriously until recently, which explains the problems the PTO has had in examining patents and the prejudice against software inventors who assert their patent rights. Many in the software community have been suckered into believing software should not be patentable, while IBM has aggressively but quietly been getting software patents and become the company with the largest software sales.

10. In considering the issues, we should deal with examples of real patents and, where possible, real infringement where facts for both sides are fairly stated.

If we are to have a meaningful debate on whether software should be patentable, I suggest we take our standards, both of debate and of where the burden of proof lies from Abraham Lincoln:

> I do not mean to say we are bound to follow implicitly in whatever our fathers did. To do so would be to discard all the lights of current experience—to reject all progress—all improvement . . . if we would supplant the opinions and policy of our fathers in any case, we should do so upon evidence so conclusive, and argument so clear, that even their great authority, fairly considered and weighed, cannot stand. . . .
>
> If any man [believes something], he is right to say so, and to enforce his position by all truthful evidence and fair argument which he can. But he has no right to mislead others, who have less access to history, and less leisure to study it, into [a] false belief . . . thus substituting falsehood and deception for truthful evidence and fair argument.

What I find most frustrating in this debate is that the mode of argument used against software patents by so many [15, 26, 27, 28] is to throw as much mud against the wall as possible and hope some of it will stick. I have expended some effort here removing some of the mud. I do not claim to have removed it all, but I hope I have wiped away enough to show you the rest will wash off too.

A Study of Nine Software Patents

In its article the League lists nine patents—mine and eight others to make its case. It is unlikely that the members of the League considered the positive side of any of the patents they cited. It is as if they went searching for quarters with heads showing, and finding several, reported their findings without turning any of them over. Here we turn over the other eight quarters in an attempt to produce some empirical results.

U.S. Patent 4,197,590

The inventor founded a company to develop and market what appears to be the first personal computer to write directly from memory to the display. This invention has been widely licensed to the personal computer industry by Cadtrak. The "XOR" is only part of the invention. Cadtrak filed and has won at least one lawsuit against a larger company. The idea behind the "XOR" claims of this patent is simple. A program XORs a cursor icon onto a display device; later a second XOR to the same place erases the cursor, restoring the original display. To move the cursor one XORs the cursor to its old location, then XORs it onto the new location. There are many ways to get around this patent. One can use an underline as a cursor or "logically or" the cursor onto the display, erasing later by rewriting the display with its original information. This approach is fast, lets you change cursor icons easily, accesses the minimum possible data, and requires no space be reserved on the screen for the cursor.

The League says this patent can be infringed in "a few lines of a program." It can be, *but not on a computer that was commercially available at the time the invention was made.* The invention is largely the invention of the frame buffer. As such, it requires hardware which has since become common, making it possible to infringe the XOR claims with a few lines of code. Many, if not most, computer manufacturers including Apple and IBM have taken out licenses which cover programs running on their computers.

This patent illustrates that it is usually easy to design around a patent one accidentally infringes. If this patent, a hardware patent, is a "bad" patent as some claim, it only demonstrates that the electronics industry tolerates "bad" patents because it finds patents beneficial on balance. Software should be able to tolerate "bad" patents similarly. To discard the patent system because some bad patents exist would be the same as suppressing free speech to stamp out lies.

U.S. Patent 4,398,249

This is what the League has mischaracterized as fact, in "Refrac recalculation patent." In 1970 Rene Pardo and Remy Landau invented the concept of an array of formulas that would enable businesspeople to write their own programs to create business applications. Although the word was not used, their invention is in essence the modern computer spreadsheet. The fact that the claims cover recalculation is an artifact of how the patent claims were written. Pardo and Landau marketed a

commercial spreadsheetlike product based on this technology. This invention has been widely adapted in the personal computer industry—over 250 spreadsheets have been marketed.

This patent was originally rejected by the PTO as a mathematical algorithm and thus unpatentable subject matter. Pardo and Landau felt so strongly about their inventive contribution that they appealed their case *pro se*, which means they, not a lawyer, wrote the brief and argued it before the appeals court. The decision, *In re Pardo* [23], is a major legal precedent which establishes that an invention is patentable whether or not the invention involves software "novelty." If their experience was typical, they were stonewalled when they tried to enforce their patents. This would explain why they approached Refac—a white knight in the fight against the patent pirates. If Pardo and Landau have the same deal Refac offered others, then they can expect to collect royalties only as Refac does.

U.S. Patent 4,633,416

This patent is held by Quantel, a company that developed a line of commercial video editing products protected by its patents. Quantel filed for patents when it was small, and it has grown from being a small to a large company because it has used its patents to prevent competitors from using its technology. The first company Quantel sued was much larger than it was.

U.S. Patent 4,777,596

The League tells us XyQuest was notified that its product XyWrite infringed Productivity Software's patent, protecting the ability to accept an abbreviation or correct a spelling error by hitting a space bar. When licensing negotiations failed, XyQuest removed the feature from future releases.

Productivity Software was founded in 1984 to develop data input systems where minimal keystroke data input is important. Based on its patented technology, Productivity Software has grown to seven employees and markets 31 specialty products. It has found niches in the medical and legal transcription and the handicapped marketplaces.

Patent problems are generally minor compared to the other problems pioneering companies face. At about the same time XyQuest had

How Patents Work

Exclusive or territorial rights bestow on their owners a long-term outlook, and create a simple test for determining whether or not to fight. This leads to stable solutions and minimizes inefficient disputes. Such rights occur in many areas. A miner stakes a claim, salespeople have exclusive territories, and professors specialize in areas and are given tenure. Such territorial rights stimulate diversity by encouraging competitors to stake their territorial claims at a distance. Lee De Forest, for example, invented the triode, the amplifying vacuum tube, to avoid infringing Fessenden's spade detector patent [29].

Many mistakenly believe that the patent system protects only "flash of genius" insights. That is not true. In 1952, Congress overrode the "flash of genius" doctrine. Patents are designed, not to stimulate invention directly, but to stimulate their commercialization by giving exclusive rights for 17 years to anyone who invents something new and not obvious. Just as an author need meet a standard of creativity to get a copyright on an original work, a patentholder need meet a standard of nonobviousness to get a patent on something new.

The following discussion of patents is useful in providing an overall understanding of how patent infringement is determined, especially where an overly broad (e.g., "bad") patent may be involved.

Before the PTO will allow your patent application, it does a search (rather like a title search when you buy a house) to find prior art on your invention—what others have done earlier that is disclosed in publications or products. The PTO examines the prior art it finds along with any you send it. If what you did is sufficiently different, it issues claims that delimit the territory of your invention.

The process is rather like finding new territory. Suppose you suddenly landed in Left Fork, North Dakota, and found that no one lived there and wanted to claim it as yours. You might try to claim all the land west of the Mississippi. The PTO will likely find that people have lived in nearby states and may issue you a claim to say, North Dakota. Of course, the PTO could allow your claim in error. The too broad claim—all the land west of the Mississippi—will look impressive and could be useful as a source of cash from people impressed by surface rather than substance.

In practice, if you try to enforce the patent against, say, Californians who just discovered gold, they would show the court that people lived in California before you landed in Left Fork. The court will declare your broad claim invalid.

Practically speaking, you would not go to court once you realized that people lived in California earlier. You might go back to the PTO to get the patent reissued, showing them the prior art and claiming a smaller territory.

The PTO might only allow narrower claims that cover eastern North Dakota, or maybe only Left Fork, North Dakota.

A patent does not necessarily give you rights to what it says it does. Undiscovered prior art might considerably narrow its scope. The advantages of being issued a too broad patent are (a) potential infringers might keep a greater distance than they have to, and (b) you can wait to define the limits of your territory until you know the terrain better. The disadvantage is that you might make business decisions based on your belief that you had rights you did not possess.

There are two ways you can respect a patent: you can avoid infringement or you can take out a license. If you are infringing, the patentholder will usually forgive past infringement if you agree to remove infringing capability. Your show of respect for the patent gives its holder credibility with other infringers.

As a possible infringer you have several courses of action when you face an overly broad (or "bad") patent, or indeed any patent:

1. *Ignore the patent problem until confronted with it.*

Why look for trouble you might never have to face? When and if a patent is brought to your attention you can decide what to do. If you do a search and find a patent, you might spend effort designing around a patent that its owner would never have asserted against you. If you do not design around it, you might be liable for treble damages because you were aware of the patent. Of course if you are competing against products protected by patents, you might want to check into their patents before you design your product, as you can expect your competitor to examine your product for infringement, thus you will probably have to face the problem one way or another. Here, it is good accounting practice to set aside a reserve for infringement.

2. *Stay outside the claimed territory.*

If the patent claims all the land west of the Mississippi and you stay on the east of the Mississippi, you will not infringe.

3. *Go where people were.*

If you know people were in Bismark before the patentholder landed in Left Fork, settling in Bismark will protect you. Distrust rumors about earlier settlers. Make sure the prior art is documented in a published paper or was obviously used in a product. If you ask around the industry you are likely to find pointers to prior art. You might want to send the prior art to the patent owner, or the PTO for insertion in the file wrapper. The file wrapper is a file containing all the correspondence on the patent with the PTO. Your patent lawyer might consult it to find prior art which might help you design around a patent or understand its scope.

4. *Make a business deal with the patentholder.*

Generally you can license, or cross-license a patent or find some other way to get rights.

5. *Break the patent.*

You can attempt to get the patent invalidated by proving it is invalid over the prior art, the disclosure was inadequate or it was otherwise invalid. This is risky and expensive where the patent is good and the patentholder determined.

A patent must claim something new, lest its owner usurp others' rights; it also must be on something nonobvious to prevent giving protection to insignificant improvements.

As technical people, we often look at a patent differently from the way entrepreneurs and judges do. We see Left Fork, North Dakota, after it has become a thriving town, and are likely to say it is obvious—there are lots of places like Left Fork, and lots of them have similar buildings; thus constructing buildings in Left Fork seems obvious. It does not belong to the entrepreneur. Invite everyone!

The Left Fork patentholder, the entrepreneur, feels this is like arguing that it was obvious that land in Silicon Valley or Microsoft stock would appreciate in value. Given the advantage of hindsight, it is obvious, but the person who invested in the land, the stock, or the technology should benefit from its appreciation in value. The entrepreneur says, "I built the buildings based on my being granted rights to them and my having a vision of what I could make of it. And now you are looking for loopholes in my deed so others can move in! It may be a poor thing, but it is my own."

As technical people our immediate bias is to find inventions "obvious," because we focus on the technical sophistication, evaluated with the advantage of hindsight. The value of land, a patent or a copyright has to do with how the market evaluates it. If the land is in the Mohave desert, the copyrighted work banal, or the patent on technology people don't want, then it may be worthless. If the land is in downtown Manhattan, the copyright on Donald Duck, or the patent on technology others want, it can be very valuable. The important thing about an invention is not so much that it be inventive, but that it be new if it is to be patented, and that it be useful if people are to buy it.

postponed introduction of its latest version for about a year so it could upgrade to IBM standards as part of an agreement in which IBM would market XyWrite exclusively. At the last minute IBM reneged on the deal [25].

While the first five patents are held by small entities, the last four patents are held by large entities and they also protected commercial products.

U.S. Patent 4,558,302
UniSys licenses this, the LZW compression patent, for 1% of sales. It has threatened a large entity with a lawsuit, but no small ones.

U.S. Patent 4,555,775
The League describes AT&T's backup store patent as "Too Obvious to Publish." Yet, in a letter in this issue of *Communications*, Dennis Richie points out that this technology was published in the ACM [36] and was recently called "a seminal paper" whose ideas are seen in X Windows Macintosh and many other windows systems [14]. While AT&T has sent notification letters on this patent, it has put the patent into reexamination and has not threatened suit or sued anyone on this patent.

U.S. Patent 4,656,583
This is an IBM patent on compiler speedup.

U.S. Patent 4,742,450
This is an IBM-shared copy on write patent. These two patents are what IBM calls Group 1 patents whose royalty is 1% of sales. They have been licensed by IBM as part of general licensing agreements but have not been licensed individually. (About 50 of IBM patents are Group 2 patents. Group 2 patents can be licensed for 2% each; the entire Group 1 portfolio, for 2%; and the entire IBM patent portfolio for 5%.)

These two patents have not been litigated and I do not believe IBM has aggressively asserted these patents against anyone. IBM, like most companies, normally files for patents only to protect what they expect to become commercial products. We treat these patents as protecting commercial products.

Most patents are never asserted. Much of the value of patents, like that of the Swiss Army, is that they act as a deterrent. The patents described here are typical of the small number where the patentholder forces a resolution: the infringer may take a license, design around the patent, or produce prior art showing there was no infringement.

Many letters asserting patents are no trespassing signs, putting potential infringers on notice should they infringe or telling them not to. They require no action. The notified companies might send prior art back to the patentholder, who might send it to the PTO for reexamination. The "infringer" may ignore the notice, waiting to see the reexamined patent or for the patentholder to become more assertive. The resolution may be hidden, in that an infringer may design around the patent. Rarely, a product is withdrawn from the market. The statistics on the cited patents are summarized in Table 4.1.

Whether any patent including those described here, is valid and infringed is a complex legal and technical question. An advantage of the patent system is that the question is an objective one based on the patent, prior art, and the "infringing" device. Such a dispute is less acrimonious than one in which the task is to evaluate testimony where one person yells "thief," and another "liar." Whether infringement actually occurred in any of the cases is irrelevant. The relevant question is did the original patentholders bring commercial products to market based on the patented technology and motivated by the rights a patent bestows?

Trademarks: Apple Paid $30 Million to Use the Name "Apple"

Trademark law is like patent law in that the first one who claims it gets to own it. The Beatles recorded on their own label, Apple Records. When Apple Computer was founded it agreed not to use the *Apple* name in the music business. Later, when the Macintosh played music, the Beatles sued. Apple Computer settled, paying about $30 million dollars to use the name "Apple."

I have been publicly accused of extorting Apple. Did I extort Apple? Did the Beatles extort Apple? Should the computer business have its own *sui generis* trademark law?

Table 4.1.
Patents cited in "Against Software Patents"

Company Size	Large		Small		Total	
Patent Activity						
Patents granted	4	100%	5	100%	9	100%
Protected commercial products	4	100%	5	100%	9	100%
License appears to be available	4	100%	4	80%	8	89%
Firm founded to develop technology	0	0%	4	80%	4	44%
Sued large entities	0	0%	4	80%	4	44%
Sued small entities	0	0%	2	50%	2	22%
First suit against small entity	0	0%	0	0%	0	0%
Suits threatened	1	25%	5	100%	5	67%
Patent asserted (notice sent)	2	50%	5	100%	7	78%
Resolution (patent asserted)						
Infringement removed	0	0%	1	20%	1	11%
Product removed from market	0	0%	1	20%	1	11%
Licenses	1	25%	2	40%	3	33%
Unresolved	1	25%	1	20%	2	22%
Nothing to Resolve (No notice)	2	50%	0	0%	2	22%
Total	4	100%	5	100%	9	100%

Note: We treat multiple patents covering the same technology as a single patent. With the exception of Quantel, the small entities were less than a dozen people at the time the patent was filed, and the large entities were Fortune 1000 companies. We assume that if a lawsuit was threatened, a notice was sent and if a lawsuit was filed, a lawsuit was first threatened. We assume that if a patent has been asserted—people have been sent notices—there is a matter to be resolved. Even if one characterizes Cadtrak and Refac as being in the business of litigating patents as some do, the relevant fact is that the original patentholders were small entities introducing commercial products protected by patents.

Analysis Results

The nine patents cited by the League summarized in Table 4.1 lead us to these conclusions:

1. Software patents stimulate companies to bring commercial products to market.
All nine patents protected commercial products.

2. Software patents stimulate new business formation.
Four of the nine patents were from startups founded to exploit the patented technology. A fifth filed for its patent in its seventh year. All five companies struggled for years.

3. Software patents stimulate the commercial introduction of fundamental advances by small entities.
The technology pioneered by at least three of the small patentholders was significant in that it started new product categories or was widely adopted in the industry.

4. Licenses are usually available where companies enforce patents.
Only Quantel seems to be unwilling to license its patent.

5. Where similar-size companies had a dispute, they settled differences quickly without litigation.
The only patent dispute between similar-size companies (XyQuest) was settled readily. No small entities were faced with a lawsuit brought by a large entity without the advantage of the patentholder having settled earlier with a large infringer.

6. Small entities incurred little if any royalty and litigation costs for infringing patents.
The only disputes in which a small entity paid patent royalties or was sued were those in which the patentholder had previously settled disputes with larger companies. No case was cited in which a big company aggressively went after a small one over patents, unless a large company had respected the patents first. The only such instance the author knows of is IBM (see Big Companies Do Sue Small Ones).

Big Companies Do Sue Small Ones

While the League says that big companies will use patents against small ones, it cites no example. Two came to my attention. (I was contacted because I had arguably relevant prior art on the first patent.) In both, IBM sued former employees to get ownership of patents on technology developed on their own time, unrelated to their work and only *after* the technology proved to have value in the market.

IBM vs. Goldwasser Civil 5:91 00021 D. Conn.

IBM encouraged employees to develop software products on their own time and seek patent protection for them so IBM could evaluate them for marketing. Goldwasser developed such a software product, but IBM rejected it; he left IBM stating he intended to pursue his technology, as he did. Six years later, another company introduced a product that seemed to be infringing his patents and he sued them. That company claimed it was covered under its cross-license with IBM; IBM sued Goldwasser to get ownership of the patent.

IBM vs. Zachariades C-91 20419

Zachariades before and while working for IBM developed on his own time a plastic valuable to the medical industry. He kept IBM informed about what he was doing, applied for patents, started his own company, and licensed the technology to a medical prosthesis company. When he was not paid, he sued and a jury awarded him $99 million. IBM "suddenly" found out what was happening and fired, and sued, him for the patents, telling him they did it in part to "terrorize" other IBM employees.

Companies like to hire litigators who know what it is like from the other side. In both cases IBM is represented by the same firm that represented Edwin Armstrong, the great inventor of modern radio when David Sarnoff and RCA were refusing to respect Armstrong's rights. Ken Burns tells the story in his PBS documentary, *Empire of the Air:* On January 31, 1954, Edwin Armstrong, under the strain of RCA's tactics—well dressed as always, in a suit, overcoat, scarf and gloves—jumped from his 13th floor apartment onto the third story roof of the River Club below [29]. His widow won all the patent suits.

Having hired a firm that experienced firsthand the tactics that caused a great inventor to kill himself, IBM should be able to, by suing Goldwasser and Zachariades, "terrorize" its employees.

7. Patent piracy by large entities appears to be common and small entities have a tough time getting their rights respected.

Four of the five small entities had large entities use their technology without first licensing it. *All four were forced to sue.* This makes it unlikely that all the patent disputes were an honest difference of opinion, although some probably were. For this reason "piracy" seems a fair characterization. These same small entities have had their patents mischaracterized and their motives impugned in the academic, trade and business press read by the software community.

It can cost over a million dollars to litigate a patent through to trial. The data shows that large entities are quick to use their power to try to intimidate small ones into abandoning their rights or accepting nuisance settlements rather than address infringement issues on their merits. It appears that this high rate of patent piracy is caused in part by the *Federal Rules of Civil Procedure* which tilt the scales of justice against the weak.

Our results confirm the League's suggestion that big companies will readily bully small ones, but refutes its suggestion that a patentholder who asserts a patent will get showered with gold. The yellow matter is not gold.

8. U.S. companies are slow to accept software innovations from outside sources.

The Japanese adapt innovations from sources outside the company twice as fast as U.S. companies [31]. The technology protected by at least two of the patents (CadTrak and HyperRacks) was exposed to companies that later became the first infringers.

Japan's ability to accommodate outside [the firm] innovation may be one of the reasons it has been so successful in dominating markets. If the U.S. is to exploit its strengths in innovation, it must learn to adapt outside innovations without the inefficiency of legal confrontation. Fast and efficient patent enforcement should encourage U.S. companies to license outside technology early rather than wait until they have an infringing product in the market and face legal exposure.

If large companies are forced to deal with infringement issues early, they might see it is to their advantage to work with the inventors, using their knowledge. Now, the legal system keeps the patentholder and infringer at war until such time as the patentholder's knowledge is of little

value to the "infringer," thus wasting one of our most valuable resources—the creativity and experience of innovators.

9. Developers do not seem to be infringing multiple patents on a single product.

The only example that was cited in which someone faced infringement issues from more than one patentholder seems to be X Windows facing the CadTrak and AT&T patents, but this has not been resolved and no lawsuits seem to have been filed or threatened.

10. The patent system seems to reject bad patents early in the patent assertion process.

We think the League is right in alleging that bad patents have been issued. The League, however, fails to identify a patent that was rejected by the courts. We think this is because issues of prior art and patent invalidity are considered early in the patent assertion process. Patentholders rarely continue to assert patents in the face of solid evidence of invalidity or noninfringement.

11. If software patents were more widely respected we would probably have had fewer variations on a theme, and more themes to vary on.

Product development effort seems to have focused on creating many versions of an invention once its value was proven. Over 250 different spreadsheets and at least four products generally considered to be HyperCard clones were marketed.

12. Big companies' patents do not seem to inhibit small developers.

The innovations protected by small entity patents listed here seem to have been more widely adopted than those of big companies in their industries. Big companies are better at commercializing and protecting their minor innovations, than their major ones.

That small entities seem to introduce the more fundamental innovations to the market is telling. Big companies are often unsuccessful in transforming innovations into commercial success: Xerox PARC pioneered much of modern-day personal computer and its software. Although IBM invented a predecessor to the spreadsheet (expired U.S. Patent 3,610,902), it did not market a commercial product based on it; it also did not assert the patent even though its claims seem to read on

(i.e., be infringed by) modern spreadsheets. These technologies became major product categories primarily through the efforts of small entities

13. Small entities using patents are exceptionally cost-effective in encouraging innovation—especially compared to federal funding.

Table 4.2 shows a rough estimate of the efficacy of three major sources of innovation: federally funded, large entity, and small entity. Our results show that small entities are 7.5 times as cost-effective at stimulating innovation as large ones, and 200,000 times as cost-effective as federal funding. The U.S. grants patent rights to universities as part of its research contracts; thus patents are issued in all these areas and patents asserted is a reasonable measure of innovation. We believe a more scientific study would refine these results, but doubt it would change the basic conclusion.

As a software developer, you might review the patents discussed and put yourself in the place of each of the parties involved. If your product finds satisfied users, do you think better financed companies with stronger marketing organizations will market competitive products, using your

Table 4.2.
Cost effectiveness to taxpayers of innovation sources

Innovation source	Commercial products	Efficacy (Patents Asserted)	Cost (000,000/yr)	Cost-effectiveness
Large entity	4	2	0.03	67.
Small entity	5	5	0.01	500.
Commercial sector	9	7	0.04	175.
Federally funded	1	0	$487.00	0.0021

Note: None of the nine patents appears to result from federal funding. However, we arbitrarily allocate one patent to this category so as to prevent zero results. While the PTO is virtually self-funded, $1.8 million of its $419 million budget comes from taxpayers. Since about 25% of the patents are from small entities and 75% from large entities we distribute the $1.8 million accordingly. We assume that 2% of the patents are software related; it is probably less. Government R & D in computer science was $487 million in fiscal 1989 out of total federally funded research of $61 billion.

innovations? If so, will patents be useful to you? If a patent is enforced against you, do you think you will be able to design around it? If you have to license it, do you think your competitors will also have to license it, thus passing the cost on to the end customers? Which problem would you rather have: a big company entering a market you developed, or finding out you were accidentally infringing a patent? Do you think the effect of software patents might be more innovation, higher software prices and an industry with more long-term profitability?

If you are protected by patents, your success depends in part on your patented inventions as others must deal with them. If you accidentally infringe a patent, designing around it is within your expertise. If you do not have patents, success depends much more on the ability to finance and market products—capabilities outside of your expertise and control. If you are a software developer, don't patents benefit you by manifesting your contributions in rights you can bring to the bargaining table, while confining the problems largely to your area of expertise and control? *Issues in Science and Technology* (Winter 1992) contains a letter from Commissioner of Patents Harry F. Manbeck who said of another article by the same authors that they [15]:

> demonstrate they do not understand the current law. . . . Most of their statements . . . do not appear to be the result of a balanced and reasoned inquiry and do not appear to be supported by the facts. . . . They cavalierly dismiss the view of those who appear to have used the patent system successfully and impugn their motives. . . .

The PTO issued about 89,000 patents in 1990 from which the League, with the advantage of hindsight, can pick and choose the ones to attack. Consider the information presented here on patents the League selected to demonstrate the PTO's mistakes. Whose standards are higher, the League's or the PTO's?

Recommendations

After reviewing our results we can make some general recommendations.

1. Policy should be made on the assumption that innovation occurs in software as in other technologies until compelling evidence to the contrary is found.

This is consistent with the results described here. The operational implications are to continue to let the system operate as it is accepting evolutionary changes based on experience rather than speculation.

2. The PTO should be viewed as a source of innovation that competes for funding with other federally funded sources of innovation.

PTO fees should be reduced, especially for small entities, and the PTO should receive a higher level of funding to improve its ability to examine patents so it can issue better quality, more timely, patents in software and other technologies. European patent offices are much better equipped and much better funded. It seems that the PTO should compete with the NSF and other organizations for federal funding on the basis of their cost effectiveness in encouraging innovation.

In 1990 only $2 million of the PTO budget of $419 million came from federal funding; the remainder came from user fees. Superficially, it might seem that investing in the patent system will have a multiplier effect of 80,000 in creating innovation as compared to federally funded science. We suggest no such thing. We do, however, ask the question: If taxpayers were to spend an additional $160 million per year to support innovation, we could either increase the $64 billion federal funding on science by one fourth of one percent (0.25%) or increase PTO funding by 40%, enabling the PTO to issue better patents and restore reduced user fees for small entities. Which will likely produce more innovation? Which will achieve a greater multiplier effect by encouraging additional private investment?

An example of federally funded science is fusion power research, which has been going on for at least 25 years, has cost hundreds of millions of dollars and has produced little practical result. Pons and Fleishman developed (and filed for patents on) cold fusion without government funding yet they, having invested their own money and not being in the mainline of governmental funding, are heavily criticized. While it is not clear that Pons and Fleishman have produced cold fusion, respected people in the field believe that they have, even if no one yet understands what is happening. This is just an example of how the system is biased in favor of government funding of expensive conventional solutions, and against individuals and small companies who risk their own time and money to innovate.

Individuals taking a contrary view have been the major source of new ideas in both science [38] and engineering [24]. This is why small entities and the patent system are so important. Most will fail, but the successes more than make up for the failures.

It would be interesting to evaluate the results of federally funded science to see which projects are worth the cost. Some projects may have become like those welfare mothers who, generation after generation, are entrapped in a governmental support system.

3. The patent laws should be modified to make it possible for small entities to assert their patent rights more effectively.

The data show it is commonplace for large companies to pirate the technology of small entities. No case was cited where a large company licensed a small entities' technology without first being sued suggests the existing laws do not motivate large companies to resolve patent disputes with small companies quickly. The issue here is not just fairness to inventors and improved efficiency in settling disputes. Rather, it is concerned with avoiding the waste that occurs because U.S. companies are so much slower at adopting new innovations than Japanese companies.

Congress responded with antipiracy legislation where software copyrights were concerned; we would hope it would similarly pass legislation to prevent patent piracy. Remedies similar to the criminal penalties for copyright infringement and Rule 11 sanctions for attorneys who file frivolous suits are worth considering. We suggest the following as possible remedies for patent disputes to stimulate discussion:

- After being put on notice, an "infringer" would have six months to file any prior art to be used to defend the infringement suit with the PTO. (I find it difficult to believe it is well-known art if it cannot be found in six months.)
- If a patentholder prevails in a lawsuit, the remedies should include an extension of the period of exclusivity against that infringer equal to the length of time the suit was in progress.
- Discovery should be limited.

These suggestions, which should induce speedier resolution of patents disputes, are suggested for all patents disputes, not just software.

The patent system, an enormously productive system for inducing innovation, is being stymied by a cumbersome dispute resolution process. Is it in the public good to have a system of conflict resolution that discourages conflict resolution? Should innovators spend their time innovating or litigating? If the courts could resolve software patent and copyrights issues more quickly, it would clarify the law so everyone can make decisions with some predictability. The problems are not unique to patents but occur in all litigation. That the judicial and even the legal, community are beginning to address the inefficiency of dispute resolution and litigation is grounds for cautious optimism.

4. Further study of the role of patents and federal funding in software innovation is useful.

We are keenly aware that the sample is small and unscientific, and thus our results should be considered suggestive rather than definitive. A more definitive study should be useful in bringing out facts that would be useful in evaluating future changes to the patent law.

These recommendations can be summarized thus: Redress the balance of incentives so innovators will prefer to develop their ideas commercially, using patent protection rather than search for federal funding.

The Software Patent Confrontation

The software industry is getting more competitive. Almost every company that has hit products uses its cash flow to develop entries in other product categories. As a result, product categories are getting very competitive. Since most software companies have confined their intellectual property to source code copyrights, user interface copyright and trademarks, whenever they come up with a successful innovation, their competitors will often quickly replicate it. As a result, the impetus is toward similarly featured products competing on price, differing only in the mistakes which the originators must maintain to support their existing customers.

Now companies are recognizing that by using patents they can compete on features and function—not just tactically, but strategically. Even if competitors do replicate the features, they will likely make them different

ALPHA: Abraham Lincoln Patent Holders Association

This organization, founded in January 1992, supports the use of, and educates people about, software patents. Already, ALPHA's members include two software patentholders whose patents have been litigated, four patentholders whose patents are mentioned here, two former board members of the Software Publishers Association and lawyers from Merchant and Gould, Baker and McKenzie, Welch and Katz and the Franklin Pierce Law Center, and a former commissioner of patents and trademarks.

enough to avoid infringement. Companies following this approach will support standards, but their products will have a substantial proprietary component engendering products with more diverse feature sets. This will enable the industry to compete more on the profitable playing field of unique capabilities and market position and less on price. This is consistent with standard business school product marketing, where product differentiation and market segmentation are basic.

Intellectual property has already driven the market for those who got in early and established standards. Lotus owns the 1-2-3 standard. Novell owns a major network standard and WordPerfect, a major word processor standard. Apple owns the Macintosh user interface standard and Intel and Microsoft own the IBM compatibility standard. Patents give new companies the opportunity to establish and own something of value in the market based on their innovativeness rather than their marketing and financial capabilities.

While the problem of people accidentally infringing software patents has been greatly exaggerated, several patents will be successfully asserted against existing products. This will be primarily between those companies that focused on innovation and have patents, and those that focused on exploiting recognized business opportunities. This kind of confrontation occurred earlier in the aircraft and other industries [21].

During these confrontations, the businesses with a large volume of infringing products will understandably feel "extorted" since they did not anticipate patent infringement. Such businesspeople will take support for their position from those who argue against software patents and advocate or suggest invalidating existing software patents [15, 39].

The software innovators who advanced the technology and made business decisions based on their patent rights will similarly feel cheated especially where they pioneered commercial products based on their inventions. When depositors made decisions based on government guarantees of S&L deposits, no one suggested that the government default on its obligations to insured depositors, as people suggest the government invalidate existing software patents. No one vilifies the S&L depositors because the government has to pay them money; yet software innovators find themselves vilified with lies and half truths.

In this confrontation, both sides start out feeling cheated. Many "infringers" will react emotionally and view it as a problem to be gotten rid of and many will fight to the bitter end. This raises the stakes, since a company having been put on notice may be liable for treble damages and attorneys' fees. Patentholders will not likely pursue these cases for four or five years to let the infringers' liability build up. After a suit is filed these companies will be getting much of their advice from those who have most to profit from the litigation: their litigators. This seems to be what is happening to Lotus.

Some software developers on finding out that the rules were not what they thought, face the problems of infringing others' patents while not having patented their own successful innovations. Some will chalk it up as one of many risks and uncertainties of business. Those who react emotionally might find it useful to first ask: Which of the players have acted in good faith? Which have not? Which have been responsible for the patent mess? Which have been innocent victims? Having answered these questions, such developers can more effectively target their wrath.

Companies that act rationally will analyze the patents to ascertain their scope and validity, whether infringement is occurring, and how easy it is to remove the "infringing" capability. They will check with other licensees. They will consider the obvious options, such as taking out a license, removing the infringing capability, finding prior art and showing it to the patentholder, or fighting in court if that is the only possibility. They will probably try to address the problem early, before the liability builds and consider negotiating a license or using some form of alternative dispute resolution to resolve infringement and validity issues. If the problem is associated with purchased products, most companies will stand in back

of their products and provide a license, warrant it against infringement, or provide guidelines on how to avoid infringement. It will put its "no problem" in writing.

Astute companies will view infringement as an opportunity in disguise. If the patent is good and competitors are, or soon will be, infringing it, the first licenses can generally get an inexpensive license forcing competitors to pay more if they want to use the technology. It may be possible to get an exclusive license on some feature which differentiates a product from the competitor's. It can be worthwhile to see if the patent covers useful capability which could be added to the product. The best time to negotiate for the license is when you do not have the liability of infringement but can offer to create a demand for that capability by incorporating it into a product. Invention being the mother of necessity, your competitors will be faced with the choice of paying a higher price to license the technology or leave it out, thus differentiating your product from theirs.

What You Can Do

Insist that the issues be debated. Don't let one side present its case unchallenged. If you need literature to distribute or a developer to join a debate on software, contact ALPHA.

One way to help correct the misinformation about software patents is to join ALPHA, The Abraham Lincoln Patent Holders Association. ALPHA is trying to correct the misinformation on software patents and to provide a forum for people to deal with issues of software patents, such as how to avoid infringement, setting licensing fees, finding prior art. Contact ALPHA at:

ALPHA
146 Main St.
Suite 404
Los Altos, CA 94022

You can help get the issue discussed intelligently in Congress.

Tell Congress it is important that the issues be discussed in open hearings where everybody gets to present his or her side, hear the issues debated and make up their own mind. Write:

House Subcommittee on Intellectual Property
2137 Rayburn Bldg.
Washington, DC 20515

Senate Subcommittee on Patents, Trademarks and Copyrights
U.S. Senate
Washington, DC 20510

In brief, what superficially looks like another problem to be dealt with in the increasingly competitive, commodities-oriented software business, might prove to be what makes products less *price* competitive. Many industries have worked on this basis all along; patents make industries more diverse in their offerings, more profitable, more innovative, and ultimately will make the U.S. more competitive.

The essence of this article is simple: Software intellectual property issues are not inherently different in substance from other technologies; what motivates people is not inherently different; industry life cycle is not inherently different; marketing and business strategies and tactics are not inherently different; the law and policy issues are not inherently different; the technology is not inherently new. Software has been around for 40 years. The issues may be new to those who had no experience with them. But the only difference is that software is a mass market industry for the first time and real money is at stake.

Acknowledgments

I would like to thank Steve Lundberg, John P. Sumner, Susan Nycum, Lewis Gable, George Gates, David Pressman and Tom Hassing for their many useful comments.

References

1. American Bar Association, *Two Hundred Years of English and American Patent, Trademark and Copyright Law*, 1977.

2. Axelrod, R. *The Evolution of Cooperation*. Basic Books, 1985.

3. Brooks, F. No silver bullet: Essence and accidents of software engineering. *Computer* (Apr. 1987).

4. Bruce, R. *Lincoln and the Tools of War*. University of Illinois Press, 1989.

5. Bulkeley, W. Will software patents cramp creativity? *Wall Street J.* (Mar. 14, 1989).

6. Bugbee, B.W. *The Genesis of American Copyright Law.* Public Affairs Press, Wash., D.C., 1967.

7. CBEMA comments on computer-related invention patents. *Comput. Lawyer* (Oct. 1991).

8. Chisum, D. The patentability of algorithms. *U Pitt. L. Review* 47, 959-971, (1986).

9. Chisum, D. *A Treatise on the Law of Patentability, Validity and Infringement,* (7 volumes) M. Bender, 1978.

10. Choate, P. *Agents of Influence.* Touchstone, 1990.

11. Clapes, A. Software copyright, and competition. *Quorem* (1989).

12. Cringely, R.X. *Accidental Empires.* Addison Wesley, 1991.

13. Fisher, L.M. Software industry in uproar over recent rush of patents. *New York Times* (May 12), 1989.

14. Foley, et al. *Computer Graphics: Principles and Practice,* Second ed. Addison Wesley, 1990.

15. Garfinkel, S., Stallman, R. and Kapor, M. Why patents are bad for software. *Issues in Science and Tech.* (Fall 1991).

16. Heckel, P. The Wright Brothers and Software Innovation, in *The Elements of Friendly Software Design,* Second ed., Sybex, 1991 (First ed. Warner Books, 1984).

17. Heckel, P. Isolating differences between files. *Commun. ACM* (Apr. 1978).

18. Heckel, P. Software patents and competitiveness. Op Ed, *San Francisco Examiner,* (July 8, 1991), A13.

19. Heckel, P. Zoomracks, designing a new software metaphor. *Dr. Dobbs J.* (Nov. 1985).

20. Heckel, P. and Lampson, B. A terminal oriented communications system. *Commun. ACM* (July 1977).

21. Heckel, P. and Schroeppel, R. Software techniques cram functions and data into pocket sized microprocessor applications. *Elect. Des.* (Apr. 12, 1980).

22. Howard, F. *Wilbur and Orville.* Knopf, 1988.

23. *In re Pardo,* 684 F.2d 912 (C.C.P.A. 1982).

24. Jewkes, J., Sawers, D. and Stillerman, R. *The Sources of Invention,* Second ed., Norton, 1969.

25. Judas, J.B. Innovation, a casualty at IBM. *Wall Street J.* (Oct. 17, 1991), A23.

26. Kahin, B. The software patent crisis. *Tech. Rev.* (Apr. 1960), 543-558.

27. League for Programming Freedom. Against software patents. *Commun. ACM* (Jan. 1992).

28. League for Programming Freedom. Software patents. *Dr. Dobbs J.* (Nov. 1990).

29. Lewis, T. *Empire of the air.* HarperCollins, 1991.

30. Lincoln, A. *Selected Speeches and Writings.* Vintage, 1992.

31. Mansfield, E. Industrial innovation in Japan and the United States. *Science* (Sept. 30, 1988).

32. McCloskey, M. Intuitive physics. *Sci. Am.* (Apr. 1983), 123.

33. Neustadt, R.E. and May, E.R. *Thinking in time: The uses of history for decisionmakers.* The Free Press, 1986.

34. Nycum, S. Legal protection for computer programs. *Comput. Law J.* 1, 1 (1978).

35. Orwell, G. *Politics and the English Language.*

36. Pike, R. Graphics in overlapped bitmap layers. *ACM Trans. Graph.* 17, 3 (July 1983), 331.

37. Ritter, T. The politics of software patents. *Midnight Eng.* (May–June 1991).

38. Root-Bernstein, R. *Discovering.* Harvard University Press, 1989.

39. Samuelson, P. Benson revisited: The case against patent protection for algorithms and other computer program-related inventions. *Emory Law J.* 39, 1025 (1990).

40. Samuelson, P. CONTU revisited: The case against copyright protection for computer programs in machine readable form. *Duke Law J.* 663 (1984), 705–53.

41. Samuelson, P. Should program algorithms be patented? *Commun. ACM* (Aug. 1990).

42. Samuelson, P. and Glushko, R. Survey on the look and feel lawsuit. *Commun. ACM* (May 1990).

43. Schon, D. *Technology and Change.* Delacorte, 1967.

44. Schwarzer, W. Science in the Courtroom. 15th Annual Intellectual Property Law Institute, Intellectual Property Section of the California State Bar, Nov. 1990.

45. Schwartz, E. The coming showdown over software patents. *Bus. Week* (May 13, 1991).

46. Sumner, J. and Lundberg, S. The versatility of software patent protection: From subroutines to look and feel. *Comput. Lawyer* (June 1986).

47. Sumner, J. and Lundberg, S. Software patents: Are they here to stay? *Comput. Lawyer* (Oct. 1991).

48. Slutsker, G. and Churbuck, D. Whose invention is it anyway? *Forbes* (Aug. 19, 1991).

II

How Should We Respond to Exploratory Hacking/Cracking/Phreaking?

Peter Ludlow

Hacker is a term that has two uses on the electronic frontier. Originally, a hacker was someone who liked to hack computer code (i.e., write programs) or, in some cases, hack electronic hardware (i.e., design and build hardware). Thanks to the news media, "hacker" has also come to have a negative connotation, usually meaning those who illicitly hack their way into other people's computer systems. Some folks have tried to preserve the original (good) sense of "hacker" by introducing the term *cracker* to cover the cases of electronic trespassers, but like all attempts to fight lexical drift, their efforts have failed. In any case, the idea that there are certain kinds of hacking that are illicit begs the central question of this section, namely, whether there is anything wrong with hacking (or cracking) your way into someone else's system.

The knee-jerk reaction is to say that trespassing is trespassing whether it is real-world trespassing or the electronic kind, but this reaction needs to be defended. There are lots of reasons real-world trespassing laws might be justified. Trespassers might hurt themselves on our property, thus exposing us to legal liability, or trespassers might pose a potential physical threat to us, or they might pose a threat to our property. These considerations do not carry over neatly to electronic trespassing. Hackers certainly aren't going to hurt themselves as they browse our system, and they do not pose an immediate threat to us, although they may try to crash the system, which can certainly ruin your entire day!

But even the system-crashing justification for electronic trespassing laws at best answers a policy question, not the central conceptual question. The policy question is whether sysops have a right to try and keep exploratory hackers out. The conceptual question is whether there is anything wrong with the hackers trying to get in (assuming that they intend no harm). Is there some sense in which (nondestructive) exploratory hacking is just plain wrong?

One possible argument against exploratory hacking is that it involves a kind of invasion of privacy. Isn't it an invasion of my privacy for you to poke around in my system and read my files? To some, however, this argument gets privacy considerations completely backward. The real invasion of privacy, they argue, occurs when corporations like TRW keep records of our personal financial transactions in a centralized data base, and sell those records to other corporations (and individuals) for a price.

In fact, according to this line of thinking, exploratory hackers have actually exposed cases of privacy invasion by uncovering files that were (illicitly) kept on their friends. Perhaps exploratory hackers help ensure that our privacy is not violated by centralized data bases.

Still, there remains the issue of the hacker who takes an interest in my system when there is no reason to suppose that my system contains detailed files on anyone except for me. Whatever the merits of hacking into a large data base, isn't the hacker invading my privacy by hacking into my system? The answer is far from clear. Consider, for example, my garbage, which I put out on the street twice a week. Legally, anyone can pick it up and go through it looking for clues and information about my life. Legally, it is not an invasion of my privacy because I have no reasonable expectation that what I place on the street in a garbage can will not be compromised. If I am concerned about my privacy I had best shred my documents or incinerate my trash. Analogously, it can be argued that if I am concerned about the security of my Internet site I had best encrypt my sensitive documents, or perhaps keep sensitive documents off Internet sites altogether.

It might be argued that exploratory hacking is wrong because it amounts to theft of proprietary information. In the previous section we saw that the very notion of theft of information is a matter of debate. If no one can own information, then how can someone steal it? Even if we agreed that it was wrong to steal information from a remote system, it would not follow that hacking into that system was wrong. One motivation for hacking in the early days was to gain access to a system such as UNIX so that one could learn how the operating system worked. Some have claimed that these considerations no longer apply because a UNIX box can be acquired for a few hundred dollars, but in the abstract the point needs to be taken seriously. Would it be wrong to hack into a system with no intention of damaging the system or even reading files, but merely to try and understand how the system works?

The above question can be sharpened with the help of an analogy. We can distinguish between car theft and joy-riding, where someone merely "borrows" my car (without permission) in order to take a spin. Ordinarily, we consider theft much more serious than joy-riding, and would deal with the perpetrators in different ways. But now consider a hypothetical

"car-hacker" who borrows cars without permission to open them up and study how they work. Let's suppose further that this car-hacker was not studying these cars for financial gain, but merely to satisfy his or her curiosity about internal combustion engines. Would we really want to treat this car-hacker on a par with a genuine car thief or even a joy-rider? Surely there seems to be something much more redeeming about the motives of the hypothetical car-hacker, even if his or her actions became something of a nuisance. Likewise, it might be argued that the electronic system hacker should not be treated as a common thief or trespasser, because the motives are, by hypothesis, simply to learn.

Of course, apprehended hackers are often treated much worse than common thieves or trespassers. In his essay on Phiber Optik, Julian Dibbell speculates that the sentencing of Phiber was not due to the moral content of the crime, but rather to the fact that hackers in general and Phiber in particular represent anarchy at a time when corporate robber barons are trying to seize control of the electronic frontier. Dibbell might also have added that hackers represent an embarrassment to these interests as well, for hackers show that one individual, armed only with a laptop computer can out-maneuver corporations with security budgets in the tens of millions of dollars. The key word here is embarrassment, and if The Mentor is right, the crux of the problem is that the true crime of the hacker is being too smart. As he says in his "Conscience of a Hacker": "My crime is that of outsmarting you, something that you will never forgive me for."

So far my remarks have suggested that hackers are only interested in exploring computer systems, but this, of course, is too narrow a view. As a brief perusal of the magazine *2600* ("The Hacker Quarterly") suggests, hacking might involve any sort of activity from building a cable-TV descrambler to constructing a red box (for simulating the tone made by a pay phone). Can these activities be part of a learning exercise? In his congressional testimony, Emmanuel Goldstein (editor of *2600*) suggests that they can. Of course, as Congressman Markey points out in his questioning, such devices can also be used to break the law, and the question arises as to how appropriate it is for Goldstein to publish information on how to build such systems. In Goldstein's view, the fact that the information might be misused is no reason to keep that information

bottled up. He also stresses that one needs to distinguish between hackers and those who use hacker-like methods to break the law. Consider the following remarkable exchange between Goldstein and Markey.

Mr. Markey . . . Let's go to the other side of the problem, the joy rider or the criminal that is using this information. What penalties would you suggest to deal with the bad hacker? Are there bad hackers?

Mr. Goldstein There are a few bad hackers. I don't know any myself, but I'm sure there are.

Mr. Markey I assume if you knew any, you would make sure we did something about them. But let's just assume there are bad people subscribing. What do we do about the bad hacker?

Mr. Goldstein Well, I just would like to clarify something. We have heard here in testimony that there are gang members and drug members who are using this technology. Now, are we going to define them as hackers because they are using the technology?

Mr. Markey Yes. Well, if you want to give them another name, fine. We will call them hackers and crackers, all right?

Mr. Goldstein I think we should call them criminals.

Mr. Markey So the crackers are bad hackers, all right? If you want another word for them, that is fine, but you have got the security of individuals decreasing with the sophistication of each one of these technologies, and the crackers are out there. What do we do with the crackers who buy your book?

Mr. Goldstein I would not call them crackers. They are criminals. If they are out there doing something for their own benefit, selling information—

Mr. Markey Criminal hackers. What do we do with them?

Mr. Goldstein There are existing laws. Stealing is still stealing.

One of the themes of Goldstein's testimony, the idea of hacking as a kind of quest for knowledge has been elevated to something of a "hacker ethic" in some quarters—an ethic in which the hacker construes his or her role as the liberator of information or as a disseminator of knowledge. To this end Dorothy Denning, in her study of hackers, has envisioned a coming ethical conflict between the bureaucratic tendency to hoard information and the hacker ethic of acquiring and sharing information. In Denning's words:

> Hackers say that it is our social responsibility to share information, and that it is information hoarding and disinformation that are the crimes. This ethic of resource and information sharing contrasts sharply with computer security policies that are based on authorization and "need to know." This discrepancy raises an interesting question: Does the hacker ethic reflect a growing force in society that stands for greater sharing of resources and information—a reaffirmation of basic values in our constitution and laws?

This nicely frames what I view as the central conceptual question of this section, whether the underlying ethic of hacking is one that we ought to encourage and indeed nurture. Can we, for example, learn something from hackers and their curiosity, or do they represent a dangerous challenge to our extant conceptions of property and information control? Again from Denning:

> What conflict in society do hackers stand at the battle lines of? Is it owning or restricting information vs. sharing information—a tension between an age-old tradition of controlling information as property and the Enlightenment tradition of sharing and disseminating information? Is it controlling access based on "need to know," as determined by the information provider, vs. "want to know," as determined by the person desiring access? Is it law enforcement vs. freedoms granted under the First and Fourth Amendments? . . . The issue is not simply hackers vs. system managers or law enforcers; it is a much larger question about values and practices in an information society.

7

The Conscience of a Hacker

The Mentor

Another one got caught today, it's all over the papers. "Teenager Arrested in Computer Crime Scandal," "Hacker Arrested after Bank Tampering"...

Damn kids. They're all alike.

But did you, in your three-piece psychology and 1950's technobrain, ever take a look behind the eyes of the hacker? Did you ever wonder what made him tick, what forces shaped him, what may have molded him?

I am a hacker, enter my world . . .

Mine is a world that begins with school . . . I'm smarter than most of the other kids, this crap they teach us bores me . . .

Damn underachievers. They're all alike.

I'm in junior high or high school. I've listened to teachers explain for the fifteenth time how to reduce a fraction. I understand it. "No, Ms. Smith, I didn't show my work. I did it in my head . . ."

Damn kid. Probably copied it. They're all alike.

I made a discovery today. I found a computer. Wait a second, this is cool. It does what I want it to. If it makes a mistake, it's because I screwed it up. Not because it doesn't like me . . .

Or feels threatened by me . . .

Or thinks I'm a smart ass . . .

Or doesn't like teaching and shouldn't be here . . .

Damn kid. All he does is play games. They're all alike.

And then it happened . . . a door opened to a world . . . rushing through the phone line like heroin through an addict's veins, an electronic pulse is sent out, a refuge from the day-to-day incompetencies is sought . . . a board is found.

"This is it . . . this is where I belong . . ."

I know everyone here . . . even if I've never met them, never talked to them, may never hear from them again . . . I know you all . . .

Damn kid. Tying up the phone line again. They're all alike . . .

You bet your ass we're all alike . . . we've been spoon-fed baby food at school when we hungered for steak . . . the bits of meat that you did let slip through were pre-chewed and tasteless. We've been dominated by sadists, or ignored by the apathetic. The few that had something to teach found us willing pupils, but those few are like drops of water in the desert.

This is our world now . . . the world of the electron and the switch, the beauty of the baud. We make use of a service already existing without paying for what could be dirt-cheap if it wasn't run by profiteering gluttons, and you call us criminals. We explore . . . and you call us criminals. We seek after knowledge . . . and you call us criminals. We exist without skin color, without nationality, without religious bias . . . and you call us criminals. You build atomic bombs, you wage wars, you murder, cheat, and lie to us and try to make us believe it's for our own good, yet we're the criminals.

Yes, I am a criminal. My crime is that of curiosity. My crime is that of judging people by what they say and think, not what they look like. My crime is that of outsmarting you, something that you will never forgive me for.

I am a hacker, and this is my manifesto. You may stop this individual, but you can't stop us all . . . after all, we're all alike.

8

The Prisoner: Phiber Optik Goes Directly to Jail

Julian Dibbell

Phiber Optik is going to prison this week and if you ask me and a whole lot of other people, that's just a goddamn shame.

To some folks, of course, it's just deserts. Talk to phone-company executives, most computer-security experts, any number of U.S. attorneys and law-enforcement agents, or Justice Louis Stanton of the Southern District of New York (who handed Phiber his year-and-a-day in the federal joint at Minorsville, Pennsylvania), and they'll tell you the sentence is nothing more than what the young hacker had coming to him. They'll tell you Phiber Optik is a remorseless, malicious invader of other people's computers, a drain on the economic lifeblood of our national telecommunications infrastructure, and/or a dangerous role model for the technoliterate youth of today.

The rest of us will tell you he's some kind of hero. Just ask. Ask the journalists like me who have come to know this 21-year-old high-school dropout from Queens over the course of his legal travails. We'll describe a principled and gruffly plain-talking spokesdude whose bravado, street-smart style, and remarkably unmanipulative accessibility have made him the object of more media attention than any hacker since Robert Morris nearly brought down the Internet. Or ask the on-line civil libertarians who felt that Phiber's commitment to nondestructive hacking and to dialogue with the straight world made him an ideal poster boy for their campaign against the repressive excesses of the government's war on hackers. You might even ask the small subset of government warriors who have arrived at a grudging respect for Phiber's expertise and the purity of his obsession with the workings of the modern computerized

phone system (a respect that has at times bordered on parental concern as it grew clear that a 1991 conviction on state charges of computer trespass had failed to curb Phiber's reckless explorations of the system).

But for a truly convincing glimpse of the high regard in which Phiber Optik is held in some quarters, you'd have to pay an on-line visit to ECHO, the liberal-minded but hardly cyberpunk New York bulletin-board system where Phiber has worked as resident technical maven since last spring. Forsaking the glories of phonephreaking for the workaday pleasures of hooking the system up to the Internet and helping users navigate its intricacies, he moved swiftly into the heart of ECHO's virtual community (which took to referring to him by the name his mother gave him—Mark—as often as by his nom de hack). So that when he was indicted again, this time on federal charges of unauthorized access to phone-company computers and conspiracy to commit further computer crimes, ECHO too was drawn into the nerve-racking drama of his case.

As the "coconspirators" named in the indictment (a group of Phiber's friends and government-friendly ex-friends) pleaded guilty one by one, there remained brave smiles and high hopes for Phiber's jury trial in July. By the time the trial date arrived, however, Phiber had made an agonizing calculus of risks and decided to plead guilty to one count each of computer intrusion and conspiracy. ECHO was left on tenterhooks waiting for the day of the sentencing. Given Mark's newfound enthusiasm for more legitimate means of working with computers and his undisputed insistence at the time of his plea that he had never damaged or intended to damage any of the systems he broke into, it seemed reasonable to wish for something lenient. A long probation, maybe, or at worst a couple months' jail time. After all, the infamous Morris had done considerably greater harm, and he got off with no jail time at all.

When the news arrived, therefore, of Phiber's 12-month prison sentence (plus three years' probation and 600 hours of service), it hit like a slap in the face, and ECHO responded with a massive outburst of dismay and sympathy. ECHO's director, Stacy Horn, posted the information at 3 PM on November 3 in the system's main conference area, and within 24 hours the place was flooded with over 100 messages offering condolences, advice on penitentiary life, and curses on Judge Stanton. Not all the messages were what you'd want to call articulate ("shit," read the first

one in its entirety; quoth another: "fuckfuckfuckfuckfuckfuckfuckfuck-fuckfuckfuckfuckfuckfuckfuckfuckfuckfuck"), nor was all the advice exactly comforting ("Try not to get killed," a sincere and apparently quite prison-savvy Echoid suggested; "Skip the country," proposed one user who connects from abroad, inviting Phiber to join him in sunny South Africa). But the sentiment throughout was unmistakably heartfelt, and when Phiber Optik finally checked in, his brief response was even more so:

"I just finished reading all this and . . . I'm speechless. I couldn't say enough to thank all of you."

He didn't have to thank anybody, of course. Motivated by genuine fellow feeling as this electronic lovefest was, it was also the last step in the long-running canonization of Phiber Optik as the digital age's first full-fledged outlaw hero, and making somebody else a hero is not necessarily the most generous of acts. For one thing, we tend to get more from our heroes than they get from us, and for another, we tend to be heedless of (when not morbidly fascinated by) the very high psychic overhead often involved in becoming a hero—especially the outlaw kind. To their credit, though, the Echoids proved themselves sensitive to the weight of the burden Phiber had been asked to take on. As one of them put it: "Sorry Mark. You've obviously been made a martyr for our generation."

There was some melodrama in that statement, to be sure, but not too much exaggeration. For ironically enough, Judge Stanton himself seemed to have endorsed its basic premise in his remarks upon passing sentence. Not unmoved by the stacks of letters sent him in support of Phiber Optik's character and motivations, the judge allowed as how a less celebrated Phiber Optik convicted of the same crimes might not deserve the severity of the discipline he was about to prescribe (and in Phiber's case it could be argued that 12 months locked up without a computer is severe enough to rate as cruel and unusual). But since Phiber had made of himself a very public advertisement for the ethic of the digital underground, the judge insisted he would have to make of the sentence an equally public countermessage. "The defendant . . . stands as a symbol here today," said Stanton, making it clear that the defendant would therefore be punished as one too.

The judge did not make it clear when exactly it was that the judicial system had abandoned the principle that the punishment fits the crime and not the status of the criminal, though I suppose that happened too long ago to be of much interest. More frustratingly, he also didn't go into much detail as to what it was that Phiber Optik was to stand as a symbol *of*. In at least one of his remarks, however, he did provide an ample enough clue:

"Hacking crimes," said Judge Stanton, "constitute a real threat to the expanding information highway."

That "real threat" bit was a nice dramatic touch, but anyone well-versed in the issues of the case could see that at this point the judge was speaking symbolically. For one thing, even as practiced by the least scrupulous joyriders among Phiber Optik's subcultural peers, hacking represents about as much of a threat to the newly rampant telecommunications juggernaut as shoplifting does to the future of world capitalism. But more to the point, everybody recognizes by now that all references to information highways, super or otherwise, are increasingly just code for the corporate wet dream of a pay-as-you-go telecom turnpike, owned by the same megabusinesses that own our phone and cable systems today and off-limits to anyone with a slender wallet or a bad credit rating. And *that,* symbolically speaking, is what Phiber Optik's transgressions threaten.

For what did his crimes consist of after all? He picked the locks on computers owned by large corporations, and he shared the knowledge of how to do it with his friends (they had given themselves the meaningless name MOD, more for the thrill of sounding like a conspiracy than for the purpose of actually acting like one). In themselves the offenses are trivial, but raised to the level of a social principle, they do spell doom for the locks some people want to put on our cyberspatial future. And I'm tempted, therefore, to close with a rousing celebration of Phiber Optik as the symbol of a spirit of anarchic resistance to the corporate Haussmannization of our increasingly information-based lives, and to cheer Phiber's hero status in places like ECHO as a sign that that spirit is thriving.

But I think I'll pass for now. Phiber Optik has suffered enough for having become a symbol, and in any case his symbolic power will always be available to us, no matter where he is. Right now, though, the man himself is going away for far too long, and like I said, that's nothing but a goddamn shame.

9

Concerning Hackers Who Break into Computer Systems

Dorothy E. Denning

1 Introduction

The world is crisscrossed with many different networks that are used to deliver essential services and basic necessities—electric power, water, fuel, food, goods, to name a few. These networks are all publicly accessible and hence vulnerable to attacks, and yet virtually no attacks or disruptions actually occur.

The world of computer networking seems to be an anomaly in the firmament of networks. Stories about attacks, breakins, disruptions, theft of information, modification of files, and the like appear frequently in the newspapers. A diffuse group called "hackers" is often the target of scorn and blame for these actions. Why are computer networks any different from other vulnerable public networks? Is the difference the result of growing pains in a young field? Or is it the reflection of deeper tensions in our emerging information society?

There are no easy or immediate answers to these questions. Yet it is important to our future in a networked, information-dependent world that we come to grips with them. I am deeply interested in them. This paper is my report of what I have discovered in the early stages of what promises to be a longer investigation. I have concentrated my attention in these early stages on the hackers themselves. Who are they? What do they say? What motivates them? What are their values? What do they have to say about public policies regarding information and computers? What do they have to say about computer security?

From such a profile I expect to be able to construct a picture of the discourses in which hacking takes place. By a discourse I mean the

invisible background of assumptions that transcends individuals and governs our ways of thinking, speaking, and acting. My initial findings lead me to conclude that this discourse belongs at the very least to the gray areas between larger conflicts that we are experiencing at every level of society and business, the conflict between the idea that information cannot be owned and the idea that it can, and the conflict between law enforcement and the First and Fourth Amendments.

But, enough of the philosophy. On with the story!

2 Opening Moves

In late fall of 1989, Frank Drake (not his real name), editor of the now defunct cyberpunk magazine *W.O.R.M.*, invited me to be interviewed for the magazine. In accepting the invitation, I hoped that something I might say would discourage hackers from breaking into systems. I was also curious about the hacker culture. This seemed like a good opportunity to learn about it.

The interview was conducted electronically. I quickly discovered that I had much more to learn from Drake's questions than to teach. For example, he asked: "Is providing computer security for large databases that collect information on us a real service? How do you balance the individual's privacy vs. the corporations?" This question surprised me. Nothing that I had read about hackers ever suggested that they might care about privacy. He also asked: "What has (the DES) taught us about what the government's (especially NSA's) role in cryptography should be?" Again, I was surprised to discover a concern for the role of the government in computer security. I did not know at the time that I would later discover considerable overlap in the issues discussed by hackers and those of other computer professionals.

I met with Drake to discuss his questions and views. After our meeting, we continued our dialog electronically with me interviewing him. This gave me the opportunity to explore his views in greater depth. Both interviews appear in "Computers Under Attack," edited by Peter Denning (DenningP90).

My dialog with Drake increased my curiosity about hackers. I read articles and books by or about hackers. In addition, I had discussions

with nine hackers whom I will not mention by name. Their ages ranged from 17 to 28.

The word "hacker" has taken on many different meanings ranging from 1) "a person who enjoys learning the details of computer systems and how to stretch their capabilities" to 2) "a malicious or inquisitive meddler who tries to discover information by poking around . . . possibly by deceptive or illegal means . . ." (Steele83). The hackers described in this paper are both learners and explorers who sometimes perform illegal actions. However, all of the hackers I spoke with said they did not engage in or approve of malicious acts that damage systems or files. Thus, this paper is not about malicious hackers. Indeed, my research so far suggests that there are very few malicious hackers. Neither is this paper about career criminals who, for example, defraud businesses, or about people who use stolen credit cards to purchase goods. The characteristics of many of the hackers I am writing about are summed up in the words of one of the hackers: "A hacker is someone who experiments with systems. . . . [Hacking] is playing with systems and making them do what they were never intended to do. Breaking in and making free calls is just a small part of that. Hacking is also about freedom of speech and free access to information—being able to find out anything. There is also the David and Goliath side of it, the underdog vs. the system, and the ethic of being a folk hero, albeit a minor one."

Richard Stallman, founder of the Free Software Foundation who calls himself a hacker according to the first sense of the word above, recommends calling security-breaking hackers "crackers" (Stallman84). While this description may be more accurate, I shall use the term "hacker" since the people I am writing about call themselves hackers and all are interested in learning about computer and communication systems. However, there are many people like Stallman who call themselves hackers and do not engage in illegal or deceptive practices; this paper is also not about those hackers.

In what follows I will report on what I have learned about hackers from hackers. I will organize the discussion around the principal domains of concerns I observed. I recommend Meyer's thesis (Meyer89) for a more detailed treatment of the hackers' social culture and networks, and Meyer

and Thomas (MeyerThomas90) for an interesting interpretation of the computer underground as a postmodernist rejection of conventional culture that substitutes "rational technological control of the present for an anarchic and playful future."

I do not pretend to know all the concerns that hackers have, nor do I claim to have conducted a scientific study. Rather, I hope that my own informal study motivates others to explore the area further. It is essential that we as computer security professionals take into account hackers' concerns in the design of our policies, procedures, laws regulating computer and information access, and educational programs. Although I speak about security-breaking hackers as a group, their competencies, actions, and views are not all the same. Thus, it is equally important that our policies and programs take into account individual differences.

In focusing on what hackers say and do, I do not mean for a moment to set aside the concerns of the owners and users of systems that hackers break into, the concerns of law enforcement personnel, or our own concerns as computer security professionals. But I do recommend that we work closely with hackers as well as these other groups to design new approaches and programs for addressing the concerns of all. Like ham radio operators, hackers exist, and it is in our best interest that we learn to communicate and work with them rather than against them.

I will suggest some actions that we might consider taking, and I invite others to reflect on these and suggest their own. Many of these suggestions are from the hackers themselves; others came from the recommendations of the ACM Panel on Hacking (Lee86) and from colleagues.

I grouped the hackers' concerns into five categories: access to computers and information for learning; thrill, excitement and challenge; ethics and avoiding damage; public image and treatment; and privacy and first amendment rights. These are discussed in the next five subsections. I have made an effort to present my findings as uncritical observations. The reader should not infer that I either approve or disapprove of actions hackers take.

3 Access to Computers and Information for Learning

Although Levy's book *Hackers* (Levy84) is not about today's security-breaking hackers, it articulates and interprets a "hacker ethic" that is

shared by many of these hackers. The ethic includes two key principles that were formulated in the early days of the AI Lab at MIT: "Access to computers—and anything which might teach you something about the way the world works—should be unlimited and total," and "All information should be free." In the context in which these principles were formulated, the computers of interest were research machines and the information was software and systems information.

Since Stallman is a leading advocate of open systems and freedom of information, especially software, I asked him what he means by this. He said: "I believe that all generally useful information should be free. By 'free' I am not referring to price, but rather to the freedom to copy the information and to adapt it to one's own uses." By "generally useful" he does not include confidential information about individuals or credit card information, for example. He further writes: "When information is generally useful, redistributing it makes humanity wealthier no matter who is distributing and no matter who is receiving." Stallman has argued strongly against user interface copyright, claiming that it does not serve the users or promote the evolutionary process (Stallman90).

I asked hackers whether all systems should be accessible and all information should be free. They said that it is OK if some systems are closed and some information, mainly confidential information about individuals, is not accessible. They make a distinction between information about security technology, e.g., the DES, and confidential information protected by that technology, arguing that it is the former that should be accessible. They said that information hoarding is inefficient and slows down evolution of technology. They also said that more systems should be open so that idle resources are not wasted. One hacker said that the high costs of communication hurts the growth of the information economy.

These views of information sharing seem to go back at least as far as the 17th and 18th centuries. Samuelson (Samuelson89) notes that "The drafters of the Constitution, educated in the Enlightenment tradition, shared that era's legacy of faith in the enabling powers of knowledge for society as well as the individual." She writes that our current copyright laws, which protect the expression of information, but not the information itself, are based on the belief that unfettered and widespread dissemination of information promotes technological progress. (Similarly for

patent laws which protect devices and processes, not the information about them.) She cites two recent court cases where courts reversed the historical trend and treated information as ownable property. She raises questions about whether in entering the Information Age where information is the source of greatest wealth, we have outgrown the Enlightenment tradition and are coming to treat information as property.

In a society where knowledge is said to be power, Drake expressed particular concern about what he sees as a growing information gap between the rich and poor. He would like to see information that is not about individuals be made public, although it could still be owned. He likes to think that companies would actually find it to their advantage to share information. He noted how IBM's disclosure of the PC allowed developers to make more products for the computers, and how Adobe's disclosure of their fonts helped them compete against the Apple-Microsoft deal. He recognizes that in our current political framework, it is difficult to make all information public, because complicated structures have been built on top of an assumption that certain information will be kept secret. He cites our defense policy, which is founded on secrecy for military information, as an example.

Hackers say they want access to information and computing and network resources in order to learn. Both Levy (Levy84) and Landreth (Landreth89) note that hackers have an intense, compelling interest in computers and learning, and many go into computers as a profession. Some hackers break into systems in order to learn more about how the systems work. Landreth says these hackers want to remain undiscovered so that they can stay on the system as long as possible. Some of them devote most of their time to learning how to break the locks and other security mechanisms on systems; their background in systems and programming varies considerably. One hacker wrote: "A hacker sees a security hole and takes advantage of it because it is there, not to destroy information or steal. I think our activities would be analogous to someone discovering methods of acquiring information in a library and becoming excited and perhaps engrossed."

We should not underestimate the effectiveness of the networks in which hackers learn their craft. They do research, learn about systems, work in groups, write, and teach others. One hacker said that he belongs

to a study group with the mission of churning out files of information and learning as much as possible. Within the group, people specialize, collaborate on research projects, share information and news, write articles, and teach others about their areas of specialization. Hackers have set up a private system of education that engages them, teaches them to think, and allows them to apply their knowledge in purposeful, if not always legal, activity. Ironically, many of our nation's classrooms have been criticized for providing a poor learning environment that seems to emphasize memorization rather than thinking and reasoning. One hacker reported that through volunteer work with a local high school, he was trying to get students turned on to learning.

Many hackers say that the legitimate computer access they have through their home and school computers do not meet their needs. One student told me that his high school did not offer anything beyond elementary courses in BASIC and PASCAL, and that he was bored by these. Hans Huebner, a hacker in Germany who goes by the name Pengo, wrote in a note to the RISKS Forum (Huebner89) : "I was just interested in computers, not in the data which has been kept on their disks. As I was going to school at that time, I didn't even have the money to buy my own computer. Since CP/M (which was the most sophisticated OS I could use on machines which I had legal access to) didn't turn me on anymore, I enjoyed the lax security of the systems I had access to by using X.25 networks. You might point out that I should have been patient and waited until I could go to the university and use their machines. Some of you might understand that waiting was just not the thing I was keen on in those days."

Brian Harvey, in his position paper (Harvey86) for the ACM Panel on Hacking, claims that the computer medium available to students, e.g., BASIC and floppy disks, is inadequate for challenging intellectual work. His recommendation is that students be given access to real computing power, and that they be taught how to use that power responsibly. He describes a program he created at a public high school in Massachusetts during the period 1979-1982. They installed a PDP-11/70 and let students and teachers carry out the administration of the system. Harvey assessed that putting the burden of dealing with the problems of malicious users on the students themselves was a powerful educational force.

He also noted that the students who had the skill and interest to be password hackers were discouraged from this activity because they also wanted to keep the trust of their colleagues in order that they could acquire "superuser" status on the system.

Harvey also makes an interesting analogy between teaching computing and teaching karate. In karate instruction, students are introduced to the real, adult community. They are given access to a powerful, deadly weapon, and at the same time are taught discipline and responsibility. Harvey speculates that the reason that students do not misuse their power is that they know they are being trusted with something important, and they want to live up to that trust. Harvey applied this principle when he set up the school system.

The ACM panel endorsed Harvey's recommendation, proposing a three-tiered computing environment with local, district-wide, and nation-wide networks. They recommended that computer professionals participate in this effort as mentors and role models. They also recommended that government and industry be encouraged to establish regional computing centers using donated or re-cycled equipment; that students be apprenticed to local companies either part-time on a continuing basis or on a periodic basis; and, following a suggestion from Felsenstein (Felsenstein86) for a "Hacker's League," that a league analogous to the Amateur Radio Relay League be established to make contributed resources available for educational purposes.

Drake said he liked these recommendations. He said that if hackers were given access to powerful systems through a public account system, they would supervise themselves. He also suggested that Computer Resource Centers be established in low-income areas in order to help the poor get access to information. Perhaps hackers could help run the centers and teach the members of the community how to use the facilities. One of my colleagues suggested cynically that the hackers would only use this to teach the poor how to hack rich people's systems. A hacker responded by saying this was ridiculous; hackers would not teach people how to break into systems, but rather how to use computers effectively and not be afraid of them. In addition, the hackers I spoke with who had given up illegal activities said they stopped doing so when they got engaged in other work.

Geoff Goodfellow and Richard Stallman have reported that they have given hackers accounts on systems that they manage, and that the hackers have not misused the trust granted to them. Perhaps universities could consider providing accounts to pre-college students on the basis of recommendations from their teachers or parents. The students might be challenged to work on the same homework problems assigned in courses or to explore their own interests. Students who strongly dislike the inflexibility of classroom learning might excel in an environment that allows them to learn on their own, in much the way that hackers have done.

4 Thrill, Excitement, and Challenge

One hacker wrote that "Hackers understand something basic about computers, and that is that they can be enjoyed. I know none who hack for money, or hack to frighten the company, or hack for anything but fun."

In the words of another hacker, "Hacking was the ultimate cerebral buzz for me. I would come home from another dull day at school, turn my computer on, and become a member of the hacker elite. It was a whole different world where there were no condescending adults and you were judged only by your talent. I would first check in to the private Bulletin Boards where other people who were like me would hang out, see what the news was in the community, and trade some info with people across the country. Then I would start actually hacking. My brain would be going a million miles an hour and I'd basically completely forget about my body as I would jump from one computer to another trying to find a path into my target. It was the rush of working on a puzzle coupled with the high of discovery many magnitudes intensified. To go along with the adrenaline rush was the illicit thrill of doing something illegal. Every step I made could be the one that would bring the authorities crashing down on me. I was on the edge of technology and exploring past it, spelunking into electronic caves where I wasn't supposed to be."

The other hackers I spoke with made similar statements about the fun and challenge of hacking. In *SPIN* magazine (Dibbell90), reporter Julian Dibbell speculated that much of the thrill comes from the dangers asso-

ciated with the activity, writing that "the technology just lends itself to cloak-and-dagger drama," and that "hackers were already living in a world in which covert action was nothing more than a game children played."

Eric Corley (Corley89) characterizes hacking as an evolved form of mountain climbing. In describing an effort to construct a list of active mailboxes on a Voice Messaging System, he writes, "I suppose the main reason I'm wasting my time pushing all these buttons is simply so that I can make a list of something that I'm not supposed to have and be the first person to accomplish this." He said that he was not interested in obtaining an account of his own on the system. Gordon Meyer says he found this to be a recurring theme: "We aren't supposed to be able to do this, but we can"—so they do.

One hacker said he was now working on anti-viral programming. He said it was almost as much fun as breaking into systems, and that it was an intellectual battle against the virus author.

5 Ethics and Avoiding Damage

All of the hackers I spoke with said that malicious hacking was morally wrong. They said that most hackers are not intentionally malicious, and that they themselves are concerned about causing accidental damage. When I asked Drake about the responsibility of a person with a PC and modem, his reply included not erasing or modifying anyone else's data, and not causing a legitimate user on a system any problems. Hackers say they are outraged when other hackers cause damage or use resources that would be missed, even if the results are unintentional and due to incompetence. One hacker wrote: "I have *always* strived to do *no* damage, and to inconvenience as few people as possible. *I never, ever, ever delete a file.* One of the first commands I do on a new system is disable the delete file command." Some hackers say that it is unethical to give passwords and similar security-related information to persons who might do damage. In the recent incident where a hacker broke into BellSouth and downloaded a text file on the emergency 911 service, hackers say that there was no intention to use this knowledge to break into or sabotage the 911 system. According to Emmanuel Goldstein (Goldstein90), the file did not even contain information about how to break into the 911 system.

The hackers also said that some break-ins were unethical, e.g., breaking into hospital systems, and that it is wrong to read confidential information about individuals or steal classified information. All said it was wrong to commit fraud for personal profit.

Although we as computer security professionals often disagree with hackers about what constitutes damage, the ethical standards listed here sound much like our own. Where the hackers' ethics differ from the standards adopted by most in the computer security community is that hackers say it is not unethical to break into many systems, use idle computer and communications resources, and download system files in order to learn. Goldstein says that hacking is not wrong: it is not the same as stealing, and uncovers design flaws and security deficiencies (Goldstein89).

Brian Reid, a colleague at Digital who has spoken with many hackers, speculates that a hacker's ethics may come from not being raised properly as a civilized member of society, and not appreciating the rules of living in society. One hacker responded to this with "What does 'being brought up properly' mean? Some would say that it is 'good' to keep to yourself, mind your own business. Others might argue that it is healthy to explore, take risks, be curious and discover." Brian Harvey (Harvey86) notes that many hackers are adolescents, and that adolescents are at a less advanced stage of moral development than adults, where they might not see how the effects of their actions hurt others. Larry Martin (Martin89) claims that parents, teachers, the press, and others in society are not aware of their responsibility to contribute to instilling ethical values associated with computer use. This could be the consequence of the youth of the computing field; many people are still computer illiterate and cultural norms may be lagging behind advances in technology and the growing dependency on that technology by businesses and society. Hollinger and Lanza-Kaduce (HollingerLanza-Kaduce88) speculate that the cultural normative messages about the use and abuse of computer technology have been driven by the adoption of criminal laws in the last decade. They also speculate that hacking may be encouraged during the process of becoming computer literate. Some of my colleagues say that hackers are

irresponsible. One hacker responded, "I think it's a strong indication of the amount of responsibility shown that so *few* actually *damaging* incidents are known."

But we must not overlook that the differences in ethics also reflect a difference in philosophy about information and information handling resources; whereas hackers advocate sharing, we seem to be advocating ownership as property. The differences also represent an opportunity to examine our own ethical behavior and our practices for information sharing and protection. For example, one hacker wrote, "I will accept that it is morally wrong to copy some proprietary software, however, I think that it is morally wrong to charge $6,000 for a program that is only around 25K long." Hence, I shall go into a few of the ethical points raised by hackers more closely. It is not a simple case of good or mature (us) against bad or immature (hackers), or of teaching hackers a list of rules.

Many computer professionals such as Martin (Martin89) argue the moral questions by analogy. The analogies are then used to justify their judgment of a hacker's actions as unethical. Breaking into a system is compared with breaking into a house, and downloading information and using computer and telecommunications services is compared with stealing tangible goods. But, say hackers, the situations are not the same. When someone breaks into a house, the objective is to steal goods, which are often irreplaceable, and property is often damaged in the process. By contrast, when a hacker breaks into a system, the objective is to learn and avoid causing damage. Downloaded information is copied, not stolen, and still exists on the original system. Moreover, as noted earlier, information has not been traditionally regarded as property. Dibbell (Dibbell90) says that when the software industries and phone companies claim losses of billions of dollars to piracy, they are not talking about goods that disappear from the shelves and could have been sold.

We often say that breaking into a system implies a lack of caring for the system's owner and authorized users. But, one hacker says that the ease of breaking into a system reveals a lack of caring on the part of the system manager to protect user and company assets, or failure on the part of vendors to warn managers about the vulnerabilities of their systems. He estimated his success rate of getting in at 10–15%, and that is without spending more than an hour on any one target system. Another hacker says that he sees messages from vendors notifying the managers, but that the managers fail to take action.

Richard Pethia of CERT (Computer Emergency Response Team) reports that they seldom see cases of malicious damage caused by hackers, but that the break-ins are nevertheless disruptive because system users and administrators want to be sure that nothing was damaged. (CERT suggests that sites reload system software from secure backups and change all user passwords in order to protect against possible back doors and Trojan Horses that might have been planted by the hacker. Pethia also noted that prosecutors are generally called for government sites, and are being called for non-government sites with increasing frequency.) Pethia says that break-ins also generate a loss of trust in the computing environment, and may lead to adoption of new policies that are formulated in a panic or management edicts that severely restrict connectivity to outside systems. Brian Harvey says that hackers cause damage by increasing the amount of paranoia, which in turn leads to tighter security controls that diminish the quality of life for the users. Hackers respond to these points by saying they are the scapegoats for systems that are not adequately protected. They say that the paranoia is generated by ill-founded fears and media distortions (I will return to this point later), and that security need not be oppressive to keep hackers out; it is mainly making sure that passwords and system defaults are well chosen.

Pethia says that some intruders seem to be disruptive to prove a point, such as that the systems are vulnerable, the security personnel are incompetent, or "it's not nice to say bad things about hackers." In the *New York Times,* John Markoff (Markoff90) wrote that the hacker who claimed to have broken into Cliff Stoll's system said he was upset by Stoll's portrayal of hackers in "The Cuckoo's Egg" (Stoll90). Markoff reported that the caller said: "He (Stoll) was going on about how he hates all hackers, and he gave pretty much of a one-sided view of who hackers are."

"The Cuckoo's Egg" captures many of the popular stereotypes of hackers. Criminologist Jim Thomas criticizes it for presenting a simplified view of the world, one where everything springs from the forces of light (us) or of darkness (hackers) (Thomas90). He claims that Stoll fails to see the similarities between his own activities (e.g., monitoring communications, "borrowing" monitors without authorization, shutting off network access

without warning, and lying to get information he wants) and those of hackers. He points out Stoll's use of pejorative words such as "varmint" to describe hackers, and Stoll's quote of a colleague: "They're technically skilled but ethically bankrupt programmers without any respect for others' work—or privacy. They're not destroying one or two programs. They're trying to wreck the cooperation that builds our networks," (Stoll90, p. 159). Thomas writes: "At an intellectual level, it (Stoll's book) provides a persuasive, but simplistic, moral imagery of the nature of right and wrong, and provides what—to a lay reader—would seem a compelling justification for more statutes and severe penalties against the computer underground. This is troublesome for two reasons. First, it leads to a mentality of social control by law enforcement during a social phase when some would argue we are already over-controlled. Second, it invokes a punishment model that assumes we can stamp out behaviors to which we object if only we apprehend and convict a sufficient number of violators. . . . There is little evidence that punishment will in the long run reduce any given offense, and the research of Gordon Meyer and I suggests that criminalization may, in fact, contribute to the growth of the computer underground."

6 Public Image and Treatment

Hackers express concern about their negative public image and identity. As noted earlier, hackers are often portrayed as being irresponsible and immoral. One hacker said that "government propaganda is spreading an image of our being at best, sub-human, depraved, criminally inclined, morally corrupt, low life. We need to prove that the activities that we are accused of (crashing systems, interfering with life support equipment, robbing banks, and jamming 911 lines) are as morally abhorrent to us as they are to the general public."

The public identity of an individual or group is generated in part by the actions of the group interacting with the standards of the community observing those actions. What then accounts for the difference between the hacker's public image and what they say about themselves? One explanation may be the different standards. Outside the hacking community, the simple act of breaking into systems is regarded as unethical by many. The use of pejorative words like "vandal" and "varmint" reflect

this discrepency in ethics. Even the word "criminal" carries with it connotations of someone evil; hackers say they are not criminal in this sense. Katie Hafner notes that Robert Morris, Jr., who was convicted of launching the Internet worm, was likened to a terrorist even though the worm did not destroy data (Hafner90).

Distortions of events and references to potential threats also create an image of persons who are dangerous. Regarding the 911 incident where a hacker downloaded a file from BellSouth, Goldstein reported "Quickly, headlines screamed that hackers had broken into the 911 system and were interfering with emergency telephone calls to the police. One newspaper report said there were no indications that anyone had died or been injured as a result of the intrusions. What a relief. Too bad it wasn't true" (Goldstein90). In fact, the hackers involved with the 911 text file had not broken into the 911 system. The dollar losses attributed to hacking incidents also are often highly inflated.

Thomas and Meyer (ThomasMeyer90) say that the rhetoric depicting hackers as a dangerous evil contributes to a "witch hunt" mentality, wherein a group is first labeled as dangerous, and then enforcement agents are mobilized to exorcise the alleged social evil. They see the current sweeps against hackers as part of a reaction to a broader fear of change, rather than to the actual crimes committed.

Hackers say they are particularly concerned that computer security professionals and system managers do not appear to understand hackers or be interested in their concerns. Hackers say that system managers treat them like enemies and criminals, rather than as potential helpers in their task of making their systems secure. This may reflect managers' fears about hackers, as well as their responsibilities to protect the information on their systems. Stallman says that the strangers he encounters using his account are more likely to have a chip on their shoulder than in the past; he attributes this to a harsh enforcer mentality adopted by the establishment. He says that network system managers start out with too little trust and a hostile attitude toward strangers that few of the strangers deserve. One hacker said that system managers show a lack of openness to those who want to learn.

Stallman also says that the laws make the hacker scared to communicate with anyone even slightly "official," because that person might try to track the hacker down and have him or her arrested. Drake raised the

issue of whether the laws could differentiate between malicious and nonmalicious hacking, in support of a "kinder, gentler" relationship between hackers and computer security people. In fact, many states such as California initially passed computer crime laws that excluded malicious hacking; it was only later that these laws were amended to include nonmalicious actions (HollingerLanza-Kaduce88). Hollinger and Lanza-Kaduce speculate that these amendments and other new laws were catalyzed mainly by media events, especially the reports on the "414 hackers" and the movie *War Games*, which created a perception of hacking as extremely dangerous, even if that perception was not based on facts.

Hackers say they want to help system managers make their systems more secure. They would like managers to recognize and use their knowledge about system vulnerabilities. Landreth (Landreth89) suggests ways in which system managers can approach hackers in order to turn them into colleagues, and Goodfellow also suggests befriending hackers (Goodfellow83). John Draper (Cap'n Crunch) says it would help if system managers and the operators of phone companies and switches could cooperate in tracing a hacker without bringing in law enforcement authorities.

Drake suggests giving hackers free access in exchange for helping with security, a suggestion that I also heard from several hackers. Drake says that the current attitude of treating hackers as enemies is not very conducive to a solution, and by belittling them, we only cause ourselves problems.

I asked some of the hackers whether they'd be interested in breaking into systems if the rules of the "game" were changed so that instead of being threatened by prosecution, they were invited to leave a "calling card" giving their name, phone number, and method of breaking in. In exchange, they would get recognition and points for each vulnerability they discovered. Most were interested in playing; one hacker said he would prefer monetary reward since he was supporting himself. Any system manager interested in trying this out could post a welcome message inviting hackers to leave their cards. This approach could have the advantage of not only letting the hackers contribute to the security of the system, but of allowing the managers to quickly recognize the potentially malicious hackers, since they are unlikely to leave their cards. Perhaps if hackers are given the opportunity to make contributions outside the underground, this will dampen their desire to pursue illegal activities.

Several hackers said that they would like to be able to pursue their activities legally and for income. They like breaking into systems, doing research on computer security, and figuring out how to protect against vulnerabilities. They say they would like to be in a position where they have permission to hack systems. Goodfellow suggests hiring hackers to work on tiger teams that are commissioned to locate vulnerabilities in systems through penetration testing. Baird Info-Systems Safeguards, Inc., a security consulting firm, reports that they have employed hackers on several assignments (Baird87). They say the hackers did not violate their trust or the trust of their clients, and performed in an outstanding manner. Baird believes that system vulnerabilities can be better identified by employing people who have exploited systems.

One hacker suggested setting up a clearinghouse that would match hackers with companies that could use their expertise, while maintaining anonymity of the hackers and ensuring confidentiality of all records. Another hacker, in describing an incident where he discovered a privileged account without a password, said, "What I (and others) wish for is a way that hackers can give information like this to a responsible source, *and have hackers given credit for helping!* As it is, if someone told them that 'I'm a hacker, and I *really* think you should know . . .' they would freak out, and run screaming to the SS (Secret Service) or the FBI. Eventually, the person who found it would be caught, and hauled away on some crazy charge. If they could only just *accept* that the hacker was trying to help!" The clearinghouse could also provide this type of service.

Hackers are also interested in security policy issues. Drake expressed concern over how we handle information about computer security vulnerabilities. He argues that it is better to make this information public than cover it up and pretend that it does not exist, and cites the CERT to illustrate how this approach can be workable. Other hackers, however, argue for restricting initial dissemination of flaws to customers and users. Drake also expressed concern about the role of the government, particularly the military, in cryptography. He argues that NSA's opinion on a cryptographic standard should be taken with a large grain of salt because of their code breaking role.

Some security specialists are opposed to hiring hackers for security work, and Eugene Spafford has urged people not to do business with any

company that hires a convicted hacker to work in the security area (ACM90). He says that "This is like having a known arsonist install a fire alarm." But, the laws are such that a person can be convicted for having done nothing other than break into a system; no serious damage (i.e., no "computer arson") is necessary. Many of our colleagues, including Geoff Goodfellow (Goodfellow83) and Brian Reid (Frenkel87), admit to having broken into systems in the past. Reid is quoted as saying that because of the knowledge he gained breaking into systems as a kid, he was frequently called in to help catch people who break in. Spafford says that times have changed, and that this method of entering the field is no longer socially acceptable, and fails to provide adequate training in computer science and computer engineering (Spafford89). However, from what I have observed, many hackers do have considerable knowledge about telecommunications, data security, operating systems, programming languages, networks, and cryptography. But, I am not challenging a policy to hire competent people of sound character. Rather, I am challenging a strict policy that uses economic pressure to close a field of activity to all persons convicted of breaking into systems. It is enough that a company is responsible for the behavior of its employees. Each hacker can be considered for employment based on his or her own competency and character.

Some people have called for stricter penalties for hackers, including prison terms, in order to send a strong deterrent message to hackers. John Draper, who was incarcerated for his activities in the 1970s, argues that in practice this will only make the problem worse. He told me that he was forced under threat to teach other inmates his knowledge of communications systems. He believes that prison sentences will serve only to spread hacker's knowledge to career criminals. He said he was never approached by criminals outside the prison, but that inside the prison they had control over him.

One hacker said that by clamping down on the hobbyist underground, we will only be left with the criminal underground. He said that without hackers to uncover system vulnerabilities, the holes will be left undiscovered, to be utilized by those likely to cause real damage.

Goldstein argues that the existing penalties are already way out of proportion to the acts committed, and that the reason is because of computers (Goldstein89). He says that if Kevin Mitnick had committed

crimes similar to those he committed but without a computer, he would have been classified as a mischief maker and maybe fined $100 for trespassing; instead, he was put in jail without bail (Goldstein89). Craig Neidorf, a publisher and editor of the electronic newsletter *Phrack,* faces up to 31 years and a fine of $122,000 for receiving, editing, and transmitting the downloaded text file on the 911 system (Goldstein90). (Since the time I wrote this, a new indictment was issued with penalties of up to 65 years in prison. Neidorf went on trial beginning July 23. The trial ended July 27 when the government dropped all charges. DED)

7 Privacy and the First and Fourth Amendments

The hackers I spoke with advocated privacy protection for sensitive information about individuals. They said they are not interested in invading people's privacy, and that they limited their hacking activities to acquiring information about computer systems or how to break into them. There are, of course, hackers who break into systems such as the TRW credit database. Emanuel Goldstein argues that such invasions of privacy took place before the hacker arrived (Harpers90). Referring to credit reports, government files, motor vehicle records, and the "megabytes of data piling up about each of us," he says that thousands of people legally can see and use this data, much of it erroneous. He claims that the public has been misinformed about the databases, and that hackers have become scapegoats for the holes in the systems. One hacker questioned the practice of storing sensitive personal information on open systems with dial-up access, the accrual of the information, the methods used to acquire it, and the purposes to which it is put. Another hacker questioned the inclusion of religion and race in credit records. Drake told me that he was concerned about the increasing amount of information about individuals that is stored in large data banks, and the inability of the individual to have much control over the use of that information. He suggests that the individual might be coowner of information collected about him or her, with control over the use of that information. He also says that an individual should be free to withhold personal information, of course paying the consequences of doing so (e.g., not getting a drivers license or credit card). In fact, all Federal Government forms are required to contain a Privacy Act Statement that states

how the information being collected will be used and, in some cases, giving the option of withholding the information.

Goldstein has also challenged the practices of law enforcement agencies in their attempt to crack down on hackers (Goldstein90). He said that all incoming and outgoing electronic mail used by *Phrack* was monitored before the newsletter was shutdown by authorities: "Had a printed magazine been shut down in this fashion after having all of their mail opened and read, even the most thick-headed sensationalist media types would have caught on: hey, isn't that a violation of the First Amendment?" He also cites the shutdown of several bulletin boards as part of Operation Sun Devil, and quotes the administrator of the bulletin board Zygot as saying "Should I start reading my users' mail to make sure they aren't saying anything naughty? Should I snoop through all the files to make sure everyone is being good? This whole affair is rather chilling." The administrator for the public system The Point wrote, "Today, there is no law or precedent which affords me . . . the same legal rights that other common carriers have against prosecution should some other party (you) use my property (The Point) for illegal activities. That worries me. . . ."

About 40 personal computer systems and 23,000 data disks were seized under Operation Sun Devil, a two-year investigation involving the FBI, Secret Service, and other federal and local law enforcement officials. In addition, the Secret Service acknowledges that its agents, acting as legitimate users, had secretly monitored computer bulletin boards (Markoff90a). Markoff reports that California Representative Don Edwards, industry leader Mitchell Kapor, and civil liberties advocates are alarmed by these government actions, saying that they challenge freedom of speech under the First Amendment and protection against searches and seizures under the Fourth Amendment. Markoff asks: "Will fear of hackers bring oppression?"

John Barlow writes: "The Secret Service may actually have done a service for those of us who love liberty. They have provided us with a devil. And devils, among their other galvanizing virtues, are just great for clarifying the issues and putting iron in your spine" (Barlow90). Some of the questions that Barlow says need to be addressed include: "What are data and what is free speech? How does one treat property which has no physical form and can be infinitely reproduced? Is a computer the same as a printing press?" Barlow urges those of us who understand the technology to address these questions, lest the answers be given to us by law makers and law enforcers who do not. Barlow and Kapor are constituting a foundation to "raise and disburse funds for education, lobbying, and litigation in the areas relating to digital speech and the extension of the Constitution into Cyberspace."

8 Conclusions

Hackers say that it is our social responsibility to share information, and that it is information hoarding and disinformation that are the crimes. This ethic of resource and information sharing contrasts sharply with computer security policies that are based on authorization and "need to know." This discrepancy raises an interesting question: Does the hacker ethic reflect a growing force in society that stands for greater sharing of resources and information—a reaffirmation of basic values in our constitution and laws? It is important that we examine the differences between the standards of hackers, systems managers, users, and the public. These differences may represent breakdowns in current practices, and may present new opportunities to design better policies and mechanisms for making computer resources and information more widely available.

The sentiment for greater information sharing is not restricted to hackers. In the best seller, *Thriving on Chaos,* Tom Peters (Peters87) writes about sharing within organizations: "Information hoarding, especially by politically motivated, power-seeking staffs, has been commonplace throughout American industry, service and manufacturing alike. It will be an impossible millstone around the neck of tomorrow's organizations. Sharing is a must." Peters argues that information flow and sharing is fundamental to innovation and competitiveness. On a broader scale, Peter Drucker (Drucker89) says that the "control of information by government is no longer possible. Indeed, information is now transnational. Like money, it has no 'fatherland.'"

Nor is the sentiment restricted to people outside the computer security field. Harry DeMaio (DeMaio89) says that our natural urge is to share information, and that we are suspicious of organizations and individuals who are secretive. He says that information is exchanged out of "want to

know" and mutual accommodation rather than "need to know." If this is so, then some of our security policies are out of step with the way people work. Peter Denning (DenningP89) says that information sharing will be widespread in the emerging worldwide networks of computers and that we need to focus on "immune systems" that protect against mistakes in our designs and recover from damage.

I began my investigation of hackers with the question, who are they and what is their culture and discourse? My investigation uncovered some of their concerns, which provided the organizational structure to this paper, and several suggestions for new actions that might be taken. My investigation also opened up a broader question: What conflict in society do hackers stand at the battle lines of? Is it owning or restricting information vs. sharing information—a tension between an age-old tradition of controlling information as property and the Englightenment tradition of sharing and disseminating information? Is it controlling access based on "need to know," as determined by the information provider, vs. "want to know," as determined by the person desiring access? Is it law enforcement vs. freedoms granted under the First and Fourth Amendments? The answers to these questions, as well as those raised by Barlow on the nature of information and free speech, are important because they tell us whether our policies and practices serve us as well as they might. The issue is not simply hackers vs. system managers or law enforcers; it is a much larger question about values and practices in an information society.

Acknowledgments

I am deeply grateful to Peter Denning, Frank Drake, Nathan Estey, Katie Hafner, Brian Harvey, Steve Lipner, Teresa Lunt, Larry Martin, Gordon Meyer, Donn Parker, Morgan Schweers, Richard Stallman, and Alex for their comments on earlier versions of this paper and helpful discussions; to Richard Stallman for putting me in contact with hackers; John Draper, Geoff Goodfellow, Brian Reid, Eugene Spafford, Dave, Marcel, Mike, RGB, and the hackers for helpful discussions; and Richard Pethia for a summary of some of his experiences at CERT. The opinions expressed here, however, are my own and do not necessarily represent those of the people mentioned above or of Digital Equipment Corporation.

References

ACM90 "Just Say No," *Comm. ACM* 33, no. 5, May 1990, p. 477.

Baird87 Bruce J. Baird, Lindsay L. Baird, Jr., and Ronald P. Ranauro, "The Moral Cracker?" *Computers and Security* 6, no. 6, December 1987, pp. 471–478.

Barlow90 John Barlow, "Crime and Puzzlement," June 1990, to appear in *Whole Earth Review*. [Appendix 1 in this volume.]

Corley89 Eric Corley, "The Hacking Fever," in Pamela Kane, *V.I.R.U.S. Protection,* Bantam Books, New York, 1989, pp. 67–72.

DeMaio89 Harry B. DeMaio, "Information Ethics, a Practical Approach," *Proc. of the 12th National Computer Security Conference, 1989,* pp. 630–633.

DenningP89 Peter J. Denning, "Worldnet," *American Scientist* 77, no. 5, Sept.-Oct. 1989.

DenningP90 Peter J. Denning, *Computers Under Attack,* ACM Press, 1990.

Dibbell90 Julian Dibbell, "Cyber Thrash," *SPIN 5*, no. 12, March 1990.

Drucker89 Peter F. Drucker, *The New Realities,* Harper and Row, New York, 1989.

Felsenstein86 Lee Felsenstein, "Real Hackers Don't Rob Banks," in full report on ACM Panel on Hacking (Lee86).

Frenkel87 Karen A. Frenkel, "Brian Reid, A Graphics Tale of a Hacker Tracker," *Comm. ACM* 30, no. 10, October 1987, pp. 820–823.

Goldstein89 Emmanuel Goldstein, "Hackers in Jail," *2600 Magazine* 6, no. 1, Spring 1989.

Goldstein90 Emmanuel Goldstein, "For Your Protection," *2600 Magazine* 7, no. 1, Spring 1990.

Goodfellow83 Geoffrey S. Goodfellow, "Testimony Before the Subcommittee on Transportation, Aviation, and Materials on the Subject of Telecommunications Security and Privacy," Sept. 26, 1983.

Hafner90 Katie Hafner, "Morris Code," *New Republic,* February 16, 1990, pp. 15–16.

Harpers90 "Is Computer Hacking a Crime?" *Harper's,* March 1990, pp. 45–57.

Harvey86 Brian Harvey, "Computer Hacking and Ethics," in full report on ACM Panel on Hacking (Lee86).

HollingerLanza-Kaduce88 Richard C. Hollinger and Lonn Lanza-Kaduce, "The Process of Criminalization: The Case of Computer Crime Laws," *Criminology* 26, no. 1, 1988, pp. 101–126.

Huebner89 Hans Huebner, "Re: News from the KGB/Wiley Hackers," *RISKS Digest* 8, no. 37, 1989.

Landreth89 Bill Landreth, *Out of the Inner Circle,* Tempus, Redmond, WA, 1989.

Lee86 John A. N. Lee, Gerald Segal, and Rosalie Stier, "Positive Alternatives: A Report on an ACM Panel on Hacking," *Comm. ACM* 29, no. 4, April 1986, pp. 297–299; full report available from ACM Headquarters, New York.

Levy84 Steven Levy, *Hackers,* Dell, New York, 1984.

Markoff90 John Markoff, "Self-Proclaimed 'Hacker' Sends Message to Critics," *New York Times,* March 19, 1990.

Markoff90a John Markoff, "Drive to Counter Computer Crime Aims at Invaders," *New York Times,* June 3, 1990.

Martin89 Larry Martin, "Unethical 'Computer' Behavior: Who Is Responsible?" *Proc. of the 12th National Computer Security Conference, 1989.*

Meyer89 Gordon R. Meyer, The Social Organization of the Computer Underground, Master's thesis, Dept. of Sociology, Northern Illinois Univ., Aug. 1989.

MeyerThomas90 Gordon Meyer and Jim Thomas, "The Baudy World of the Byte Bandit: A Postmodernist Interpretation of the Computer Underground," Dept. of Sociology, Northern Illinois Univ., DeKalb, IL, March 1990.

Peters87 Tom Peters, *Thriving on Chaos,* Harper & Row, New York, Chapter VI, S-3, p. 610, 1987.

Spafford89 Eugene H. Spafford, "The Internet Worm, Crisis and Aftermath," *Comm. ACM* 32, no. 6, June 1989, pp. 678–687.

Stallman84 Richard M. Stallman, Letter to ACM Forum, *Comm. ACM* 27, no. 1, January 1984, pp. 8–9.

Stallman90 Richard M. Stallman, "Against User Interface Copyright" to appear in *Comm. ACM.*

Steele83 Guy L. Steele, Jr., Donald R. Woods, Raphael A. Finkel, Mark R. Crispin, Richard M. Stallman, and Geoffrey S. Goodfellow, *The Hacker's Dictionary,* Harper & Row, New York, 1983.

Stoll90 Clifford Stoll, *The Cuckoo's Egg,* Doubleday, 1990.

Thomas90 Jim Thomas, "Review of *The Cuckoo's Egg,*" *Computer Underground Digest* 1, no. 6, April 27, 1990.

ThomasMeyer90 Jim Thomas and Gordon Meyer, "Joe McCarthy in a Leisure Suit: (Witch)Hunting for the Computer Underground," Unpublished manuscript, Department of Sociology, Northern Illinois University, DeKalb, IL, 1990; see also the *Computer Underground Digest* 1, no. 11, June 16, 1990.

Postscript, June 11, 1995

After completing the article five years ago, I interviewed people in law enforcement and industry who investigated cases of system intrusion. I found that many of the claims made by hackers were not substantiated by the evidence collected and that with few exceptions, the cases were handled competently and professionally. First and Fourth Amendment rights were not being trampled, and the issue was not law enforcement vs. civil liberties. As a result of my continued research, I developed a better understanding of all sides of the hacker issue, and came to disagree with some of my earlier interpretations and conclusions. The purpose of this postscript is to summarize some of my current thoughts on hackers.

Hacking is a serious and costly problem. Even when there is no malicious intent, intrusions can be extremely disruptive if not outright damaging. A system administrator must assess whether passwords or sensitive information might have been compromised, check for altered files and Trojan horses, and, when necessary, restore the system to a previous "safe" state or change passwords. A system might be down for hours or more than a day while these activities take place. At one university I know, a full-time person is needed just to respond to intruders. Hackers either do not appreciate the consequences of their "non-malicious" hacking on system administrators and users, or else they deny these negative effects in order to justify their actions.

Hackers place responsibility for their intrusions on system developers and administrators for not making their systems secure. They do not seem to appreciate that security is only one factor that must be considered in the design and operation of a system. Real-world requirements, constraints, and budgets can lead to tradeoffs with other factors such as ease of use, network access, development time, and system or administration overhead. One system administrator I know spends about a third of his time keeping up with and responding to security threats. That is time that otherwise could be spent installing new software or making other improvement to the system. Even when security is of high priority, it is difficult to fully achieve since new designs and protocols can introduce new vulnerabilities. In one recent case, a network security tool (SATAN) that had been developed by security experts to detect vulnerabilities was found to introduce one of its own. I do not mean to suggest that system developers, administrators, and users have no responsibility for making their systems secure, but rather that those who carry out an attack are responsible for the attack itself in the same way that robbers and other criminals are responsible for their deeds. It is unrealistic to expect or demand that all systems will be fully secure.

In placing the blame for their intrusions on their victims, hackers fail to acknowledge how their own actions have contributed to the security problem. They spread knowledge about how to penetrate systems through electronic publications and bulletin board systems, and by teaching novices. The current issue of *Phrack* (vol. 6, no. 47), for example, contains articles on how to crack Unix and VMS passwords, gain root access, erase one's tracks from system logs, send fake mail, and defeat copy protection. Many articles contain code for implementing an attack or point the reader to sites where penetration software can be downloaded and run. Many attacks have been sufficiently automated that novices can perform them with little effort or understanding of the systems they are attacking.

Hackers justify their illegal or unethical actions by appealing to the First Amendment and by claiming that the vulnerabilities they find need to be widely exposed lest they be exploited by "real criminals" or "malicious hackers." In fact, information disseminated through hacker publications and bulletin boards has frequently been used to commit serious crimes, with losses sometimes reaching millions of dollars. Hackers do not acknowledge the value of information to those that produce it (even while jealously guarding access to some of their own files), using the hacker ethic that "all information should be free" as a convenient rationale for disseminating whatever they please. They do not distinguish between the dissemination of information about system vulnerabilities and attacks for the purpose of preventing attacks vs. performing them, a distinction that leads to considerably different articles and publications (e.g., CERT advisories vs. *Phrack*'s hacker tutorials). Hackers do not see that in many cases, they are the biggest threat. Were it not for hackers, many systems might never be attacked despite their weaknesses, just as many of us are never robbed even though we are vulnerable.

I do not have a solution to the hacker problem, but I no longer recommend working closely with hackers towards one. I doubt that many hackers have any serious interest in seeing their attacks successfully thwarted, as it would destroy a "game" they enjoy. Moreover, working with people who flagrantly violate the law sends the wrong message and rewards the wrong behavior. Computer ethics education might deter some potential hackers, but it will not deter those hackers who are

determined to pursue their trade and take advantage of computer networks to spread their knowledge far and wide. Better security and law enforcement are the best approaches, so that the chances of penetration are reduced while those for detection and prosecution are increased. However, neither will solve the problem completely. There is no "silver bullet" that will stop hacking.

Are Computer Hacker Break-ins Ethical?*

Eugene H. Spafford
Department of Computer Sciences
Purdue University
West Lafayette, IN 47907–1398
spaf@cs.purdue.edu

Abstract

Recent incidents of unauthorized computer intrusion have brought about discussion of the ethics of breaking into computers. Some individuals have argued that as long as no significant damage results, break-ins may serve a useful purpose. Others counter with the expression that the break-ins are almost always harmful and wrong.

This article lists and refutes many of the reasons given to justify computer intrusions. It is the author's contention that break-ins are ethical only in extreme situations, such as a life-critical emergency. The article also discusses why no break-in is "harmless."

1 Introduction

On November 2, 1988, a program was run on the Internet that replicated itself on thousands of machines, often loading them to the point where they were unable to process normal requests. [1, 2, 3] This *Internet Worm* program was stopped in a matter of hours, but the controversy engendered by its release raged for years. Other recent incidents, such as the "wily hackers"[1] tracked by Cliff Stoll [4], the "Legion of Doom" members who are alleged to have stolen telephone company 911 software [5], and the growth of the computer virus problem [6, 7, 8, 9] have added to the discussion. What constitutes improper access to computers? Are some break-ins ethical? Is there such a thing as a "moral hacker"?[10]

It is important that we discuss these issues. The continuing evolution of our technological base and our increasing reliance on computers for critical tasks suggests that future incidents may well

 See page 12 for publication history.

[1] I realize that many law-abiding individuals consider themselves *hackers* — a term formerly used as a compliment. The press and general public have co-opted the term, however, and it is now commonly viewed as a pejorative. Here, I will use the word as the general public now uses it.

have more serious consequences than those we have seen to date. With human nature as varied and extreme as it is, and with the technology as available as it is, we must expect to experience more of these incidents.

In this article, I will introduce a few of the major issues that these incidents have raised, and present some arguments related to them. For clarification, I have separated a few issues that often have been combined when debated; it is possible that most people are in agreement on some of these points once they are viewed as individual issues.

2 What is Ethical?

Webster's Collegiate Dictionary defines ethics as: "The discipline dealing with what is good and bad and with moral duty and obligation." More simply, it is the study of what is *right* to do in a given situation—what we *ought* to do. Alternatively, it is sometimes described as the study of what is *good* and how to achieve that good. To suggest whether an act is right or wrong, we need to agree on an ethical system that is easy to understand and apply as we consider the ethics of computer break-ins.

Philosophers have been trying for thousands of years to define right and wrong, and I will not make yet another attempt at such a definition. Instead, I will suggest that we make the simplifying assumption that we can judge the ethical nature of an act by applying a deontological assessment: regardless of the effect, is the act itself ethical? Would we view that act as sensible and proper if **everyone** were to engage in it? Although this may be too simplistic a model (and it can certainly be argued that other ethical philosophies may also be applied), it is a good first approximation for purposes of discussion. If you are unfamiliar with any other formal ethical evaluation method, try applying this assessment to the points I raise later in this paper. If the results are obviously unpleasant or dangerous in the large, then they should be considered unethical as individual acts.

Note that this philosophy assumes that *right* is determined by actions and not by results. Some ethical philosophies assume that the ends justify the means; our current society does not operate by such a philosophy, although many individuals do. As a society, we profess to believe that "it isn't whether you win or lose, it's how you play the game." This is why we are concerned with issues of due process and civil rights, even for those espousing repugnant views and committing heinous acts. The process is important no matter the outcome, although the outcome may help to resolve a choice between two almost equal courses of action.

Philosophies that consider the results of an act as the ultimate measure of good are often impossible to apply because of the difficulty in understanding exactly what results from any arbitrary activity. Consider an extreme example: the government orders a hundred cigarette smokers, chosen at random, to be beheaded on live nationwide television. The result might well be that many hundreds of thousands of other smokers would quit "cold turkey," thus prolonging their lives. It might also prevent hundreds of thousands of people from ever starting to smoke, thus improving the health and longevity of the general populace. The health of millions of other people would improve as they

would no longer be subjected to secondary smoke, and the overall impact on the environment would be very favorable as tons of air and ground pollutants would no longer be released by smokers or tobacco companies.

Yet, despite the great good this might hold for society, everyone, except for a few extremists, would condemn such an *act* as immoral. We would likely object even if only one person was executed. It would not matter what the law might be on such a matter; we would not feel that the act was morally correct, nor would we view the ends as justifying the means.

Note that we would be unable to judge the morality of such an action by evaluating the results, because we would not know the full scope of those results. Such an act might have effects favorable or otherwise, on issues of law, public health, tobacco use, and daytime TV shows for decades or centuries to follow. A system of ethics that considered primarily only the results of our actions would not allow us to evaluate our current activities at the time when we would need such guidance; if we are unable to discern the appropriate course of action prior to its commission, then our system of ethics is of little or no value to us. To obtain ethical guidance, we must base our actions primarily on evaluations of the actions and not on the possible results.

More to the point of this paper, if we attempt to judge the morality of a computer break-in based on the sum total of all future effect, we would be unable to make such a judgement, either for a specific incident or for the general class of acts. In part, this is because it is so difficult to determine the long-term effects of various actions, and to discern their causes. We cannot know, for instance, if increased security awareness and restrictions are better for society in the long-term, or whether these additional restrictions will result in greater costs and annoyance when using computer systems. We also do not know how many of these changes are directly traceable to incidents of computer break-ins.

One other point should be made here: it is undoubtedly possible to imagine scenarios where a computer break-in would be considered to be the preferable course of action. For instance, if vital medical data were on a computer and necessary to save someone's life in an emergency, but the authorized users of the system cannot be located, breaking into the system might well be considered the right thing to do. However, that action does not make the break-in ethical. Rather, such situations occur when a greater wrong would undoubtedly occur if the unethical act were not committed. Similar reasoning applies to situations such as killing in self-defense. In the following discussion, I will assume that such conflicts are not the root cause of the break-ins; such situations should very rarely present themselves.

3 Motivations

Individuals who break into computer systems or who write *vandalware* usually use one of a few rationalizations for their actions. (See, for example, [11] and the discussion in [12].) Most of these individuals would never think to walk down a street, trying every door to find one unlocked, then search through the drawers of the furniture inside. Yet, these same people seem to give no second

thought to making repeated attempts at guessing passwords to accounts they do not own, and once on to a system, browsing through the files on disk.

These computer burglars often present the same reasons for their actions in an attempt to rationalize their activities as morally justified. I present and refute some of the most commonly used ones in what follows; motives involving theft and revenge are not uncommon, and their moral nature is simple to discern, so I shall not include them here.

3.1 The Hacker Ethic

Many hackers argue that they follow an ethic that both guides their behavior and justifies their break-ins. This hacker ethic states, in part, that all information should be free.[10] This view holds that information belongs to everyone, and there should be no boundaries or restraints to prevent anyone from examining information. Richard Stallman states much the same thing in his GNU Manifesto.[13] He and others have further stated in various forums that if information is free, it logically follows that there should be no such thing as intellectual property, and no need for security.

What are the implications and consequences of such a philosophy? First and foremost, it raises some disturbing questions of privacy. If all information is (or should be) free, then privacy is no longer a possibility. For information to be free to everyone, and for individuals to no longer be able to claim it as property, means that anyone may access the information if they please. Furthermore, as it is no longer property of any individual, that means that anyone can alter the information. Items such as bank balances, medical records, credit histories, employment records, and defense information all cease to be controlled. If someone controls information and controls who may access it, the information is obviously not free. But without that control, we would no longer be able to trust the accuracy of the information.

In a perfect world, this lack of privacy and control might not be a cause for concern. However, if all information were to be freely available and modifiable, imagine how much damage and chaos would be caused in our real world by such a philosophy! Our whole society is based on information whose accuracy must be assured. This includes information held by banks and other financial institutions, credit bureaus, medical agencies and professionals, government agencies such as the IRS, law enforcement agencies, and educational institutions. Clearly, treating all their information as "free" would be unethical in any world where there might be careless and unethical individuals.

Economic arguments can be made against this philosophy, too, in addition to the overwhelming need for privacy and control of information accuracy. Information is not universally free. It is held as property because of privacy concerns, and because it is often collected and developed at great expense. Development of a new algorithm or program, or collection of a specialized database, may involve the expenditure of vast sums of time and effort. To claim that it is free or should be free is to express a naive and unrealistic view of the world. To use this as a justification for computer break-ins is clearly unethical. Although not all information currently treated as private or controlled as proprietary needs such protection, that does not justify unauthorized access to it or to any other data.

3.2 The Security Arguments

These arguments are the most common ones within the computer community. One common argument was the same one used most often by people attempting to defend the author of the Internet Worm program in 1988: break-ins illustrate security problems to a community that will otherwise not note the problems.

In the Worm case, one of the first issues to be discussed widely in Internet mailing lists dealt with the intent of the perpetrator — exactly why the worm program had been written and released. Explanations put forth by members of the community ranged from simple accident to the actions of a sociopath. A common explanation was that the Worm was designed to illustrate security defects to a community that would not otherwise pay attention. This was not supported by the testimony during the author's trial, nor is it supported by past experience of system administrators.

The Worm author, Robert T. Morris, appears to have been well-known at some universities and major companies, and his talents were generally respected. Had he merely explained the problems or offered a demonstration to these people, he would have been listened to with considerable attention. The month before he released the Worm program on the Internet, he discovered and disclosed a bug in the file transfer program *ftp*; news of the flaw spread rapidly, and an official fix was announced and available within a matter of weeks. The argument that no one would listen to his report of security weaknesses is clearly fallacious.

In the more general case, this security argument is also without merit. Although some system administrators might have been complacent about the security of their systems before the Worm incident, most computer vendors, managers of government computer installations, and system administrators at major colleges and universities have been attentive to reports of security problems. People wishing to report a problem with the security of a system need not exploit it to report it. By way of analogy, one does not set fire to the neighborhood shopping center to bring attention to a fire hazard in one of the stores, and then try to justify the act by claiming that firemen would otherwise never listen to reports of hazards.

The most general argument that some people make is that the individuals who break into systems are performing a service by exposing security flaws, and thus should be encouraged or even rewarded. This argument is severely flawed in several ways. First, it assumes that there is some compelling need to force users to install security fixes on their systems, and thus *computer burglars* are justified in "breaking and entering" activities. Taken to extremes, it suggests that it would be perfectly acceptable to engage in such activities on a continuing basis, so long as they might expose security flaws. This completely loses sight of the purpose of the computers in the first place — to serve as tools and resources, not as exercises in security. The same reasoning would imply that vigilantes have the right to attempt to break into the homes in my neighborhood on a continuing basis to demonstrate that they are susceptible to burglars.

Another flaw with this argument is that it completely ignores the technical and economic factors that prevent many sites from upgrading or correcting their software. Not every site has the resources to install new system software or to correct existing software. At many sites, the systems are run as

turnkey systems — employed as tools and maintained by the vendor. The owners and users of these machines simply do not have the ability to correct or maintain their systems independently, and they are unable to afford custom software support from their vendors. To break into such systems, with or without damage, is effectively to trespass into places of business; to do so in a vigilante effort to force the owners to upgrade their security structure is presumptuous and reprehensible. A burglary is not justified, morally or legally, by an argument that the victim has poor locks and was therefore "asking for it."

A related argument has been made that vendors are responsible for the maintenance of their software, and that such security breaches should immediately require vendors to issue corrections to their customers, past and present. The claim is made that without highly-visible break-ins, vendors will not produce or distribute necessary fixes to software. This attitude is naive, and is neither economically feasible nor technically workable. Certainly, vendors should bear some responsibility for the adequacy of their software,[14] but they should not be responsible for fixing every possible flaw in every possible configuration.

Many sites customize their software or otherwise run systems incompatible with the latest vendor releases. For a vendor to be able to provide quick response to security problems, it would be necessary for each customer to run completely standardized software and hardware mixes to ensure the correctness of vendor-supplied updates. Not only would this be considerably less attractive for many customers and contrary to their usual practice, but the increased cost of such "instant" fix distribution would add to the price of such a system — greatly increasing the cost borne by the customer. It is unreasonable to expect the user community to sacrifice flexibility **and** pay a much higher cost per unit simply for faster corrections to the occasional security breach. That assumes it was even possible for the manufacturer to find those customers and supply them with fixes in a timely manner, something unlikely in a market where machines and software are often repackaged, traded, and resold.

The case of the Internet Worm is a good example of the security argument and its flaws. It further stands as a good example of the conflict between ends and means valuation of ethics. Various people have argued that the Worm's author did us a favor by exposing security flaws. At Mr. Morris's trial on Federal charges stemming from the incident, the defense attorneys also argued that their client should not be punished because of the good the Worm did in exposing those flaws. Others, including the prosecuting attorneys for the government, argued that the act itself was wrong no matter what the outcome. Their contention has been that the result does not justify the act itself, nor does the defense's argument encompass all the consequences of the incident.

This is certainly true; the complete results of the incident are still not known. There have been many other break-ins and network worms since November 1988, perhaps inspired by the media coverage of that incident. More attempts will possibly be made, in part inspired by Mr. Morris's act. Some sites on the Internet have restricted access to their machines, and others were removed from the network; I have heard of sites where a decision has been made not to pursue a connection, even though this will hinder research and operations. Combined with the many decades of person-hours devoted to cleaning up afterwards, this seems to be a high price to pay for a claimed "favor."

The legal consequences of this act are also not yet known. For instance, many bills were introduced into Congress and state legislatures in subsequent years as a (partial) result of these incidents. One piece of legislation introduced into the House of Representatives, HR-5061, entitled "The Computer Virus Eradication Act of 1988," was the first in a series of legislative actions that had the potential to affect significantly the computer profession. In particular, HR-5061 was notable because its wording would have prevented it from being applied to true computer viruses,[2] The passage of similar well-intentioned but poorly-defined legislation could have a major negative effect on the computing profession as a whole.

3.3 The Idle System Argument

Another argument put forth by system hackers is that they are simply making use of idle machines. They argue that because some systems are not used at any level near their capacity, the hacker is somehow entitled to use them.

This argument is also flawed. First of all, these systems are usually not in service to provide a general-purpose user environment. Instead, they are in use in commerce, medicine, public safety, research, and government functions. Unused capacity is present for future needs and sudden surges of activity, not for the support of outside individuals. Imagine if large numbers of people without a computer were to take advantage of a system with idle processor capacity: the system would quickly be overloaded and severely degraded or unavailable for the rightful owners. Once on the system, it would be difficult (or impossible) to oust these individuals if sudden extra capacity was needed by the rightful owners. Even the largest machines available today would not provide sufficient capacity to accommodate such activity on any large scale.

I am unable to think of any other item that someone may buy and maintain, only to have others claim a right to use it when it is idle. For instance, the thought of someone walking up to my expensive car and driving off in it simply because it is not currently being used is ludicrous. Likewise, because I am away at work, it is not proper to hold a party at my house because it is otherwise not being used. The related positions that unused computing capacity is a shared resource, and that my privately-developed software belongs to everyone, are equally silly (and unethical) positions.

3.4 The Student Hacker Argument

Some trespassers claim that they are doing no harm and changing nothing — they are simply learning about how computer systems operate. They argue that computers are expensive, and that they are merely furthering their education in a cost-effective manner. Some authors of computer viruses claim that their creations are intended to be harmless, and that they are simply learning how to write complex programs.

[2] It provided penalties only in cases where **programs** were introduced into computer systems; a computer virus is a segment of code attached to an existing program that modifies other programs to include a copy of itself.[6]

There are many problems with these arguments. First, as an educator, I claim that writing vandalware or breaking into a computer and looking at the files has almost nothing to do with computer education. Proper education in computer science and engineering involves intensive exposure to fundamental aspects of theory, abstraction, and design techniques. Browsing through a system does not expose someone to the broad scope of theory and practice in computing, nor does it provide the critical feedback so important to a good education (cf. [15, 16]). Neither does writing a virus or worm program and releasing it into an unsupervised environment provide any proper educational experience. By analogy, stealing cars and joyriding does not provide one with an education in mechanical engineering, nor does pouring sugar in the gas tank.

Furthermore, individuals "learning" about a system cannot know how everything operates and what results from their activities. Many systems have been damaged accidentally by ignorant (or careless) intruders; most of the damage from computer viruses (and the Internet Worm) appear to be caused by unexpected interactions and program faults. Damage to medical systems, factory control, financial information, and other computer systems could have drastic and far-ranging effects that have nothing to do with education, and could certainly not be considered harmless.

A related refutation of the claim has to do with knowledge of the extent of the intrusion. If I am the person responsible for the security of a critical computer system, I cannot assume that *any* intrusion is motivated solely by curiosity and that nothing has been harmed. If I know that the system has been compromised, I must fear the worst and perform a complete system check for damages and changes. I cannot take the word of the intruder, for any intruder who actually caused damage would seek to hide it by claiming that he or she was "just looking." In order to regain confidence in the correct behavior of my system, I must expend considerable energy to examine and verify every aspect of it.

Apply our universal approach to this situation and imagine if this "educational" behavior was widespread and commonplace. The result would be that we would spend all our time verifying our systems and never be able to trust the results fully. Clearly, this is not good, and thus we must conclude that these "educational" motivations are also unethical.

3.5 The Social Protector Argument

One last argument, more often heard in Europe than the U.S. is that hackers break into systems to watch for instances of data abuse and to help keep "Big Brother" at bay. In this sense, the hackers are protectors rather than criminals. Again, this assumes that the ends justify the means. It also assumes that the hackers are actually able to achieve some good end.

Undeniably, there is some misuse of personal data by corporations and by the government. The increasing use of computer-based record systems and networks may lead to further abuses. However, it is not clear that breaking into these systems will aid in righting the wrongs. If anything, it will cause those agencies to become even more secretive and use the break-ins as an excuse for more restricted access. Break-ins and vandalism have not resulted in new open-records laws, but they have resulted in the introduction and passage of new criminal statutes. Not only has such

activity failed to deter "Big Brother," but it has also resulted in significant segments of the public urging more laws and more aggressive law enforcement — the direct opposite of the supposed goal.

It is also not clear that these are the individuals we want "protecting" us. We need to have the designers and users of the systems — trained computer professionals — concerned about our rights and aware of the dangers involved with the inappropriate use of computer monitoring and record-keeping. The threat is a relatively new one, as computers and networks have become widely used only in the last few decades. It will take some time for awareness of the dangers to spread throughout the profession. Clandestine efforts to breach the security of computer systems do nothing to raise the consciousness of the appropriate individuals. Worse, they associate that commendable goal (heightened concern) with criminal activity (computer break-ins), discouraging proactive behavior by the individuals in the best positions to act in our favor. Perhaps it is in this sense that computer break-ins and vandalism are most unethical and damaging.

4 Concluding Remarks

I have argued here that computer break-ins, even when no obvious damage results, are unethical. This must be the considered conclusion even if the result is an improvement in security, because the activity itself is disruptive and immoral. The results of the act should be considered separately from the act itself, especially when we consider how difficult it is to understand all the effects resulting from such an act.

Of course, I have not discussed every possible reason for a break-in. There might well be an instance where a break-in might be necessary to save a life or to preserve national security. In such cases, to perform one wrong act to prevent a greater wrong may be the right thing to do. It is beyond the scope or intent of this paper to discuss such cases, especially as no known hacker break-ins have been motivated by such instances.

Historically, computer professionals as a group have not been overly concerned with questions of ethics and propriety as they relate to computers. Individuals and some organizations have tried to address these issues, but the whole computing community needs to be involved to address the problems in any comprehensive manner. Too often, we view computers simply as machines and algorithms, and we do not perceive the serious ethical questions inherent in their use.

When we consider, however, that these machines influence the quality of life of millions of individuals, both directly and indirectly, we understand that there are broader issues. Computers are used to design, analyze, support, and control applications that protect and guide the lives and finances of people. Our use (and misuse) of computing systems may have effects beyond our wildest imagining. Thus, we must reconsider our attitudes about acts demonstrating a lack of respect for the rights and privacy of other people's computers and data.

We must also consider what our attitudes will be towards future security problems. In particular, we should consider the effect of **widely** publishing the source code for worms, viruses, and other threats to security. Although we need a process for rapidly disseminating corrections and security

information as they become known, we should realize that widespread publication of details will imperil sites where users are unwilling or unable to install updates and fixes.[3] Publication should serve a useful purpose; endangering the security of other people's machines or attempting to force them into making changes they are unable to make or afford is not ethical.

Finally, we must decide these issues of ethics as a community of professionals and then present them to society as a whole. No matter what laws are passed, and no matter how good security measures might become, they will not be enough for us to have completely secure systems. We also need to develop and act according to some shared ethical values. The members of society need to be educated so that they understand the importance of respecting the privacy and ownership of data. If locks and laws were all that kept people from robbing houses, there would be many more burglars than there are now; the shared mores about the sanctity of personal property are an important influence in the prevention of burglary. It is our duty as informed professionals to help extend those mores into the realm of computing.

References

[1] Donn Seeley. A tour of the worm. In *Proceedings of the Winter 1989 Usenix Conference*. The Usenix Association, January 1989.

[2] Eugene H. Spafford. The Internet Worm: Crisis and aftermath. *Communications of the ACM*, 32(6):678–698, June 1989.

[3] Eugene H. Spafford. An analysis of the Internet Worm. In C. Ghezzi and J. A. McDermid, editors, *Proceedings of the 2nd European Software Engineering Conference*, pages 446–468. Springer-Verlag, September 1989.

[4] Clifford Stoll. *Cuckoo's Egg*. Doubleday, New York, NY, 1989.

[5] John Schwartz. The hacker dragnet. *Newsweek*, 65(18), April 1990.

[6] Eugene H. Spafford, Kathleen A. Heaphy, and David J. Ferbrache. *Computer Viruses: Dealing with Electronic Vandalism and Programmed Threats*. ADAPSO, Arlington, VA, 1989.

[7] Lance Hoffman, editor. *Rogue Programs: Viruses, Worms, and Trojan Horses*. Van Nostrand Reinhold, 1990.

[8] David J. Stang. *Computer Viruses*. National Computer Security Association, Washington, DC, 2nd edition edition, March 1990.

[3] To anticipate the oft-used comment that the "bad guys" already have such information: not every computer burglar knows or will know *every* system weakness — unless we provide them with detailed analyses.

[9] Peter J. Denning, editor. *Computers Under Attack: Intruders, Worms, and Viruses*. ACM Books/Addison-Wesley, 1991.

[10] Bruce J. Baird, Lindsay L. Baird Jr., and Ronald P. Ranauro. The moral cracker? *Computers and Security*, 6(6):471–478, December 1987.

[11] Bill Landreth. *Out of the Inner Circle: a Hacker's Guide to Computer Security*. Microsoft Press, New York, 1984.

[12] Adelaide, John Perry Barlow, Robert Jacobson Bluefire, Russell Brand, Clifford Stoll, Dave Hughes, Frank Drake, Eddie Joe Homeboy, Emmanuel Goldstein, Hank Roberts, Jim Gasperini JIMG, Jon Carroll JRC, Lee Felsenstein, Tom Mandel, Robert Horvitz RH, Richard Stallman RMS, Glenn Tenney, Acid Phreak, and Phiber Optik. Is computer hacking a crime? *Harper's Magazine*, 280(1678):45–57, March 1990.

[13] Richard Stallman. *GNU EMacs Manual*, chapter The GNU Manifesto, pages 239–248. Free Software Foundation, 1986.

[14] M. Douglas McIlroy. Unsafe at any price. *Information Technology Quarterly*, IX(2):21–23, 1990.

[15] P. J. Denning, D. E. Comer, D. Gries, M. C. Mulder, A. Tucker, A. J. Turner, and P. R. Young. Computing as a discipline. *Communications of the ACM*, 32(1):9–23, January 1989.

[16] Allen B. Tucker, Bruce H. Barnes, Robert M. Aiken, Keith Barker, Kim B. Bruce, J. Thomas Cain, Susan E. Conry, Gerald L. Engel, Richard G. Epstein, Doris K. Lidtke, Michael C. Mulder, Jean B. Rogers, Eugene H. Spafford, and A. Joe Turner. Computing curricula 1991, 1991. Published by the IEEE Society Press.

[17] Eugene H. Spafford. Is a computer break-in ever ethical? *Information Technology Quarterly*, IX(2):9–14, 1990.

[18] Eugene H. Spafford. Are computer hacker break-ins ethical? *Journal of Systems and Software*, 17(1):41–48, January 1992.

About the Author

Gene Spafford received a Ph.D. in 1986 from the School of Information and Computer Sciences at Georgia Institute of Technology. In 1987, Professor Spafford joined the faculty of the Department of Computer Sciences at Purdue University. He is an active researcher with the NSF/Purdue/University of Florida Software Engineering Research Center (SERC) there.

Besides a continuing widely-known and respected involvement in the Usenet and other forms of electronic conferencing, Dr. Spafford does research on issues relating to increasing the reliability of computer systems, and the consequences of computer failures. This includes work with computer security, software testing and debugging, and issues of liability and professional ethics.

Professor Spafford's involvement with software engineering and security include positions on the editorial boards of the journals *International Journal of Computer and Software Engineering*, *Computers & Security*, the *Virus Bulletin*, the *Journal of Artificial Life*, *Network Security*, and the *Journal of Information Systems Security*. He is coauthor (with S. Garfinkel) of the books *Practical Unix and Internet Security*, published by O'Reilly and Associates (1991, 1996), and of *Computer Viruses: Dealing with Electronic Vandalism and Programmed Threats*, ADAPSO (now ITAA) (1989); and was contributing editor on *Computer Crime: A Crime-fighter's Handbook* (1995) and *Web Security and Commerce* (1997), both also published by O'Reilly. He has written over 100 papers and reports on his research, and contributed chapters to many other books on computer security and professional ethics, and has spoken internationally at conferences, symposia, and colloquia on these and related issues.

Among many other activities, Dr. Spafford is a member of the Association for Computing Machinery (ACM), where he has served as chair of the ACM Self-Assessment Committee, and as a member of the Technical Standards Committee. He is currently a member of the ACM's U.S. Public Policy Committee. He is also a Senior Member of the IEEE and Computer Society of the IEEE, and a charter recipient of the IEEE Computer Society's *Golden Core* award for service to the profession.

Publication History

This paper originally began as parts of two invited talks: A panel presentation on ethics and security at the Winter 1989 Usenix Conference following the Morris Internet Worm incident, and a presentation by the Harvard Office of Information Technology in November of 1989 on computer rights and responsibilities.

The first written version of these comments appeared in the Harvard INFORMATION TECHNOLOGY QUARTERLY as *Is a Computer Break-in Ever Ethical?*[17]. An expanded version of the paper was then solicited for a special issue of the JOURNAL OF SYSTEMS AND SOFTWARE, where it was published in 1992.[18]

Since then, the article has been widely reprinted in works on computing ethics:

- Reprinted (pp. 125–134) in COMPUERS, ETHICS, & SOCIAL VALUES; D. G. Johnson and

easy it is for anybody to listen to them. Now if I say that by tuning to 871 megahertz you can listen to a cellular phone call, I don't think I am committing a crime, I think I am explaining to somebody. What I have done at previous conferences is hold up this scanner and press a button and show people how easy it is to listen, and those people, when they get into their cars later on in the day, they do not use their cellular telephones to make private calls of a personal nature because they have learned something, and that is what we are trying to do, we are trying to show people how easy it is. Now, yes, that information can be used in a bad way, but to use that as an excuse not to give out the information at all is even worse, and I think it is much more likely that things may be fixed, the cellular industry may finally get its act together and start protecting phone calls. The phone companies might make red boxes harder to use or might make it easier for people to afford phone calls, but we will never know if we don't make it public.

Mr. Fields I want to be honest with you, Mr. Goldstein. I think it is frightening that someone like you thinks there is a protected right in invading someone else's privacy.

III

Encryption, Privacy, and Crypto-Anarchism

Peter Ludlow

This section begins with the question of whether individuals should be allowed to use military-grade encryption technology to encrypt their electronic communications. It should be obvious why good encryption technology is attractive to many people. Current encryption programs like PGP (Pretty Good Privacy; see the readings by Zimmermann) make it possible to communicate with friends and business partners without allowing inquisitive enemies to "eavesdrop" on our communications. On the other hand, it should also be obvious why certain government agencies are cool to the idea of widespread encryption technology. If encryption is so good that government security forces cannot break the code, then in principle criminals can communicate freely over the Internet without fear of having their plans compromised. The question is, which set of concerns should weigh more heavily, those of individuals, or those of government security forces?

Some have argued that the dilemma just posed is a false one—that there is a solution that allows individuals to have military-grade encryption while at the same time giving the government a "back door" which allows it to decipher the encrypted communications of potential terrorists, for example. In current versions of this proposal a chip (called the "Clipper Chip"), designed to encrypt and decipher digital communications, could be installed in all phones and computers and would provide standardized, military-grade encryption to users (see, for example, the reading by Denning). Yet the government would have a "key" that would be held in escrow by a government agency (or agencies) and that could only be used when a court order was issued. Thus, if the FBI had evidence that potential terrorists were exchanging encrypted plans, the government could apply for a court order, retrieve the "key," and begin monitoring the communications of the terrorists. So everyone should be happy, right? Wrong.

A number of commentators (see for example, the reading by Barlow) are highly suspicious of any plan that gives the government a "built-in" way of tapping into our communications. In the view of these commentators, our government has given us little reason to trust it with control over our secret encryption keys. When corporations, for example, can exchange business information worth millions of dollars, is it not plausible that someone would try to acquire these keys from poorly paid government bureaucrats through bribery? Or why suppose that our government might not exploit

its key ownership to gather information on political foes? This is not a wildly implausible scenario. Government officials are notorious for mistaking their political well-being with the well-being of the nation. In recent years the Nixon administration was famous for blurring this distinction, but the problem goes back much further.

For example, according to Dorothy Fowler (in her book *Unmailable*), in 1785 at the request of the secretary for the Department of Foreign Affairs, a resolution was passed authorizing that office to inspect any mail when it thought the safety and interest of the United States required such inspection. Congress was exempted from this ruling; it appears, however, that their mail was opened and read. For example, George Washington complained that his mail that went through the post office was opened and its content made known to everyone. Supposedly the problem was so bad that Madison, Jefferson, and Monroe took to using a cipher to communicate.

In 1792 an act was passed officially prohibiting this letter opening, but it didn't seem to help. Most of the postmasters were Federalists and had little problem with opening the mail. For example, in 1798 Jefferson wrote to John Taylor that he owed him a "political letter" but that "the infidelities of the post office and the circumstances of the times are against my writing fully and freely." Jefferson anonymously wrote another letter to a colleague saying "you will know from whom this comes without a signature; the omission of which has rendered almost habitual with me by the curiosity of the post offices." Who says only criminals have need of encryption and anonymous remailers?

Advocates of the Clipper chip maintain that the proposed safeguards should be adequate to prevent most abuses, and add that the potential consequences of unrestrained encryption could be devastating. Suppose terrorists used PGP encryption software to plot the construction and planting of a nuclear bomb. Of course, one might ask why terrorists smart enough to build such a device would be dumb enough to communicate their plans using an encryption devise for which the United States government has a key.

Clipper-type strategies also reflect a certain peculiar view about the nature of communications in the global marketplace. It is one thing to allow the United States government to be free to intercept all communications between its citizens, but what happens when those citizens work

for corporations based in other countries, or when U.S. corporations communicate with corporations in other countries? For example, suppose that Smith works for a Japanese auto manufacturer here in the United States. Is it appropriate that the U.S. government be able to spy on the communications between Smith and Smith's employer, particularly if the information being exchanged includes valuable trade secrets that might be of value to U.S. auto makers? Or what if the sales office of a German auto maker wants to communicate sales info to its home office in Stuttgart. Is it appropriate that the U.S. Government be able to eavesdrop on those communications? The questions need to be seriously considered, lest we lapse into a sort of myopic thinking about our own interests in a global marketplace filled with competing interests.

The discussion thus far has merely taken up some of the obvious benefits and problems attributed to the use of encryption technologies, but it is arguable that there are some more far-reaching consequences to consider. Technologically it is possible not only to encrypt simple messages, but to effectively digitize and encrypt our financial transactions as well (see the reading from Chaum). So, for example, it is possible to set up an electronic bank somewhere on the Internet (the exact location could be protected by an anonymous remailer), which could pay "info credits" to other accounts upon receiving an encrypted order from the payer's account. In effect, we could have a network of financial transactions taking place entirely in encrypted communications with a bank of unknown location.

It is interesting to speculate on the consequences of such a banking arrangement. One immediate consequence might be the emergence of underground black-market economies engaged in the swapping of proprietary information (see the Timothy May readings on one such hypothetical network, "blacknet"). Are such scenarios utopian or antiutopian? That issue is apparently subject to debate (May himself seems to take the utopian view).

But there are even more far-reaching possibilities than the mere emergence of black-market economies (which will always be with us to some degree in any case). Some of the cypherpunks have hypothesized that the emergence of encrypted banking may eventually lead to the death of the nation-state. According to this line of thinking, as more transactions take place in the underground banking networks, more money will escape

traditional attempts at taxation. As this happens the nation-states will lose more power or be forced to impose higher taxes, forcing even more corporations into the underground economy.

Are predictions about the death of the nation-state just speculative science fiction? Not necessarily. If my business is information intensive, there is no reason I cannot conduct my business from an underground computer account, trade with underground partners, and use underground banks (all via encrypted communications). At times, I will need to buy tangible goods, and these transactions will certainly be visible to the government; but why would the government need to know about the rest of my transactions? It is inevitable that there will be future information barons who amass billion-dollar fortunes, and who conduct their business using underground banks on the Internet. This does not make for a mere billion-dollar underground economy, however. The underground electronic bank will potentially invest in other ventures, thus expanding the monetary supply in the underground economy. At a certain crucial threshold, enough money could escape the taxation net of the nation-state so that its abilities to operate effectively will erode. If the nation-state chooses to raise taxes, more businesses will slip into the electronic underground, further eroding the viability of the national government.

Taxes, contrary to what some of the cypherpunks think, are still inevitable. New underground trading confederations would probably require new security arrangements (such as hacker defense), and those will, of course, have to be paid for. The future does not promise to be tax free. Nevertheless, taxation authority will be radically restructured without reference to traditional nation-state boundaries. The significance? The cypherpunks may not be too far off base when they prophesize the end of the nation-state.

So far this is just an observation, not a judgment, and we might well recoil in horror at such scenarios. In any case it is time to take such possibilities seriously and ask ourselves the following questions: Will encryption technologies hasten the demise of national governments as we know them? Is this a bad thing, or a good thing? If it is a bad thing, is there anything that can prevent it? If it is a good thing, what can be done to speed matters along?

11

How PGP Works/Why Do You Need PGP?

Philip R. Zimmermann

How It Works

It would help if you were already familiar with the concept of cryptography in general and public key cryptography in particular. Nonetheless, here are a few introductory remarks about public key cryptography.

First, some elementary terminology. Suppose I want to send you a message, but I don't want anyone but you to be able to read it. I can "encrypt," or "encipher" the message, which means I scramble it up in a hopelessly complicated way, rendering it unreadable to anyone except you, the intended recipient of the message. I supply a cryptographic "key" to encrypt the message, and you have to use the same key to decipher or "decrypt" it. At least that's how it works in conventional "single-key" cryptosystems.

In conventional cryptosystems, such as the U.S. Federal Data Encryption Standard (DES), a single key is used for both encryption and decryption. This means that a key must be initially transmitted via secure channels so that both parties can know it before encrypted messages can be sent over insecure channels. This may be inconvenient. If you have a secure channel for exchanging keys, then why do you need cryptography in the first place?

In public key cryptosystems, everyone has two related complementary keys, a publicly revealed key and a secret key. Each key unlocks the code that the other key makes. Knowing the public key does not help you deduce the corresponding secret key. The public key can be published and widely disseminated across a communications network. This protocol

provides privacy without the need for the same kind of secure channels that a conventional cryptosystem requires.

Anyone can use a recipient's public key to encrypt a message to that person, and that recipient uses her own corresponding secret key to decrypt that message. No one but the recipient can decrypt it, because no one else has access to that secret key. Not even the person who encrypted the message can decrypt it.

Message authentication is also provided. The sender's own secret key can be used to encrypt a message, thereby "signing" it. This creates a digital signature of a message, which the recipient (or anyone else) can check by using the sender's public key to decrypt it. This proves that the sender was the true originator of the message, and that the message has not been subsequently altered by anyone else, because the sender alone possesses the secret key that made that signature. Forgery of a signed message is infeasible, and the sender cannot later disavow his signature.

These two processes can be combined to provide both privacy and authentication by first signing a message with your own secret key, then encrypting the signed message with the recipient's public key. The recipient reverses these steps by first decrypting the message with her own secret key, then checking the enclosed signature with your public key. These steps are done automatically by the recipient's software.

Because the public key encryption algorithm is much slower than conventional single-key encryption, encryption is better accomplished by using a high-quality fast conventional single-key encryption algorithm to encipher the message. This original unenciphered message is called "plaintext." In a process invisible to the user, a temporary random key, created just for this one "session," is used to conventionally encipher the plaintext file. Then the recipient's public key is used to encipher this temporary random conventional key. This public-key-enciphered conventional "session" key is sent along with the enciphered text (called "ciphertext") to the recipient. The recipient uses her own secret key to recover this temporary session key, and then uses that key to run the fast conventional single-key algorithm to decipher the large ciphertext message.

Public keys are kept in individual "key certificates" that include the key owner's user ID (which is that person's name), a timestamp of when the key pair was generated, and the actual key material. Public key

certificates contain the public key material, while secret key certificates contain the secret key material. Each secret key is also encrypted with its own password, in case it gets stolen. A key file, or "key ring" contains one or more of these key certificates. Public key rings contain public key certificates, and secret key rings contain secret key certificates.

The keys are also internally referenced by a "key ID," which is an "abbreviation" of the public key (the least significant 64 bits of the large public key). When this key ID is displayed, only the lower 24 bits are shown for further brevity. While many keys may share the same user ID, for all practical purposes no two keys share the same key ID.

PGP uses "message digests" to form signatures. A message digest is a 128-bit cryptographically strong one-way hash function of the message. It is somewhat analogous to a "checksum" or CRC error checking code, in that it compactly "represents" the message and is used to detect changes in the message. Unlike a CRC, however, it is computationally infeasible for an attacker to devise a substitute message that would produce an identical message digest. The message digest gets encrypted by the secret key to form a signature.

Documents are signed by prefixing them with signature certificates, which contain the key ID of the key that was used to sign it, a secret-key-signed message digest of the document, and a timestamp of when the signature was made. The key ID is used by the receiver to look up the sender's public key to check the signature. The receiver's software automatically looks up the sender's public key and user ID in the receiver's public key ring.

Encrypted files are prefixed by the key ID of the public key used to encrypt them. The receiver uses this key ID message prefix to look up the secret key needed to decrypt the message. The receiver's software automatically looks up the necessary secret decryption key in the receiver's secret key ring.

These two types of key rings are the principal method of storing and managing public and secret keys. Rather than keep individual keys in separate key files, they are collected in key rings to facilitate the automatic lookup of keys either by key ID or by user ID. Each user keeps his own pair of key rings. An individual public key is temporarily kept in a separate file long enough to send to your friend who will then add it to her key ring.

Why Do You Need PGP?

It's personal. It's private. And it's no one's business but yours. You may be planning a political campaign, discussing your taxes, or having an illicit affair. Or you may be doing something that you feel shouldn't be illegal, but is. Whatever it is, you don't want your private electronic mail (e-mail) or confidential documents read by anyone else. There's nothing wrong with asserting your privacy. Privacy is as apple-pie as the Constitution.

Perhaps you think your e-mail is legitimate enough that encryption is unwarranted. If you really are a law-abiding citizen with nothing to hide, then why don't you always send your paper mail on postcards? Why not submit to drug testing on demand? Why require a warrant for police searches of your house? Are you trying to hide something? You must be a subversive or a drug dealer if you hide your mail inside envelopes. Or maybe a paranoid nut. Do law-abiding citizens have any need to encrypt their e-mail?

What if everyone believed that law-abiding citizens should use postcards for their mail? If some brave soul tried to assert his privacy by using an envelope for his mail, it would draw suspicion. Perhaps the authorities would open his mail to see what he's hiding. Fortunately, we don't live in that kind of world, because everyone protects most of their mail with envelopes. So no one draws suspicion by asserting their privacy with an envelope. There's safety in numbers. Analogously, it would be nice if everyone routinely used encryption for all their e-mail, innocent or not, so that no one drew suspicion by asserting their e-mail privacy with encryption. Think of it as a form of solidarity.

Today, if the Government wants to violate the privacy of ordinary citizens, it has to expend a certain amount of expense and labor to intercept and steam open and read paper mail, and listen to and possibly transcribe spoken telephone conversation. This kind of labor-intensive monitoring is not practical on a large scale. This is only done in important cases when it seems worthwhile.

More and more of our private communications are being routed through electronic channels. Electronic mail is gradually replacing conventional paper mail. E-mail messages are just too easy to intercept and

scan for interesting keywords. This can be done easily, routinely, automatically, and undetectably on a grand scale. International cablegrams are already scanned this way on a large scale by the NSA.

We are moving toward a future when the nation will be crisscrossed with high capacity fiber optic data networks linking together all our increasingly ubiquitous personal computers. E-mail will be the norm for everyone, not the novelty it is today. The Government will protect our e-mail with Government-designed encryption protocols. Probably most people will trust that. But perhaps some people will prefer their own protective measures.

Senate Bill 266, a 1991 omnibus anticrime bill, had an unsettling measure buried in it. If this non-binding resolution had become real law, it would have forced manufacturers of secure communications equipment to insert special "trap doors" in their products, so that the Government can read anyone's encrypted messages. It reads: "It is the sense of Congress that providers of electronic communications services and manufacturers of electronic communications service equipment shall insure that communications systems permit the Government to obtain the plain text contents of voice, data, and other communications when appropriately authorized by law." This measure was defeated after rigorous protest from civil libertarians and industry groups.

In 1992, the FBI Digital Telephony wiretap proposal was introduced to Congress. It would require all manufacturers of communications equipment to build in special remote wiretap ports that would enable the FBI to remotely wiretap all forms of electronic communication from FBI offices. Although it never attracted any sponsors in Congress because of citizen opposition, it will be reintroduced in 1993.

Most alarming of all is the White House's bold new encryption policy initiative, under development at NSA for four years, and unveiled April 16th, 1993. The centerpiece of this initiative is a Government-built encryption device, called the "Clipper" chip, containing a new classified NSA encryption algorithm. The Government is encouraging private industry to design it into all their secure communication products, like secure phones, secure FAX, etc. AT&T is now putting the Clipper into all their secure voice products. The catch: At the time of manufacture, each Clipper chip will be loaded with its own unique key, and the

Government gets to keep a copy, placed in escrow. Not to worry, though—the Government promises that they will use these keys to read your traffic only when duly authorized by law. Of course, to make Clipper completely effective, the next logical step would be to outlaw other forms of cryptography.

If privacy is outlawed, only outlaws will have privacy. Intelligence agencies have access to good cryptographic technology. So do the big arms and drug traffickers. So do defense contractors, oil companies, and other corporate giants. But ordinary people and grassroots political organizations mostly have not had access to affordable "military grade" public-key cryptographic technology. Until now.

PGP empowers people to take their privacy into their own hands. There's a growing social need for it. That's why I wrote it.

12

Crypto Rebels

Steven Levy

The office atmosphere of Cygnus Support, a fast-growing Silicon Valley company that earns its dollars by providing support to users of free software, seems like a time warp to the days when hackers ran free. Though Cygnus is located in a mall-like business park within earshot of U.S. 101, it features a spacious cathedral ceiling overhanging a cluttered warren of workstation cubicles arranged in an irregular spherical configuration. A mattress is nestled in the rafters. In a hallway behind the reception desk is a kitchen laden with snack food and soft drinks.

Today, a Saturday, only a few show up for work. The action instead is in a small conference room overlooking the back of the complex—a "physical meeting" of a group whose members most often gather in the corridors of cyberspace. Their mutual interest is the arcane field of cryptography—the study of secret codes and cyphers. The very fact that this group exists, however, is indication that the field is about to shift into overdrive. This is crypto with an attitude, best embodied by the group's moniker: Cypherpunks.

The one o'clock meeting doesn't really get underway until almost three. By that time around fifteen techie-cum-civil libertarians are sitting around a table, wandering around the room, or just lying on the floor staring at the ceiling while listening to the conversations. Most have beards and long hair—Smith Brothers gone digital.

The talk today ranges from reports on a recent cryptography conference to an explanation of how entropy degrades information systems. There is an ad hoc demonstration of a new product, an AT&T "secure" phone, supposedly the first conversation-scrambler that's as simple to use

13

Jackboots on the Infobahn

John Perry Barlow

On January 11, I managed to schmooze myself aboard Air Force 2. It was flying out of LA, where its principal passenger had just outlined his vision of the information superhighway to a suited mob of television, show-biz, and cable types who fervently hoped to own it one day—if they could ever figure out what the hell it was.

From the standpoint of the Electronic Frontier Foundation the speech had been wildly encouraging. The administration's program, as announced by Vice President Al Gore, incorporated many of the concepts of open competition, universal access, and deregulated common carriage that we'd been pushing for the previous year.

But he had said nothing about the future of privacy, except to cite among the bounties of the NII its ability to "help law enforcement agencies thwart criminals and terrorists who might use advanced telecommunications to commit crimes."

On the plane I asked Gore what this implied about administration policy on cryptography. He became as noncommittal as a cigar-store Indian. "We'll be making some announcements. . . . I can't tell you anything more." He hurried to the front of the plane, leaving me to troubled speculation.

Despite its fundamental role in assuring privacy, transaction security, and reliable identity within the NII, the Clinton administration has not demonstrated an enlightenment about cryptography up to par with the rest of its digital vision.

The Clipper Chip—which threatens to be either the goofiest waste of federal dollars since President Gerald Ford's great Swine Flu program or,

if actually deployed, a surveillance technology of profound malignancy—seemed at first an ugly legacy of the Reagan-Bush modus operandi. "This is going to be our Bay of Pigs," one Clinton White House official told me at the time Clipper was introduced, referring to the disastrous plan to invade Cuba that Kennedy inherited from Eisenhower.

(Clipper, in case you're just tuning in, is an encryption chip that the National Security Agency and FBI hope will someday be in every phone and computer in America. It scrambles your communications, making them unintelligible to all but their intended recipients. All, that is, but the government, which would hold the "key" to your chip. The key would be separated into two pieces, held in escrow, and joined with the appropriate "legal authority.")

Of course, trusting the government with your privacy is like having a Peeping Tom install your window blinds. And, since the folks I've met in this White House seem like extremely smart, conscious freedom-lovers—hell, a lot of them are Deadheads—I was sure that after they were fully moved in, they'd face down the National Security Agency and the FBI, let Clipper die a natural death, and lower the export embargo on reliable encryption products.

Furthermore, the National Institutes of Standards and Technology and the National Security Council have been studying both Clipper and export embargoes since April. Given that the volumes of expert testimony they had collected overwhelmingly opposed both, I expected the final report would give the administration all the support it needed to do the right thing.

I was wrong. Instead, there would be no report. Apparently, they couldn't draft one that supported, on the evidence, what they had decided to do instead.

The Other Shoe Drops

On Friday, February 4, the other jackboot dropped. A series of announcements from the administration made it clear that cryptography would become their very own "Bosnia of telecommunications" (as one staffer put it). It wasn't just that the old Serbs in the National Security Agency and the FBI were still making the calls. The alarming new

reality was that the invertebrates in the White House were only too happy to abide by them. Anything to avoid appearing soft on drugs or terrorism.

So, rather than ditching Clipper, they declared it a Federal Data Processing Standard, backing that up with an immediate government order for 50,000 Clipper devices. They appointed the National Institutes of Standards and Technology and the Department of Treasury as the "trusted" third parties that would hold the Clipper key pairs. (Treasury, by the way, is also home to such trustworthy agencies as the Secret Service and the Bureau of Alcohol, Tobacco, and Firearms.)

They reaffirmed the export embargo on robust encryption products, admitting for the first time that its purpose was to stifle competition to Clipper. And they outlined a very porous set of requirements under which the cops might get the keys to your chip. (They would not go into the procedure by which the National Security Agency could get them, though they assured us it was sufficient.)

They even signaled the impending return of the dread Digital Telephony, an FBI legislative initiative requiring fundamental reengineering of the information infrastructure; providing wiretapping ability to the FBI would then become the paramount design priority.

Invasion of the Body Snatchers

Actually, by the time the announcements thudded down, I wasn't surprised by them. I had spent several days the previous week in and around the White House.

I felt like I was in another remake of the *Invasion of the Body Snatchers*. My friends in the administration had been transformed. They'd been subsumed by the vast mindfield on the other side of the security clearance membrane, where dwell the monstrous bureaucratic organisms that feed on fear. They'd been infected by the institutionally paranoid National Security Agency's *Weltanschauung*.

They used all the telltale phrases. Mike Nelson, the White House point man on the NII, told me, "If only I could tell you what I know, you'd feel the same way I do." I told him I'd been inoculated against that argument during Vietnam. (And it does seem to me that if you're going to initiate

a process that might end freedom in America, you probably need an argument that isn't classified.)

Besides, how does he know what he knows? Where does he get his information? Why, the National Security Agency, of course. Which, given its strong interest in the outcome, seems hardly an unimpeachable source.

However they reached it, Clinton and Gore have an astonishingly simple bottom line, to which even the future of American liberty and prosperity is secondary: They believe that it is their responsibility to eliminate, by whatever means, the possibility that some terrorist might get a nuke and use it on, say, the World Trade Center. They have been convinced that such plots are more likely to ripen to hideous fruition behind a shield of encryption.

The staffers I talked to were unmoved by the argument that anyone smart enough to steal a nuclear device is probably smart enough to use PGP or some other uncompromised crypto standard. And never mind that the last people who popped a hooter in the World Trade Center were able to get it there without using any cryptography and while under FBI surveillance.

We are dealing with religion here. Though only ten American lives have been lost to terrorism in the last two years, the primacy of this threat has become as much an article of faith with these guys as the Catholic conviction that human life begins at conception or the Mormon belief that the Lost Tribe of Israel crossed the Atlantic in submarines.

In the spirit of openness and compromise, they invited the Electronic Frontier Foundation to submit other solutions to the "problem" of the nuclear-enabled terrorist than key escrow devices, but they would not admit into discussion the argument that such a threat might, in fact, be some kind of phantasm created by the spooks to ensure their lavish budgets into the post-Cold War era.

As to the possibility that good old-fashioned investigative techniques might be more valuable in preventing their show-case catastrophe (as it was after the fact in finding the alleged perpetrators of the last attack on the World Trade Center), they just hunkered down and said that when wiretaps were necessary, they were damned well necessary.

When I asked about the business that American companies lose because of their inability to export good encryption products, one staffer essentially dismissed the market, saying that total world trade in crypto goods was still less than a billion dollars. (Well, right. Thanks more to the diligent efforts of the National Security Agency than to dim sales potential.)

I suggested that a more immediate and costly real-world effect of their policies would be to reduce national security by isolating American commerce, owing to a lack of international confidence in the security of our data lines. I said that Bruce Sterling's fictional data-enclaves in places like the Turks and Caicos Islands were starting to look real-world inevitable.

They had a couple of answers to this, one unsatisfying and the other scary. The unsatisfying answer was that the international banking community could just go on using DES, which still seemed robust enough to them. (DES is the old federal Data Encryption Standard, thought by most cryptologists to be nearing the end of its credibility.)

More frightening was their willingness to counter the data-enclave future with one in which no data channels anywhere would be secure from examination by one government or another. Pointing to unnamed other countries that were developing their own mandatory standards and restrictions regarding cryptography, they said words to the effect of, "Hey, it's not like you can't outlaw the stuff. Look at France."

Of course, they have also said repeatedly—and for now I believe them—that they have absolutely no plans to outlaw non-Clipper crypto in the U.S. But that doesn't mean that such plans wouldn't develop in the presence of some pending "emergency." Then there is that White House briefing document, issued at the time Clipper was first announced, which asserts that no U.S. citizen "as a matter of right, is entitled to an unbreakable commercial encryption product."

Now why, if it's an ability they have no intention of contesting, do they feel compelled to declare that it's not a right? Could it be that they are preparing us for the laws they'll pass after some bearded fanatic has gotten himself a surplus nuke and used something besides Clipper to conceal his plans for it?

If they are thinking about such an eventuality, we should be doing so as well. How will we respond? I believe there is a strong, though currently untested, argument that outlawing unregulated crypto would violate the First Amendment, which surely protects the manner of our speech as clearly as it protects the content.

But of course the First Amendment is, like the rest of the Constitution, only as good as the government's willingness to uphold it. And they are, as I say, in the mood to protect our safety over our liberty.

This is not a mind-frame against which any argument is going to be very effective. And it appeared that they had already heard and rejected every argument I could possibly offer.

In fact, when I drew what I thought was an original comparison between their stand against naturally proliferating crypto and the folly of King Canute (who placed his throne on the beach and commanded the tide to leave him dry), my government opposition looked pained and said he had heard that one almost as often as jokes about roadkill on the information superhighway.

I hate to go to war with them. War is always nastier among friends. Furthermore, unless they've decided to let the National Security Agency design the rest of the National Information Infrastructure as well, we need to go on working closely with them on the whole range of issues like access, competition, workplace privacy, common carriage, intellectual property, and such. Besides, the proliferation of strong crypto will probably happen eventually no matter what they do.

But then again, it might not. In which case we could shortly find ourselves under a government that would have the automated ability to log the time, origin and recipient of every call we made, could track our physical whereabouts continuously, could keep better account of our financial transactions than we do, and all without a warrant. Talk about crime prevention!

Worse, under some vaguely defined and surely mutable "legal authority," they also would be able to listen to our calls and read our e-mail without having to do any backyard rewiring. They wouldn't need any permission at all to monitor overseas calls.

If there's going to be a fight, I'd rather it be with this government than the one we'd likely face on that hard day.

Hey, I've never been a paranoid before. It's always seemed to me that most governments are too incompetent to keep a good plot strung together all the way from coffee break to quitting time. But I am now very nervous about the government of the United States of America.

Because Bill 'n' Al, whatever their other new-paradigm virtues, have allowed the very old-paradigm trogs of the Guardian Class to define as their highest duty the defense of America against an enemy that exists primarily in the imagination—and is therefore capable of anything.

To assure absolute safety against such an enemy, there is no limit to the liberties we will eventually be asked to sacrifice. And, with a Clipper Chip in every phone, there will certainly be no technical limit on their ability to enforce those sacrifices.

14

The Clipper Chip Will Block Crime

Dorothy E. Denning

Hidden among the discussions of the information highway is a fierce debate, with huge implications for everyone. It centers on a tiny computer chip called the Clipper, which uses sophisticated coding to scramble electronic communications transmitted through the phone system.

The Clinton administration has adopted the chip, which would allow law enforcement agencies with court warrants to read the Clipper codes and eavesdrop on terrorists and criminals. But opponents say that, if this happens, the privacy of law-abiding individuals will be at risk. They want people to be able to use their own scramblers, which the government would not be able to decode.

If the opponents get their way, however, all communications on the information highway would be immune from lawful interception. In a world threatened by international organized crime, terrorism, and rogue governments, this would be folly. In testimony before Congress, Donald Delaney, senior investigator with the New York State Police, warned that if we adopted an encoding standard that did not permit lawful intercepts, we would have havoc in the United States.

Moreover, the Clipper coding offers safeguards against casual government intrusion. It requires that one of the two components of a key embedded in the chip be kept with the Treasury Department and the other component with the Commerce Department's National Institute of Standards and Technology. Any law enforcement official wanting to wiretap would need to obtain not only a warrant but the separate components from the two agencies. This, plus the superstrong code and key system would make it virtually impossible for anyone, even corrupt government officials, to spy illegally.

But would terrorists use Clipper? The Justice Department has ordered $8 million worth of Clipper scramblers in the hope that they will become so widespread and convenient that everyone will use them. Opponents say that terrorists will not be so foolish as to use encryption to which the government holds the key but will scramble their calls with their own code systems. But then who would have thought that the World Trade Center bombers would have been stupid enough to return a truck that they had rented?

Court-authorized interception of communications has been essential for preventing and solving many serious and often violent crimes, including terrorism, organized crime, drugs, kidnaping, and political corruption. The FBI alone has had many spectacular successes that depended on wiretaps. In a Chicago case code-named RUKBOM, they prevented the El Rukn street gang, which was acting on behalf of the Libyan government, from shooting down a commercial airliner using a stolen military weapons system.

To protect against abuse of electronic surveillance, federal statutes impose stringent requirements on the approval and execution of wiretaps. Wiretaps are used judiciously (only 846 installed wiretaps in 1992) and are targeted at major criminals.

Now, the thought of the FBI wiretapping my communications appeals to me about as much as its searching my home and seizing my papers. But the Constitution does not give us absolute privacy from court-ordered searches and seizures, and for good reason. Lawlessness would prevail.

Encoding technologies, which offer privacy, are on a collision course with a major crime-fighting tool: wiretapping. Now the Clipper chip shows that strong encoding can be made available in a way that protects private communications but does not harm society if it gets into the wrong hands. Clipper is a good idea, and it needs support from people who recognize the need for both privacy and effective law enforcement on the information highway.

15

The Denning–Barlow Clipper Chip Debate

Dorothy E. Denning and John Perry Barlow

OnlineHost Good evening and welcome to the Time Online Odeon! Tonight we look from both sides at the Clipper Chip, a semiconductor device that the National Security Agency developed and wants installed in every telephone, computer modem and fax machine.

OnlineHost In his article in the current issue of *Time,* Philip Elmer-DeWitt writes: "The chip combines a powerful encryption algorithm with a "back door"—the cryptographic equivalent of the master key that opens schoolchildren's padlocks when they forget their combinations. A "secure" phone equipped with the chip could, with proper authorization, be cracked by the government.

OnlineHost "Law-enforcement agencies say they need this capability to keep tabs on drug runners, terrorists and spies. Critics denounce the Clipper—and a bill before Congress that would require phone companies to make it easy to tap the new digital phones—as Big Brotherly tools that will strip citizens of whatever privacy they still have in the computer age.

OnlineHost "Lined up on one side are the three-letter cloak-and-dagger agencies—the NSA, the CIA and the FBI—and key policymakers in the Clinton Administration (who are taking a surprisingly hard line on the encryption issue). Opposing them is an equally unlikely coalition of computer firms, civil libertarians, conservative columnists and a strange breed of cryptoanarchists who call themselves the cypherpunks."

RPTime Lined up on our stage tonight are John Perry Barlow, Dr. Dorothy Denning and Philip Elmer-DeWitt. Barlow is co-founder of the Electronic Frontier Foundation, which promotes freedom in digital media. A recognized commentator on computer security, he is arguing

against the Clipper Chip. Dr. Denning is the chairperson of the Computer Science Department at Georgetown University. A leading expert on cryptography and data security, she favors the adoption of the Clipper Chip. Philip Elmer-DeWitt, *Time*'s technology editor will lead the questioning of our guests. Audience questions may be sent up using the Interact with Host function . . . Phil?

PhilipED Dr. Denning, could you *briefly* make the case for why we need the key escrow encryption system.

DDenning The government needs a new encryption standard to replace DES. They came up with a very strong algorithm called SKIPJACK. In making that available, they didn't want to do it in a way that could ultimately prove harmful to society. So they came up with the idea of key escrow so that if SKIPJACK were used to conceal criminal activity, they would be able to get access to the communications.

PhilipED Thanks. Mr. Barlow, could you briefly make the case *against* Clipper.

Barlow1 We'll see if I can be brief. We oppose Clipper in large part because of the traffic analysis which it makes possible. We believe that it is in the functional nature of the chip as designed to greatly enhance the ability of government to observe who we are calling, when, and from where, all fairly automatically and centrally. We also oppose Clipper because of the many ways in which we believe the escrow system could be compromised, by people and institutions both inside and outside of government.

PhilipED Dr. Denning, what about John's contention that Clipper makes it easier to detect calling patterns.

DDenning I don't buy this. First off, for law enforcement to access any communications, they need a court order. Even if the communications are encrypted. Second, with a court order, they can get access to call setup information and find out what other lines the subject of the investigation is talking to. This is of much more use than anything in the encrypted stream.

PhilipED John, is Dorothy right that you need a court order for call setup info?

Barlow1 Dorothy, the government asked for and received over 100,000 calling records last year without a court order. I see nothing in the Clipper documents which indicates that they would require a court order to get this kind of information, which each chip would make readily available to the entire network.

DDenning You need a court order to do implement pen registers and dialed number recorders in order to find out who is talking to whom.

Barlow1 Furthermore, my faith in court orders has been eroded by 30 years of government wiretap abuse.

PhilipED Aren't we talking about three different hurdles here, one for a wiretap . . .

Barlow1 But that's only with the present system where putting a pen register on a line requires physical entrance to a phone company site.

PhilipED One for a pen register (to track calling patterns in real time) and one for phone records.

RPTime Let's take a question from the audience . . . How would you guarantee that this facility will never be misused? If you can't make that guarantee, why should a democratic society, with a prohibition against prior restraint, consent to this? John Barlow?

Barlow1 There are three different sources of information, as you say. But there are not three "hurdles." That sounds like a question for Dorothy. I don't think we should, obviously.

RPTime Dr. Denning?

DDenning First of all, there has been no evidence of widespread abuse of wiretaps since passage of the 1968 and 1978 wiretap statutes. Second, there are a lot of security mechanisms going into it to protect against abuse. Third, it will provide much greater protection against illegal wiretaps than we have now, since almost all phone conversations are in the clear. It will make virtually all illegal wiretaps impossible. Fourth, if for some reason it doesn't provide adequate protection, we can destroy the key databases and everyone will have absolute privacy against government wiretaps. I don't think our society will tolerate that kind of abuse.

PhilipED John, isn't Dorothy right that you're better off with compromised encryption than none?

Barlow1 Gee, where to begin . . . First of all, there was plenty of abuse after 1968. Remember Watergate, Dorothy? Second, I believe that Clipper in the Net will dramatically *enhance* certain powers of . . .

DDenning I was talking specifically about wiretap abuses. And there hasn't been any evidence since the 1978 law.

Barlow1 . . . surveillance over current technical abilities. One of the reasons that wiretap hasn't been more abused is the bureaucratic overhead of current practices. Make it so that it doesn't require 50 agents to conduct a wire tap and you'll see a lot more of it. And Watergate included quite a number of wiretap violations. Indeed, the burglers were caught trying to install one. As to the assertion that we can always back up and destroy the databases if we don't like it, I can't imagine that someone as bright as yourself would believe that this is possible. Technology and power ratchet into positions which almost never retract without a complete change in the system of authority

RPTime Care to respond Dr. Denning?

DDenning Clipper would prevent the Watergate burglars from getting anywhere since they wouldn't have a court order. Clipper will not make wiretaps cheaper or easier. Wiretaps are becoming more difficult. And there will always be more agents involved because they have to follow exacting procedures, including minimization (throw out all conversations that are not specific to the crime at hand).

Barlow1 Dorothy, they were from the *Government,* remember? I can't imagine that Nixon wouldn't have been able to find a sympathetic ear from somebody at NIST and somebody else at Treasury. Further, you're not talking about the truly insidious element of this, which is dramatically improved traffic analysis. Content is less important than context, and most agents will support this.

RPTime Another question from the audience. JCMaille asks . . . Does the government have a constitutional right of access to my personal communications? Dr. Denning, why don't you go first?

DDenning The Supreme Court ruled that wiretaps with a court order are Constitutional. At one time, communications were not even protected under the 4th Amendment. The government could wiretap without a court order! Now a court order is required.

PhilipED To put the question another way, do citizens have a right to use powerful encryption?

DDenning Right now there are no laws preventing the use of any encryption. Clipper is voluntary. You can still use something else.

RPTime We have to apologize. John Barlow has temporarily lost his connection . . .

PhilipED Dr. Denning, in your opinion . . . would a law outlawing powerful encryption be unconstitutional?

DDenning I don't think so. But that doesn't mean it will happen.

RPTime John Barlow is back with us. Sorry for the interruption! Barlow, Denning just said she didn't think a law banning powerful encryption would be unconstitutional. What do you say?

Barlow1 Actually, I believe that our current export embargoes are a violation of the 1st Amendment which specifies speech without regard to the manner of speech. If we could restrict manner of speech, it would be constitutional to require that everyone speak English. Which of course it isn't.

PhilipED John, can you make the case why ordinary law-abiding citizens need powerful encryption?

Barlow1 Because it is in the nature of digitally networked communications to be quite visible. Every time we make any sort of transaction in a digital environment, we smear our fingerprints all over Cyberspace. If we are to have any privacy in the future, we will need virtual "walls" made of cryptography.

RPTime Another audience question . . . Isn't this like the gun argument? If guns are outlawed only criminals will have guns? Well, if Clipper is standardized, won't criminals be the ones *not* using it?

RPTime Dr. Denning?

DDenning If Clipper becomes the de facto standard, then it will be the chief method of encryption. That would be what you'd get at Radio Shack. What criminals use will depend on what is readily available and what their cohorts are using. Both parties of a conversation have to use the same thing. Criminals also talk to a lot of people outside their immediate circle—e.g., to buy goods and services. Also, they can be quite stupid at times. But the main thing is that criminals will not be able to take advantage of the SKIPJACK algorithm as a way of concealing their conversations. This is the whole point. It is not to catch criminals. It is to allow people access to a really high quality algorithm in a way that someone cannot use it to conceal criminal activity.

Barlow1 The gun analogy is excellent up to a point. I can't for the life

of me imagine why we would think that even a stupid criminal would use Clipper if something else were available. And when I talk to people in the administration their big hobgoblin is the "nuclear-armed" terrorist. Any fanatic smart enough to assemble and detonate a nuclear device is going to be smart enough to download PGP from a bulletin board somewhere. Also, I'd like to point out that the gun analogy doesn't go the whole distance. Crypto is by its nature a purely *defensive* technology. You can't shoot people with it.

PhilipED Speaking of PGP, Dr. Denning, is that encryption system secure, in your opinion?

DDenning I don't know of anyone who's been able to break the IDEA algorithm that it uses.

RPTime Back to the audience for a question from SteveHW . . . This is for Dr. Denning. What is the evidence of harm if the Clipper proposal is not adopted?

DDenning The harm would be to the government. They would not be able to use it and would have to resort to something less secure. Also, Clipper is part of a larger project to make hardware available for encryption and digital signatures. This will be used, for example, in the Defense Message System. The goverment needs a new standard. I personally believe that making really powerful encryption like SKIPJACK available without key escrow could be harmful to society. Wiretaps have been essential for preventing and solving many serious crimes and terrorist activities.

Barlow1 Why on earth would the government have to use something else if they failed to get the rest of rest of us to buy into this folly? Hey, they are already using SKIPJACK. It's a government algorithm and has been in use for a . . .

DDenning CPSR and others are asking the government to drop Clipper.

Barlow1 . . . long time. There are plenty other algorithms which we can use which are truly protected . . . unless of course, this is only the first step in a process which will outlaw other forms of crypto. And I believe that it must be. Makes absolutely no sense otherwise. EFF is not asking the Government to drop Clipper, though we would vastly prefer they did. We're merely asking that no steps be taken to require it either by law or

practice . . . as, for example, would be the case if you had to use a Clipper chip to file your tax return.

PhilipED Dr. Denning, do you think this is the "first step in a process to outlaw crypto"?

DDenning No I do not. The government has not been using SKIPJACK to my knowledge. The Clipper initiative represents the first time that the government has put one of their really good algorithms out there in the unclassified arena. They are trying to do this in a way that won't backfire against the public. Other NSA developed algorithms are not available for purchase by the public.

Barlow1 I appreciate their willingness to make some of that crypto research available to a public which has paid so much for it, but I'm afraid that I would never trust an algorithm which was given to me by any government. And I certainly don't trust a classified algorithm like SKIPJACK, even without a back door which everyone can see. I think I'll stick to systems which have been properly vetted to be clear of such compromises, like RSA. I hope others will do likewise and that RSA will become the standard which Clipper shouldn't be.

RPTime Time for one more question from our audience . . . To John Barlow. Isn't society becoming increasingly vulnerable to concerted criminal/terrorist disruption, requiring *stronger* law enforcement tools?

Barlow1 Gee. I don't know. It's a scary world. However, I'm willing to take my chances with the few terrorists and drug lords there are out there rather than trusting government with the kind of almost unlimited surveillance power which Clipper and Digital Telephony would give them. It's a tough choice. But when you look at the evil perpetrated by government over this century in the name of stopping crime, it far exceeds that done by other organized criminals.

RPTime Dr. Denning, hasn't remote listening technology enhanced police abilities to eavesdrop to the point . . . where the loss of a few wire taps won't mean much?

DDenning No. They need to get the cooperation of the service providers to implement a wiretap. The loss of some wiretaps could be costly indeed. As an example, wiretaps were used to help solve a case that involved plans by a Chicago gang from shooting down a commercial airliner.

There have been 2 cases where they helped save the lives of kids who were going to be kidnaped for the making of a snuff murder film. They helped solve a case where a man's house was going to be bombed. I could go on. If we take John's arguments about law enforcement to their logical conclusion, we'd just get rid of law enforcement. I think it's better to have it. The people in law enforcement hate it as much as the rest of us when some member of the community does something wrong. And they correct it, design new procedures and laws where necessary, and go on.

Barlow1 Oh, please. I'm not proposing eliminating police. I'm opposing giving them unlimited powers. Also, these are the same cases cited over and over by everyone from you to Judge Freeh. Surely, we aren't going to fundamentally change the balance of power in this country because of these two (undocumented, to my knowledge) stories.

DDenning Clipper is not going to change the balance of power. It does not give law enforcement any additional authority to do wiretaps.

Barlow1 Well, this is where we basically disagree, Dorothy. If we could continue the same level of LE capacity we presently have, I'd have no objection. But I believe, for reasons I'm not sure we have the bandwidth to discuss here, that we are talking about dramatically enhancing their abilities. For one thing, we would greatly reduce the bureaucratic overhead involved in wiretap, which is what keeps it under 900 cases nationwide at the present.

RPTime And that will have to be the last word on the matter for tonight . . .

DDenning The overhead of a wiretap is more likely to increase, not decrease.

PhilipED Not quite! Maybe not! ;-)

RPTime *That* will be the final word!

Barlow1 Well, let's get together and talk, Dorothy.

RPTime *Time* thanks Dr. Dorothy Denning and John Perry Barlow for being with us tonight . . . along with Philip Elmer-DeWitt. Thank you all, and goodnight! Thank you both. This was very interesting.

DDenning Thank you for the opportunity to be here!

16

Achieving Electronic Privacy

David Chaum

Every time you make a telephone call, purchase goods using a credit card, subscribe to a magazine or pay your taxes, that information goes into a data base somewhere. Furthermore, all these records can be linked so that they constitute in effect a single dossier on your life, not only your medical and financial history but also what you buy, where you travel and whom you communicate with. It is almost impossible to learn the full extent of the files that various organizations keep on you, much less to assure their accuracy or to control who may gain access to them.

Organizations link records from different sources for their own protection. Certainly it is in the interest of a bank looking at a loan application to know that John Doe has defaulted on four similar loans in the past two years. The bank's possession of that information also helps its other customers, to whom the bank passes on the cost of bad loans. In addition, these records permit Jane Roe, whose payment history is impeccable, to establish a charge account at a shop that has never seen her before.

That same information in the wrong hands, however, provides neither protection for businesses nor better service for consumers. Thieves routinely use a stolen credit card number to trade on their victims' good payment records; murderers have tracked down their targets by consulting government-maintained address records. On another level, the U.S. Internal Revenue Service has attempted to single out taxpayers for audits based on estimates of household income compiled by mailing-list companies.

The growing amounts of information that different organizations collect about a person can be linked because all of them use the same key in the U.S.—the social security number—to identify the individual in ques-

Hiding Crimes in Cyberspace[1]

Dorothy E. Denning and William E. Baugh, Jr.

July 1999

[To appear in *Information, Communication and Society*, Vol. 2, No 3, Autumn 1999, and in *Cybercrime*, B. D. Loader and D. Thomas (eds.), Routledge, 1999.

INTRODUCTION

The growth of telecommunications and electronic commerce has led to a growing commercial market for digital encryption technologies. Business needs encryption to protect intellectual property and to establish secure links with their partners, suppliers, and customers. Banks need it to ensure the confidentiality and authenticity of financial transactions. Law enforcement needs it to stop those under investigation from intercepting police communications and obstructing investigations. Individuals need it to protect their private communications and confidential data. Encryption is critical to building a secure and trusted global information infrastructure for communications and electronic commerce.

Encryption also gives criminals and terrorists a powerful tool for concealing their activities. It can make it impossible for law enforcement agencies to obtain the evidence needed for a conviction or the intelligence vital to criminal investigations. It can frustrate communications intercepts, which have played a significant role in averting terrorist attacks and in gathering information about specific transnational threats, including terrorism, drug trafficking, and organized crime (White House 1995). It can delay investigations and add to their cost.

The use of encryption to hide criminal activity is not new. The April 1970 issue of the FBI Law Enforcement Bulletin reports on several cases where law enforcement agencies had to break codes in order to obtain evidence or prevent violations of the law. None of the cases, however, involved electronic information or computers. Relatively simple substitution ciphers were used to conceal speech.

Digital computers have changed the landscape considerably. Encryption and other advanced technologies increasingly are used, with direct impact on law enforcement. If all communications and stored information in criminal cases were encrypted, it would be a nightmare for investigators. It would not be feasible to decrypt everything, even if technically possible. How would law enforcement agencies know where to spend limited resources?

We address here the use of encryption and other information technologies to hide criminal activities. Numerous case studies are presented for illustration. We first examine encryption and the options available to law enforcement for dealing with it. Next we discuss a variety of other tools for concealing information: passwords, digital compression, steganography, remote storage, and audit disabling. Finally we discuss tools for hiding crimes through anonymity: anonymous remailers, anonymous digital cash, computer penetration and looping, cellular phone cloning, and cellular phone cards.

ENCRYPTION IN CRIME AND TERRORISM

This section describes criminal use of encryption in four domains: voice, fax, and data communications; electronic mail; files stored on the computers of individual criminals and criminal enterprises; and information posted in public places on computer networks.

Voice, Fax, and Real-Time Data Communications

Criminals can use encryption to make their real-time communications inaccessible to law enforcement. The effect is to deny law enforcement one of the most valuable tools in fighting organized crime - the court-ordered wiretap. In March 1997, the director of the Federal Bureau of Investigation, Louis J. Freeh, testified that the FBI was unable to assist with 5 requests for decryption assistance in communications intercepts in 1995 and 12 in 1996 (US Congress 1997a). Such wiretaps can be extremely valuable as they capture the subjects' own words, which generally holds up much better in court than information acquired from informants, for example, who are often criminals themselves and extremely unreliable. Wiretaps also provide valuable information regarding the intentions, plans, and members of criminal conspiracies, and in providing leads in criminal investigations. Drug cartels and organizations rely heavily on communications networks; monitoring of these networks has been critical for identifying those at the executive level and the organizations' illegal proceeds. Communications intercepts have also been useful in terrorism cases, sometimes helping to avoid a deadly attack. They have helped prevent the bombing of a foreign consulate in the United States and a rocket attempt against a U.S. ally, among other things (ibid).

There is little case information in the public domain on the use of communications encryption devices by criminal enterprises. The Cali cartel is reputed to be using sophisticated encryption to conceal their telephone communications. Communications devices seized from the cartel in 1995 included radios that distort voices, video phones which provide visual authentication of the caller's identity, and instruments for scrambling transmissions from computer modems (Grabosky and Smith 1997).

We understand that some terrorist groups are using high-frequency encrypted voice/data links with state sponsors of terrorism. Hamas reportedly is using encrypted

Internet communications to transmit maps, pictures, and other details pertaining to terrorist attacks. The Israeli General Security Service believes that most of the data is being sent to the Hamas worldwide center in Great Britain (IINS 1997).

The lack of universal interoperability and cost of telephone encryption devices - several hundred dollars for a device that provides strong security - has likely slowed their adoption by criminal enterprises. The problems to law enforcement could get worse as prices drop and Internet telephony becomes more common. Criminals can conduct encrypted voice conversations over the Internet at little or no cost. This impact on law enforcement, however, may be balanced by the emergence of digital cellular communications. These phones encrypt the radio links between the mobile devices and base stations, which is where the communications are most vulnerable to eavesdroppers. Elsewhere, the communications travel in the clear (or are separately encrypted while traversing microwave or satellite links), making court-ordered interception possible in the switches. The advantage to users is that they can protect their local over-the-air communications even if the parties they are conversing with are using phones with no encryption or with incompatible methods of encryption. The benefit to law enforcement is that plaintext can be intercepted in the base stations or switches. Although there are devices for achieving end-to-end encryption with cellular phones, they are more costly and require compatible devices at both ends.

Hackers use encryption to protect their communications on Internet Relay Chat (IRC) channels from interception. They have also installed their own encryption software on computers they have penetrated. The software is then used to set up a secure channel between the hacker's PC and the compromised machine. This has complicated, but not precluded, investigations.

Electronic Mail

Law enforcement agencies have encountered encrypted e-mail and files in investigations of pedophiles and child pornography, including the FBI's Innocent Images national child pornography investigation. In many cases, the subjects were using Pretty Good Privacy (PGP) to encrypt files and e-mail. PGP uses conventional cryptography for data encryption and public-key cryptography for key distribution. The investigators thought this group favored PGP because they are generally educated, technically knowledgeable, and heavy Internet users. PGP is universally available on the Internet, and they can download it for free. Investigators say, however, that most child pornography traded on the Internet is not encrypted.

One hacker used encrypted e-mail to facilitate the sale of credit card numbers he had stolen from an Internet service provider and two other companies doing business on the Web. According to Richard Power, editorial director of the Computer Security Institute, Carlos Felipe Salgado Jr. had acquired nearly 100,000 card numbers by penetrating the computers from an account he had compromised at the University of California at San Francisco. Using commonly available hacking tools, he exploited known security flaws in order to go around firewalls and bypass encryption and other

security measures. Boasting about his exploits on Internet Relay Chat, Salgado, who used the code name SMAK, made the mistake of offering to sell his booty to someone on the Internet. He conducted on-line negotiations using encrypted e-mail and received initial payments via anonymous Western Union wire transfer. Unknown to him, he had walked right into an FBI sting. After making two small buys and checking the legitimacy of the card numbers, FBI agents arranged a meeting at San Francisco airport. Salgado was to turn over the credit cards in exchange for $260,000. He arrived with an encrypted CD-ROM containing about 100,000 credit card numbers and a paperback copy of Mario Puzo's The Last Don. The key to decrypting the data was given by the first letter of each sentence in the first paragraph on page 128. Salgado was arrested and waived his rights. In June 1997, he was indicted on three counts of computer crime fraud and two counts of trafficking in stolen credit cards. In August, he pled guilty to four of the five counts. Had he not been caught, the losses to the credit card companies could have run from $10 million to over $100 million (Power 1997).

We were told of another case in which a terrorist group that was attacking businesses and state officials used encryption to conceal their messages. At the time the authorities intercepted the communications, they were unable to decrypt the messages, although they did perform some traffic analysis to determine who was talking with whom. Later they found the key on the hard disk of a seized computer, but only after breaking through additional layers of encryption, compression, and password protection. The messages were said to have been a great help to the investigating task force. We also received an anonymous report of a group of terrorists encrypting their e-mail with PGP.

Stored Data

In many criminal cases, documents and other papers found at a subject's premises provide evidence crucial for successful prosecution. Increasingly, this information is stored electronically on computers. Computers themselves have posed major challenges to law enforcement, and encryption has only compounded these challenges.

The FBI found encrypted files on the laptop computer of Ramsey Yousef, a member of the international terrorist group responsible for bombing the World Trade Center in 1993 and a Manila Air airliner in late 1995. These files, which were successfully decrypted, contained information pertaining to further plans to blow up eleven U.S.-owned commercial airliners in the Far East (US Congress 1997a). Although much of the information was also available in unencrypted documents, the case illustrates the potential threat of encryption to public safety if authorities cannot get information about a planned attack and some of the conspirators are still at large.

Successful decryption of electronic records can be important to an investigation. Such was the case when Japanese authorities seized the computers of the Aum Shinrikyo cult - the group responsible for gassing the Tokyo subway in March 1995,

killing 12 people and injuring 6,000 more (Kaplan and Marshall 1996). The cult had stored their records on computers, encrypted with RSA. Authorities were able to decrypt the files after finding the key on a floppy disk. The encrypted files contained evidence that was said to be crucial to the investigation, including plans and intentions to deploy weapons of mass destruction in Japan and the United States.

In the Aum cult case, the authorities were lucky to find the key on a disk. In other cases, the subjects turned over their keys. For example, the Dallas Police Department encountered encrypted data in the investigation of a national drug ring which was operating in several states and dealing in Ecstasy. A member of the ring, residing within their jurisdiction, had encrypted his address book. He turned over the password, enabling the police to decrypt the file. Meanwhile, however, the subject was out on bond and alerted his associates, so the decrypted information was not as useful as it might have been. The detective handling the case said that in the ten years he had been working drug cases, this was the only time he had encountered encryption, and that he rarely even encountered computers. He noted that the Ecstasy dealers were into computers more than other types of drug dealers, most likely because they are younger and better educated. They were using the Internet for sales, but they were not encrypting electronic mail. The detective also noted that the big drug dealers were not encrypting phone calls. Instead, they were swapping phones (using cloned phones - see later discussion) to stay ahead of law enforcement (Manning 1997).[2]

In many cases, investigators have had to break the encryption system in order to get at the data. For example, when the FBI seized the computers of CIA spy Aldrich Ames, they found encrypted computer files, but no keys. Fortunately, Ames had used standard commercial off-the-shelf software, and the investigator handling the computer evidence was able to break the codes using software supplied by AccessData Corporation of Orem, Utah. The key was Ames's Russian code name, KOLOKOL (bell). According to investigators, failure to recover the encrypted data would have weakened the case. Ames was eventually convicted of espionage against the United States (CSI 1997).[3]

Code breaking is not always so easy. In his book about convicted hacker Kevin Poulsen, Jonathan Littman reported that Poulsen had encrypted files documenting everything from the wiretaps he had discovered to the dossiers he had compiled about his enemies. The files were said to have been encrypted several times using the 'Defense Encryption Standard' [sic]. According to Littman, a Department of Energy supercomputer was used to find the key, a task that took several months at an estimated cost of hundreds of thousands of dollars. The effort apparently paid off, however, yielding nearly ten thousand pages of evidence (Littman 1997).

A substantial effort was also required to break the encryption software used by the New York subway bomber, Leary. In that case, the result yielded child pornography and personal information, which was not particularly useful to the case.

Investigators, however, retrieved other evidence from the computer that was used at the trial. Leary was found guilty and sentenced to 94 years in jail.

Timeliness is critical in some investigations. Several years ago, a Bolivian terrorist organization assassinated four U.S. Marines, and AccessData was brought in to decrypt files seized from a safe house. With only twenty four hours to perform this task, they decrypted the custom-encrypted files in twelve, and the case ended with one of the largest drug busts in Bolivian history. The terrorists were caught and put in jail (CSA 1997). In such cases, an effort that requires months or years to complete might be useless.

In other cases, the ability to successfully decrypt files proved unessential, as when a Durham priest was sentenced to six years in jail for sexually assaulting minors and distributing child pornography (Akdeniz). The priest was part of an international pedophile ring that communicated and exchanged images over the Internet. When U.K. authorities seized his computers, they found files of encrypted messages. The encryption was successfully broken, however, the decrypted data did not affect the case.

Even when decrypted material has little or no investigative value, considerable resources are wasted reaching that determination. If all information were encrypted, it would be extremely difficult for law enforcement to decide where to spend precious resources. It would not be practical or even possible to decrypt everything. Yet if nothing were decrypted, many criminals would go free.

Some investigations have been derailed by encryption. For example, at one university, the investigation of a professor thought to be trafficking in child pornography was aborted because the campus police could not decrypt his files. In another case, an employee of a company copied proprietary software to a floppy disk, took the disk home, and then stored the file on his computer encrypted under PGP. Evidently, his intention was to use the software to offer competing services, which were valued at tens of millions of dollars annually (the software itself cost over a million dollars to develop). At the time we heard about the case, the authorities had not determined the passphrase needed to decrypt the files. Information contained in logs had led them to suspect the file was the pilfered software.

At Senate hearings in September 1997, Jeffery Herig, special agent with the Florida Department of Law Enforcement, testified that they were unable to access protected files within a personal finance program in an embezzlement case at Florida State University. He said the files could possibly hold useful information concerning the location of the embezzled funds (US Congress 1997b).

Herig also reported that they had encountered unbreakable encryption in a U.S. Customs case involving an illegal, world-wide advanced fee scheme. At least 300 victims were allegedly bilked out of over $60 million. Herig said they had encountered three different encryption systems. Although they were able to defeat the

first two, they were unsuccessful with the third. The vendor told them that there no backdoors. “Although I have been able to access some of the encrypted data in this case,” Herig said, “we know there is a substantial amount of incriminating evidence which has not been recovered” (ibid).

In early 1997, we were told that Dutch organized crime had received encryption support from a group of skilled hackers who themselves used PGP and PGPfone to encrypt their communications. The hackers had supplied the mobsters with palmtop computers on which they installed Secure Device, a Dutch software product for encrypting data with IDEA. The palmtops served as an unmarked police/intelligence vehicles database. In 1995, the Amsterdam Police captured a PC in the possession of one organized crime member. The PC contained an encrypted partition, which they were unable to recover at the time. Nevertheless, there was sufficient other evidence for conviction. The disk, which was encrypted with a U.S. product, was eventually decrypted in 1997 and found to be of little interest.

There have been a few reported cases of company insiders using encryption as a tool of extortion. The employees or former employees threatened to withhold the keys to encrypted data unless payment was made. In these cases, encryption is not used to conceal evidence of crimes, but rather to intimidate the organization. We are not aware of any extortion attempts of this nature that succeeded.

The use of encryption by the victims of crime can also pose a problem for law enforcement. At hearings in June 1997, Senator Charles Grassley told of an 11-year-old boy in the Denver area who committed suicide after being sexually molested. The boy had left behind a personal organizer, which investigators believed might contain information about the man whom his mother believed molested him. The organizer was encrypted, however, and the police had been unable to crack the password. The investigation had been on hold since February 1996.

In April 1998, the FBI’s Computer Analysis Response Team (CART) forensics laboratory started collecting data on computer forensics cases handled at headquarters or in one of the FBI’s field offices. As of December 9, they had received 299 examination reporting forms, of which 12 (4%) indicated use of encryption.[4] This is slightly lower than CART’s estimate of 5-6% for 1996 (Denning and Baugh 1997). There are at least three possible explanations. One is that the 1996 estimate, which was made before the FBI began collecting hard data, was somewhat high. A second is that as computers have become more common and user friendly, they are increasingly being used by criminals who lack the knowledge or skills to encrypt their files. Hence, the percentage of computer forensics cases involving encryption is staying about the same or decreasing even as the total number of forensics cases (and encryption cases) is growing. A third is that the early reports are skewed; as more come in, the percentage could approach 5-6%.

Public Postings

Criminals can use encryption to communicate in secrecy through open forum such as computer bulletin boards and Internet Web sites. Although many people might see the garbled messages, only those with the key would be able to determine the plaintext.

This technique was used by an extortionist who threatened to kill Microsoft president and chief executive officer Bill Gates in spring 1997.[5] The extortionist transmitted his messages to Gates via letter, but then asked Gates to acknowledge acceptance by posting a specified message on the America Online Netgirl bulletin board. Gates then received a letter with instructions to open an account for a Mr. Robert M. Rath in a Luxemburg bank and to transfer $5,246,827.62 to that account. The money was to be transferred by April 26 in order "to avoid dying, among other things." Gates was reminded that April 26 was his daughter's birthday. The letter came with a disk, which contained an image of Elvira and the key to a simple substitution cipher. Gates was told to use the code to encrypt instructions for accessing the Rath account via telephone or facsimile. He was then to attach the ciphertext to the bottom of the image and post the image to numerous image libraries within the Photography Forum of America Online (AOL). The graphic image with ciphertext was uploaded to AOL at the direction of the FBI on April 25. Figure 1 shows the image as posted and translation code.

Although Gates complied with the requests, he did not lose his money. The extortion threat was traced to Adam Quinn Pletcher in Long Grove, Illinois. On May 9, Pletcher admitted writing and mailing the threatening letters (there were four altogether) to Gates.

LAW ENFORCEMENT OPTIONS

The majority of investigations we heard about were not stopped by encryption. Authorities obtained the key by consent, found it on disk, or cracked the system in some way, for example, by guessing a password or exploiting a weakness in the overall system. Alternatively, they used other evidence such as printed copies of encrypted documents, other paper documents, unencrypted conversations and files, witnesses, and information acquired through other, more intrusive, surveillance technologies such as bugs. We emphasize, however, that these were cases involving computer searches and seizures, not wiretaps. This section discusses the options available to law enforcement for dealing with encryption.

Getting Key From Subject

In many cases, subjects have cooperated with the police and disclosed their keys or passwords, sometimes as part of a plea bargain. One hacker who had encrypted his files with the Colorful File System confessed to his crimes and revealed his CFS passphrase:

ifyoucanreadthisyoumustbeerikdale--**oragoodcypherpunk

He (Erik) wanted to speed the process along. The decrypted files contained evidence that was important to the case.[6]

A question that frequently arises is whether a court can compel the disclosure of plaintext or keys, or whether the defendants are protected by the 5th Amendment. Philip Reitinger, an attorney with the Department of Justice Computer Crime Unit, studied this question and concluded that a grand jury subpoena can direct the production of plaintext or of documents that reveal keys, although a limited form of immunity may be required (Reitinger 1996). He left open the question of whether law enforcement could compel production of a key that has been memorized but not recorded. He also observed that faced with the choice of providing a key that unlocks incriminating evidence or risking contempt of court, many will choose the latter and claim loss of memory or destruction of the key.

In People v. Price in Yolo County, California Superior Court prosecutors successfully compelled production of the passphrase protecting the defendant's PGP key. In this case, however, the key was not sought for the purpose of acquiring evidence for conviction, but rather to determine whether the defendant's computer should be released from police custody. He had already been convicted of annoying children and wanted his computer back. The police argued it should not be released as
there was reason to believe it contained contraband, specifically PGP-encrypted files containing child pornography. This determination was based on the existence of a pair of files named "Boys.gif" and "Boys.pgp" (when PGP encrypts a plaintext file, it automatically gives the ciphertext file the same name but with the extension ".pgp").[7]

The defendant was unsuccessful in arguing a 5th Amendment privilege. The prosecution argued that the contents of the file had already been uttered and, therefore, were not protected under the 5th Amendment. As long as prosecutors did not try to tie the defendant to the file by virtue of his knowing the passphrase, no incrimination was implied by disclosing the passphrase.

To handle the passphrase, a court clerk was sworn in as a special master. An investigator activated the PGP program to the point where it prompted for the passphrase. He left the room while the defendant disclosed the passphrase to the special master, who typed it into the computer. The investigator was then brought back into the room to hit the Enter key and complete the decryption process. As expected, child pornography fell out. The judge then ordered the computer, its peripherals, and all diskettes destroyed. The defendant argued that the computer contained research material, but the judge admonished him for commingling it with the contraband.

Getting Access Through a Third Party

Some encryption products have a key recovery system which enables access to plaintext through a means other than the normal decryption process. The key needed to decrypt the data is recovered using information stored with the ciphertext plus information held by a trusted agent, which could be an officer of the organization owning the data or a third party. The primary objective is to protect organizations and individuals using strong encryption from loss or destruction of encryption keys, which could render valuable data inaccessible.

Key recovery systems can accommodate lawful investigations by proving authorities with a means of acquiring the keys needed. If the keys are held by a third party, this can be done without the knowledge of the criminal group under investigation. Of course, if criminal enterprises operate their own recovery services, law enforcement may be no better off. Indeed, they could be worse off because the encryption will be much stronger, possibly uncrackable, and the criminals might not cooperate with the authorities. Moreover, with wiretaps, which must be performed surreptitiously to have value, investigators cannot go to the subjects and ask for keys to tap their lines. Key recovery systems could also encourage the use of encryption in organized crime to protect electronic files, as criminal enterprises need not worry about loss of keys.

Because of the potential benefits of key recovery to law enforcement, the Clinton Administration has encouraged the development of key recovery products by offering an export advantages to companies making such products. Beginning in December 1996, products with key recovery systems could be readily exported with unlimited key lengths. The Administration has retained restrictions on non-recoverable products that use keys longer than 56 bits, but even here export controls have been liberalized to allow ready export under certain conditions.

Breaking the Codes

It is often possible to obtain the key needed to decrypt data by exploiting a weakness in the encryption algorithm, implementation, key management system, or some other system component. Indeed, there are software tools on the Internet for cracking the encryption in many commercial applications. One site on the World Wide Web lists freeware crackers and products from AccessData Corp. and CRAK Software for Microsoft Word, Excel, and Money; WordPerfect, Data Perfect, and Professional Write; Lotus 1-2-3 and Quattro Pro; Paradox; PKZIP; Symantex Q&A, and Quicken.[8]

Eric Thompson, president of AccessData, reported that they had a recovery rate of 80-85 per cent with the encryption in large-scale commercial commodity software applications. He also noted that 90 per cent of the systems are broken somewhere other than at the crypto engine level, for example, in the way the text is pre-processed (CSI 1997). A passphrase or key might be found in the swap space on disk.

In those cases where there is no shortcut attack, the key might be determined by brute force search, that is, by trying all possible keys until one is found that yields

known plaintext or, if that is not available, meaningful data. Keys are represented as strings of 0s and 1s (bits), so this means trying every possible bit combination. This is relatively easy if the keys are no more than 40 bits, and somewhat longer keys can be broken given enough horsepower. In July 1998, John Gilmore, a computer privacy and civil liberties activist, and Paul Kocher, president of Cryptography Research in California, won $10,000 for designing a supercomputer that broke a 56-bit DES challenge cipher in record time, in their case 56 hours or less than three days. The EFF DES Cracker was built by a team of about a dozen computer researchers with funds from the Electronic Frontier Foundation. It took less than a year to build and cost less than $250,000. It tested keys at a rate of almost 100 billion per second (EFF 1998; Markoff 1998).

Unfortunately, criminals can protect against such searches by using methods that take longer keys, say 128 bits with the RC4, RC5, or IDEA encryption algorithm or 168 bits with Triple DES. Because each additional bit doubles the number of candidates to try, a brute force search quickly becomes intractable. To crack a 64-bit key, it would take 10 EFF DES Crackers operating for an entire year. At 128 bits, it is totally infeasible to break a key by brute force, even if all the computers in the world are put to the task. To break one in a year would require, say, 1 trillion computers (more than 100 computers for every person on the globe), each running 10 billion times faster than the EFF DES Cracker. Put another way, it would require the equivalent of 10 billion trillion DES Crackers! Many products, including PGP, use 128-bit keys or longer.

With many encryption systems, for example PGP, a user's private key (which unlocks message keys) is computed from or protected by a passphrase chosen by the user. In that case, it may be easier to brute force the password than the key because it will be limited to ASCII characters and be less random than an arbitrary stream of bits. Eric Thompson reports that the odds are about even of successfully guessing a password. They use a variety of techniques including Markov chains, phonetic generation algorithms, and concatenation of small words (CIS 1997).

Often, investigators will find multiple encryption systems on a subject's computer. For example, PGP might be used for e-mail, while an application's built-in encryption might be used to protect documents within the application. In those cases, the subject might use the same password with all systems. If investigators can break one because the overall system is weak, they might be able to break the other, more difficult system by trying the same password.

To help law enforcement develop the capability to stay abreast of new technologies, including encryption, the Federal Bureau of Investigation proposes to establish a technical support center. The center would maintain a close working relationship with the encryption vendors. The Clinton Administration announced support for the center in its September 1998 update on encryption policy (White House 1998).

One issue raised by the development and use of tools for breaking codes is how law enforcement can protect its sources and methods. If investigators must reveal in court the exact methods used to decipher a message, future use of such methods could be jeopardized.

Finding an Access Point

Another strategy for acquiring plaintext is to find an access point that provides direct access to the plaintext before encryption or after decryption. In the area of communications, a router or switch might offer such access to communications that traverse the switch. If the communications are encrypted on links coming into and going out of the switch, but in the clear as they pass through the switch, then a wiretap placed in the switch will give access to the plaintext communications. We noted earlier how digital cellular communications could be intercepted in this manner, while at the same time offering users considerably greater security and privacy than offered by analog phones that do not use encryption.

Network encryption systems which offer access points of this nature are given an export advantage over those that do not (ibid). The approach was initially called a "private doorbell" approach to distinguish it from one that uses key recovery agents (Corcoran 1998; CISCO 1998). Now it is considered a form of recoverable encryption.

For stored data, Codex Data Systems of Bardonia, New York, advertises a product called Data Interception by Remote Transmission (D.I.R.T.) which is designed to allow remote monitoring of a subject's personal computer by law enforcement and other intelligence gathering agencies. Once D.I.R.T. is installed on the subject's machine, the software will surreptitiously log keystrokes and transmit captured data to a pre-determined Internet address that is monitored and decoded by D.I.R.T. Command Center Software. D.I.R.T. add-ons include remote file access, real-time capture of keystrokes, remote screen capture, and remote audio and video capture. The software could be used to capture a password and read encrypted e-mail traffic and files.

When All Else Fails

The inability to break through encryption does not always spell doom. Investigators may find printed copies of encrypted documents. They may find the original plaintext version of an encrypted file, for example, if the subject forgot to delete the original file or if it was not thoroughly erased from the disk. They may obtain incriminating information from unencrypted conversations, witnesses, informants, and hidden microphones. They may conduct an undercover or sting operation to catch the subject. These other methods do not guarantee success, however.

If there is sufficient evidence of some crime, but not the one believed to be concealed by encryption, a conviction may be possible on lesser charges. This happened in Maryland when police encountered an encrypted file in a drug case. Allegations were raised that the subject had been involved in document counterfeiting and file names were consistent with formal documents. Efforts to decrypt the files failed, however, so the conviction was on the drug charges only.[9]

In another case, a 15-year-old boy came to the child abuse bureau of the Sacramento County Sheriff's Department with his mother, who desired to file a complaint against an adult who had met her son in person, befriending the boy and his friends and buying them pizza. The man had sold her son $500-$1000 worth of hardware and software for $1.00 and given him lewd pictures on floppy disks. The man subsequently mailed her son pornographic material on floppy disk and sent her son pornographic files over the Internet using America Online. After three months of investigation, a search warrant was issued against a man in Campbell, California and the adoption process of a 9-year-old boy was stopped. Eventually, the subject was arrested, but by this time he had purchased another computer system and traveled to England to visit another boy. Within ten days of acquiring the system, he had started experimenting with different encryption systems, eventually settling on PGP. He had encrypted a directory on the system. There was information indicating that the subject was engaged in serious corporate espionage, and it was thought that the encrypted files might have contained evidence of that activity. They were never able to decrypt the files, however, and after the subject tried unsuccessfully to put a contract out on the victim from jail, he pled no contest to multiple counts of distribution of harmful material to a juvenile and the attempt to influence, dissuade, or harm a victim/witness.[10]

If encryption precludes access to all evidence of wrongdoing, then a case is dropped (assuming other methods of investigation have failed as well). Several cases that had been aborted or put on hold because of encryption were noted earlier.

OTHER TECHNOLOGIES FOR HIDING EVIDENCE

The modern day criminal has access to a variety of tools for concealing information besides encryption:

Passwords

Criminals, like law abiding persons, often password protect their machines to keep others out. In one gambling operation with connections to New York's Gambino, Genovese, and Colombo crime families, bookies had password-protected a computer used to cover bets at the rate of $65 million a year (Ramo 1996). After discovering that the password was one of the henchmen's mother's name, the cops found 10,000 digital betting slips worth $10 million.

Another gambling enterprise operated multiple sites linked by a computer system, with drop-offs and pick-ups spanning three California counties. The ring leader managed his records with a commercial accounting program, using a password to control access to his files. Although the software manufacturer refused to assist law enforcement, police investigators were able to gain access by zeroing out the passwords in the data files. They found the daily take on bets, payoffs, persons involved, amounts due and paid or owed, and so forth. The printed files showed the results of four years of bookmaking, and resulted in a plea of guilty to the original charges and a sizeable payment of back taxes, both state and federal.[11]

Passwords are encountered much more often than encryption in computer forensics cases. Of the 299 computer examination reports received by the FBI's CART between April and December 1998, 60 (20 per cent) indicated use of passwords. This was five times as many as had indicated use of encryption.[12]

Digital Compression

Digital compression is normally used to reduce the size of a file or communication without losing information content, or at least significant content. The greatest reductions are normally achieved with audio, image, and video data; however, substantial savings are possible even with text data. Compression can benefit the criminal trying to hide information in two ways. First, it makes the task of identifying and accessing information more difficult for the police conducting a wiretap or seizing files. Second, when used prior to encryption, it can make cracking an otherwise weak cipher difficult. This is because the compressed data is more random in appearance than the original data, making it less susceptible to techniques that exploit the redundancy in languages and multimedia formats.

Steganography

Steganography refers to methods of hiding secret data in other data such that its existence is even concealed. One class of methods encodes the secret data in the low-order bit positions of image, sound, or video files. There are several tools for doing this, many of which can be downloaded for free off the Internet. With S-tools, for example, the user hides a file of secret data in an image by dragging the file over the image. The software will optionally encrypt the data before hiding it for an extra layer of security. S-tools will also hide data in sound files or in the unallocated sectors of a disk. Figure 2 shows the effect of using S-tools to hide a 17-page book chapter inside an image file that is less than four times the size; that is, about a quarter of the file contains a hidden document. The difference between the before and after images is barely noticeable.

There have been a few reported cases of criminals using steganography to facilitate their crimes. One credit card thief, for example, used it to hide stolen card numbers on a hacked Web page. He replaced bullets on the page with images that looked the same but contained the credit card numbers, which he then offered to

associates. This case illustrates the potential of using Web images as "digital dead drops" for information brokering. Only a handful of people need even know the drop exists.

Steganography can be used to hide the existence of files on a computer's hard disk. Ross Anderson, Roger Needham, and Adi Shamir propose a steganographic file system that would make a file invisible to anyone who does not know the file name and a password. An attacker who does not know this information gains no knowledge about whether the file exists, even given complete access to all the hardware and software. One simple approach creates cover files so that the user's hidden files are the exclusive or (XOR) of a subset of the cover files. The subset is chosen by the user's password (Anderson et al 1998).

Remote Storage

Criminals can hide data by storing it on remote hosts, for example, a file server at their Internet Service Provider (ISP). Jim McMahon, former head of the High Technology Crimes Detail of the San Jose Police Department, reported that he had personally seen suspects hiding criminal data on non-local disks, often at ISP locations, but sometimes on the systems of innocent third parties with poor security, leaving them open to intrusions and subsequent abuse. Eugene Schultz, former manager of the Department of Energy's Computer Incident Advisory Capability, said that a group of hackers from the Netherlands had taken so much information from Defense Department computers that they could not store it all on their own disks. So they broke into systems at Bowling Green University and the University of Chicago and downloaded the information to these sites, figuring they could transfer it somewhere else later.[13] Software pirates have been known to stash their pilfered files in hidden directories on systems they have hacked.

Data can be hidden on removable disks and kept in a physical location away from the computers. Don Delaney, a detective with the New York State Police, told us in early 1997 that in one Russian organized crime case involving more than $100 million in state sales tax evasion, money laundering, gasoline bootlegging, and enterprise corruption, they had to obtain amendments to their search warrants in order to seize disks and records from handbags and locked briefcases in the offices at two locations. After an exhaustive six month review of all computer evidence, they determined that the largest amount of the most damaging evidence was on the diskettes. The crooks did their work in Excel and then saved it on floppies. The lesson they learned from this was to execute the search warrant with everyone present and look for disks in areas where personal property is kept. As storage technologies continue to get smaller, criminals will have even more options for hiding data.

Audit Disabling

Most systems keep a log of activity on the system. Perpetrators of computer crimes have, in many cases, disabled the auditing or deleted the audit records pertaining to

their activity. The hacking tool RootKit, for example, contains Trojan horse system utilities which conceal the presence of the hacker and disable auditing. ZAP is another tool for erasing audit records. Both of these can be downloaded for free on the Internet.

CONCEALING CRIMES THROUGH ANONYMITY

Crimes can be concealed by hiding behind a cloak of anonymity. A variety of technologies are available:

Anonymous Remailers

An anonymous remailer is a service that allows someone to send an electronic mail message without the receiver knowing the sender's identity. The remailer may keep enough information about the sender to enable the receiver to reply to the message by way of the remailer. To illustrate, suppose Alice wishes to send an anonymous e-mail message to Bob. Instead of e-mailing to Bob directly, Alice sends the message to a remailer (an e-mail server), which strips off the headers and forwards the contents to Bob. When Bob gets the message, he sees that it came via the remailer, but he cannot tell who the sender was. Some remailers give users pseudonyms so that recipients can reply to messages by way of the remailer. The remailer forwards the replies to the owners of the pseudonyms. These pseudo anonymous remailers do not provide total anonymity because the remailer knows who the parties are. Other remailers offer full anonymity, but they cannot support replies. All they do is act as a mail forwarder.

A remailer can accumulate batches of messages before forwarding them to their destinations. That way, if someone is intercepting encrypted Internet messages for the purpose of traffic analysis, the eavesdropper would not be able to deduce who is talking to whom.

There are numerous anonymous and pseudo anonymous remailers on the Internet. Some provide encryption services (typically using PGP) in addition to mail forwarding so that messages transmitted to and from the remailer can be encrypted. Users who don't trust the remailers can forward their messages through multiple remailers.

Anonymous remailers allow persons to engage in criminal activity while concealing their identities. President Clinton, for example, has received e-mail death threats that were routed through anonymous remailers. In one case involving remailers, an extortionist threatened to fly a model airplane into the jet engine of an airplane during takeoff at a German airport, the objective being to cause the plane to crash. The threats were sent as e-mail through an anonymous remailer in the United States. The messages were traced to introductory accounts on America Online, but the person had provided bogus names and credit card numbers. He was caught, however, before carrying out his threat.[14]

Anonymous Digital Cash

Digital cash enables users to buy and sell information goods and services. It is particularly useful with small transactions, serving the role of hard currency. Some methods allow users to make transactions with complete anonymity; others allow traceability under exigent circumstances, for example, a court order.

Total anonymity affords criminals the ability to launder money and engage in other illegal activity in ways that circumvent law enforcement. Combined with encryption or steganography and anonymous remailers, digital cash could be used to traffic in stolen intellectual property on the Web or to extort money from victims.

In May 1993, Timothy May wrote an essay about a hypothetical organization, BlackNet, which would buy and sell information using a combination of public-key cryptography, anonymous remailers, and anonymous digital cash.

> 'BlackNet can make anonymous deposits to the bank account of your choice, where local banking laws permit, can mail cash directly ..., or can credit you in 'CryptoCredits,' the internal currency of BlackNet ... If you are interested, do not attempt to contact us directly (you'll be wasting your time), and do not post anything that contains your name, your e-mail address, etc. Rather, compose your message, encrypt it with the public key of BlackNet (included below), and use an anonymous remailer chain of one or more links to post this encrypted, anonymized message on one of the locations listed ..." (May 1996a).

Although May said he wrote the essay to point out the difficulty of "bottling up" new technologies (May 1996b), rumors spread shortly after May's essay appeared on the Internet of actual BlackNets being used for the purpose of selling stolen trade secrets.

In an essay called 'Assassination Politics,' James Dalton Bell suggested using cyber betting pools to kill off Internal Revenue Service (IRS) agents and other 'hated government employees and officeholders' (Bell 1996).[15] The idea was simple: using the Internet, encryption, and untraceable digital cash, anyone could contribute anonymously to a pool of digital cash. The person, presumably the assassin, correctly guessing the victim's time of death wins. After spending nearly two years peddling his ideas on Internet discussion groups and mailing lists, Bell was arrested and pled guilty to two felony charges: obstructing and impeding the IRS and falsely using a social security number with the intent to deceive. In his plea agreement, he admitted to conducting a "stink bomb" attack on an IRS office in Vancouver (McCullah 1997).[16] He also disclosed the passphrase required to decrypt e-mail messages that had been sent to Bell by his associates encrypted under PGP.

Although Bell did not implement any betting pools, an anonymous message was posted to the Cypherpunks Internet mailing list announcing an Assassination Politics Bot (program) called Dead Lucky that did. The message also listed four potential targets. A related message pointed to an interactive Web page titled Dead Lucky,

which contained the statement 'If you can correctly predict the date and time of death of others then you can win large prizes payable in untaxable, untraceable eca$h.' The page also stated 'Contest will officially begin after Posting of Rules and Announcement of Official Starting Date (Until then it is for Entertainment Purposes Only).' Another anonymous message posted to Cypherpunks had the subject 'Encrypted InterNet DEATH THREAT!!! / ATTN: Ninth District Judges / PASSWORD: sog.' The PGP encrypted message, when decrypted with 'sog,' contained death threats and a claim to authorship of the Assassination Bot. Investigators linked the messages and Bot to an individual by the name of Carl Edward Johnson. In August 1998, a warrant was issued charging Johnson with threatening 'to kill certain law enforcement officers and judges of the United States, with intent to impede, intimidate, or interfere with said officers and judges on account of their official duties.'[17]

Computer Penetrations and Looping

By breaking into someone's computer account and issuing commands from that account, a criminal can hide behind the account holder's identity. In one such case, two hackers allegedly penetrated the computers of Strong Capital Management and sent out 250,000 ads with fraudulent headers that bore the company's name. The ads were for on-line striptease services ('cyber stripping'), computer equipment, and sports betting. SCM filed a $125 million lawsuit against the hackers, demanding penalties of $5,000 per message (Kabay 1997).

Hackers can make it difficult for investigators to discover their true identity by using a technique called looping.' Instead of penetrating a particular system directly, they can enter one system and use that as a springboard to penetrate another, use the second system to penetrate a third, and so forth, eventually reaching their target system. The effect is to conceal the intruder's location and complicate an investigation. In order to trace the connection, investigators need the help of systems administrators along the path. If the path crosses several national borders, getting that cooperation may be impossible.

Cellular Phones and Cloning

Drug lords, gangsters, and other criminals regularly use "cloned" cell phones to evade the police. Typically, they buy the phones in bulk and discard them after use. A top Cali cartel manager might use as many as 35 different cell phones a day (Ramo 1996). In one case involving the Colombia cartel, DEA officials discovered an unusual number of calls to Colombia on their phone bills. It turned out that cartel operatives had cloned the DEA's own number! Some cloned phones, called 'lifetime phones,' hold up to 99 stolen numbers. New numbers can be programmed into the phone from a keypad, allowing the user to switch to a different cloned number for each and every call. With cloning, whether cellular communications are encrypted may have little impact on law enforcement, as they do not even know which numbers to tap.

Digital cellular phones use stronger methods of authentication that protect against cloning. As this technology replaces analog cell phones, cloning may be less of a problem for law enforcement.

Cellular Phone Cards

A similar problem occurs with cellular phone cards. These pre-paid cards, which are inserted into a mobile phone, specify a telephone number and amount of air time. In Sweden, phone cards can be purchased anonymously, which has made wiretapping impossible. The narcotics police have asked that purchasers be required to register in a database that would be accessible to the police (Minow 1997). A similar card is used in France, however buyers must show an identification card at the time of purchase. In Italy, a pre-paid card must be linked to an identity, which must be linked to an owner.

CONCLUSIONS

Criminals and terrorists are using encryption and other advanced technologies to hide their activities. Indications are that use of these technologies will continue and expand, with a growing impact on law enforcement. Although the majority of investigations we heard about were not stopped by encryption, we heard about a few cases that were effectively derailed or put on hold by encryption. Even when the encryption was broken, however, it delayed investigations, sometimes by months or years, and added to their cost, in a few cases costing agencies hundreds of thousands of dollars to crack open encrypted files.

Efforts to decrypt data for law enforcement agencies or corporations in need of recovering from lost keys have been largely successful because of weaknesses in the systems as a whole. That success rate is likely to drop, however, as vendors integrate stronger encryption into their products and get smarter about security. It is not possible to break well-designed cryptosystems that use key lengths of 128 bits or more. It is not just a matter of paying enough money or getting enough people on the Internet to help out. The resources simply do not exist - anywhere.

Most of the investigators we talked to said that they had not yet detected substantial use of encryption by large organized crime groups. This can be attributed to several factors, including the difficulty and overhead of using encryption (particularly the personnel time involved) and a general sense that their environments are already reasonably isolated and protected from law enforcement.

Maria Christina Ascents, who runs the Italian state police’s crime and technology center, said that the Italian Mafia is increasingly looking to use encryption to help protect it from the government. She cited encryption as their greatest limit on investigations, and noted that instead of hiring cryptographers to create their codes,

mobsters download copies of Pretty Good Privacy (PGP) off the Internet (Ramo 1996).

As the population becomes better educated about technology and encryption, more and more criminals will have the knowledge and skills needed to evade law enforcement, particularly given the ease with which unbreakable, user-friendly software encryption can be distributed and obtained on the Internet. We recommend ongoing collection of data on the use of encryption and other advanced technologies in crime. We need to know how encryption is impacting cases - whether it is broken or circumvented, whether cases are successfully investigated and prosecuted despite encryption, and costs to investigators.

Encryption is a critical international issue with severe impact and benefits to business and order. National policy must recognize not only the threat to law enforcement and intelligence operations, but also the need to protect the intellectual property and economic competitiveness of industry. Encryption policy must also respect consumer needs for encryption and basic human rights, including privacy and freedom of expression. Addressing all of these interests is enormously challenging.

NOTES

[1] The chapter is an update of a study we conducted in 1997 at the invitation of the U.S. Working Group on Organized Crime, National Strategy Information Center, Washington, DC.

[2] Additional information was provided by Detective R. J. Montemayor in the Dallas Police Department.

[3] The key used by Ames was disclosed to us by Robert Reynard on February 18, 1998.

[4] Data provided by CART on December 9, 1998.

[5] United States District Court, Northern District of Illinois, Eastern Division, Search Warrant, Case Number 97-157M, May 8, 1997; United States of America v. Adam Quinn Pletcher, United States District Court, Western District of Washington at Seattle, Magistrate's Docket No. Case No. 97-179M, May 9, 1997.

[6] Byron W. Thompson, presentation at HTCIA/FBI Training Seminar, Perspectives on Computer Crime, November 12-13, 1998.

[7] Information on this case was provided by Fred B. Cotton of SEARCH Group, Inc. Cotton was the investigator who activated the PGP program on the defendant's computer.

[8] http://www.hiwaay.net/boklr/bsw_crak.html as of February 1997.

[9] This case was reported to us by Howard Schmidt.

[10] This case was reported by Brian Kennedy of the Sacramento County Sheriff's Department.

[11] This case was first reported to us on February 22, 1997 by Jim McMahon, former head of the High Technology Crimes Detail of the San Jose Police Department. We received additional information from Robert Reynard on June 10, 1998.

[12] Data provided by CART on December 9, 1998.

[13] Communication from Eugene Schultz, May 15, 1998.

[14] Presentation by Christoph Fischer at Georgetown University, July 22, 1998.

[15] A version of Bell's essay on Assassination Politics is in Winn Schwartau, *Information Warfare*, 2nd ed., Thunder's Mouth Press, 1996, pp. 420-425.

[16] http://jya.com/jimbell3.htm.

[17] United States of America v. Carl Edward Johnson, Warrant for Arrest, Case No. 98-430M, United States District Court, Western District of Washington, August 19, 1998.

References

Akdeniz, Y., 'Regulation of Child Pornography on the Internet,' http://www.leads.ac.uk/law/pgs/yaman/child.htm.

Anderson, R., Needham, R., and Shamir, A. (1998) 'The Steganographic File System,' presented at the Workshop on Information Hiding, Portland, OR, April 14-17.

Cisco Systems Inc. (1998) 'Thirteen High-Tech Leaders Support Alternative Solution to Network Encryption Stalemate,' Press Release, July 13.

Corcoran, E. (1998) 'Breakthrough Possible in Battle over Encryption Technology,' Washington Post, July 12.p. A8.

CSI (1997) 'Can your crypto be turned against you? An interview with Eric Thompson of AccessData,' Computer Security Alert, No. 167, February.

Denning, D. E. and Baugh, W. E., Jr. (1997) 'Encryption and Evolving Technologies as Tools of Organized Crime and Terrorism,' National Strategy Information Center, Washington, DC, July.

EEF (1998) '"EFF DES Cracker" Machine Brings Honesty to Crypto Debate,' press announcement from the Electronic Frontier Foundation, July 17.

Fischer, C. (1998) Presentation at Georgetown University, July 22.

FBI Law Enforcement Bulletin (1970) 'Crime and Cryptology', April, 13-14.

Grabosky, P. N. and Smith, R. G. (1998) Crime in the Digital Age: Controlling Telecommunications and Cyberspace Illegalities, Transaction Publishers.

IINS News Service, (1997) 'Hamas Using Internet for Attack Instructions', Israel, September 28.

Kaplan, D. E. and Marshall, A. (1996) The Cult at the End of the World, Crown Publishers.

Littman, J. (1997) The Watchman: The Twisted Life and Crimes of Serial Hacker Kevin Poulson, Little, Brown and Co.

Manning, W. M. (1997) 'Should You Be on the Net?' FBI Law Enforcement Bulletin, January, 18-22.

Markoff, J. (1998) 'U.S. Data-Scrambling Code Cracked with Homemade Equipment,' New York Times, July 17.

May, T. C. (1996a) 'Introduction to BlackNet,' reprinted in, Ludlow, P (ed), High Noon on the Electronic Frontier, MIT Press, pp. 241-243.

May, T. C. (1996b) 'BlackNet Worries,' in Peter Ludlow, (ed), High Noon on the Electronic Frontier, MIT Press, pp. 245-249.

McCullah, D. (1997) 'IRS Raids a Cypherpunk,' The Netly News, April 4.

Minow, M. (1997) 'Swedish Narcotics Police Demand Telephone Card Database,' Risks-Forum Digest, Vol. 19, Issue 07, April 14.

Power, R. (1997) 'CSI Special Report: Salgado Case Reveals Darkside of Electronic Commerce,' Computer Security Alert, No. 174, September.

Ramo, J. C. (1996) 'Crime Online,' Time Digital, September 23, pp. 28-32.

Reitinger, P. R. (1996) 'Compelled Production of Plaintext and Keys.'

US Congress (1997a) Statement of Louis J. Freeh, Director FBI, before the Senate Committee on Commerce, Science, and Transportation, regarding the Impact of Encryption on Law Enforcement and Public Safety, March 19.

US Congress (1997b) Jeffrey A. Herig, Special Agent, Florida Department of Law Enforcement, "The Encryption Debate: Criminals, Terrorists, and the Security Needs of Business and Industry," testimony before the Senate Judiciary Subcommittee on Technology, Terrorism, and Government Information, September 3.

White House (1995) Remarks by the President to Staff of the CIA and Intelligence Community, Central Intelligence Agency, McLean, VA, July 14.

White House (1998) 'Administration Updates Encryption Policy', statement by the Press Secretary and fact sheet, September.

ABOUT THE AUTHORS

Dorothy E. Denning is professor of Computer Science and Communication, Culture, and Technology at Georgetown University. She is author of *Information Warfare and Security*, Addison Wesley, 1999. E-mail: denning@cs.georgetown.edu. Web: www.cs.georgetown.edu/~denning.

William E. Baugh, Jr. is vice president, Science Applications International Corporation, and general manager, Advanced Network Technologies and Security Operations. He is former assistant director, Federal Bureau of Investigation. E-mail: William.E.Baugh.Jr@cpmx.saic.com.

IV

Censorship and Sysop Liability

Peter Ludlow

This section takes up the difficult issue of censorship on the Internet. As the reading from *Computer and Academic Freedom News* shows, even in academia any number of items have been subject to censorship by various universities. Sometimes sexually explicit materials are censored, sometimes hate speech is censored, sometimes students are barred from having certain kinds of computer code in their accounts.

In the main, however, censorship tends to be directed at sexually explicit materials. One recent case of such censorship is the banning of certain Usenet user groups by Carnegie Mellon University (see the *Time* article by Philip Elmer-Dewitt). Fearing a lawsuit, the university originally banned some 80 newsgroups in the alt.sex hierarchy. Some of these usergroups contained digitized photographs, but they also included newsgroups that were dedicated to discussions of safe sex, sexual problems, and such matters. The university soon restored the groups devoted to discussion, but many observers found it remarkable that the administration of a university such as CMU did not initially see that such discussions were constitutionally protected free speech (see the ACLU press release).

In a sense, the initial action by CMU belies an unspoken assumption that seems to have widespread currency—an assumption that electronic forms of speech should not enjoy the same protections that the printed and spoken word do. There is a reluctance to see that the sexually explicit material being banned in electronic form is often available in the campus library or the campus bookstore, or is possibly being assigned in freshman literature courses. Why should the electronic word be exempt from constitutional protections?

Strictly speaking, of course, there is no serious answer to the above question. Protected speech should be protected speech, no matter what its form (see the reading by Shallit for a discussion of this point). Speech in electronic form presents more of a threat, however, and this is perhaps what lies at the bottom of a number of censorship actions. The threat is that electronic words and images are so readily located, copied, and transmitted.

Consider the college student set on extracting graphic S&M literature from the campus library. Where would the student begin? That is not so clear. On the Internet, however, the student might simply drop in on the

alt.sex.bondage usergroup. There is a sense in which the Internet makes information too available, too "in your face."

This, of course, suggests that free access to information was never taken seriously. It was perhaps acceptable to allow underground presses to grind out a few hundred pamphlets, but it is quite another for such radical information to be available to everyone in the world, and at virtually no expense. Likewise, it is one thing to allow certain books to be housed in libraries so long as they are lost in the stacks with millions of other volumes, but to have sexually explicit or radical material so readily available on the Internet appears to evoke considerable concern, not only from university administrators, but from government legislators as well. It may be that many institutions tolerate free expression only so long as there are structural barriers limiting its distribution and influence. In other words, perhaps they were never advocates of free expression in the first place.

A good example of this is the recent media hand-wringing over the presence of bomb recipes on the Internet. Sometimes the concern is that minors might get access to such recipes, but one also hears the concern voiced that this sort of information need not be protected and that it ought to be banned. Of course bomb recipes have been circulated among children for years, and I remember injuries to children from pipe bombs and so forth being common when I was in grade school (long before the Internet came into existence). For all we know, the distribution of technical information about bomb making has slowed since children got onto the Internet and began pursuing other technical interests. So what is the real concern? It appears to be that the Internet makes this information just too *available* (quite apart from the issue of whether the information is actually used). That is, the real danger is perceived to lie in the information itself (not in the bombs that may be constructed) and in particular in the ready availability of that information.

Ease of access seems to be the subtext in a number of discussions about troublesome materials on the Internet. Parents will claim that pornographic materials, bomb recipes, and so on, are, in a sense, being broadcast directly into their homes. It is not like the old days when Junior had to cross the tracks and go to the other side of town to find such materials. Now the materials are available to children with computers with Internet access, and that means upper-middle-class children (hence the keen interest by the media).

In a sense the parents are right when they think that the materials are coming right into their homes. True, in some cases the child must FTP or Gopher to some remote site, but it is increasingly possible to use a graphical interface such as Netscape, in which one merely points and clicks on an icon. Phenomenologically, the selected materials may as well be in a desktop computer. Indeed, users may have no idea what the remote location is, nor, in many cases, will they even care.

The transparent feel of the new interfaces is bound to exacerbate other problems—in particular the problem of maintaining community standards. For example, it is possible that a web site in New York could contain images that would violate community standards in Eagle Grove, Iowa. What happens when the citizens of Eagle Grove learn that these materials (which local standards take to be obscene) are available at the click of a button?

Even without transparent interfaces, problems of this nature have arisen. In "Virtual Community Standards" Godwin discusses one such recent case, in which a U.S. attorney in Memphis charged the operators of a Milpitas, California, BBS with violating Memphis community standards. The BBS did indeed carry hard-core pornography, some of which allegedly violated Memphis standards of decency and which was downloaded by a Tennessee postal inspector. In this case, however, the California site was not on the Internet, and hence there was no transparent interface. Indeed, the postal inspector had to telephone California in order to log on and download the materials. In Godwin's judgment, the decision to prosecute in this case turned the whole community standards principle on its head.

> The case sends a frightening message to virtual communities: "It doesn't matter if you're abiding by your own community's standards—you have to abide by Memphis's as well."

But, on the assumption that there is a genuine problem here, what is to be done about it? Should we ban everything that isn't safe for children and the residents of Memphis? That hardly seems like an appropriate response, but it is not without advocates (although the advocates would probably describe their positions in other ways). Alternatively, should

children and the residents of Memphis be banned from the electronic frontier? Again, this hardly seems appropriate.

One approach has been to hold that BBS system operators should be responsible for who logs onto the system. In the article on the Memphis case, Godwin touches on the problems with trying to screen users in this way. Consider the difficulty of keeping minors off of the system. In a world where minors routinely enter bars with fake IDs, there are bound to be loopholes in any such screening process.

That isn't the only problem that system operators have, however. System operators are also routinely held responsible for ensuring that illegal material such as child pornography is not uploaded onto their systems. But it is far from clear that a sysop is able to screen everything that is uploaded onto a system. Some electronic bulletin boards are vast, with hundreds if not thousands of files being posted daily. It is possible that a user might upload pilfered credit information, even child pornography, without the knowledge of the sysop (see Godwin's "Sex and the Single Sysadmin"). Yet it appears that, for all practical purposes, the government holds the system operator responsible for the material posted on the system. Just how fair is this? After all, we wouldn't hold the telephone company responsible if pilfered credit information were transmitted via telephone, nor would we hold the post office responsible if child pornography were sent via mail. So why is the sysop responsible for everything that appears on his or her system (a point that applies also to pirated software, by the way)?

There are intermediate positions here. One might argue that sysops be responsible for seeing that their systems are not routinely used for distributing certain materials, while excusing them for the odd file that gets posted to the system. Drawing the line is difficult here, however, as it is not always clear where the odd upload ends and a policy of tolerating such uploads begins.

This discussion has granted a certain assumption—that censorship is possible on the electronic frontier. There are those who would argue otherwise, suggesting that there are just too many ways to bypass attempted censorship. For example, if CMU had banned the alt.sex hierarchy, it would have been a simple matter to telnet to a remote system that continued to carry it. Even if we were to ban the alt.sex hierarchy from every system in the United States it would still be possible to telnet to a remote location in, say Denmark, which continued to carry the "offending" material.

This, then, is the problem that concerned parents and university administrators must face. They can hardly police every Usenet and FTP site in the world. Nor is it reasonable to suppose that every country in the world is going to be interested in preventing 17-year-old CMU students from seeing pictures of naked bodies. It seems that the only solution is to block Internet access altogether, or to constantly monitor the activities of each student. Whether such draconian measures are feasible remains to be seen. If they are, one has to ask the question whether such measures do not do greater harm than potential encounters with images and ideas that some perceive to be dangerous.

20

Censoring Cyberspace

Philip Elmer-Dewitt

The steam began rising for Carnegie Mellon University four weeks ago, when one of its research associates, Martin Rimm, informed the administration that a draft of his study of pornography on the computer networks was about to be released. Rimm had made an elaborate analysis of the sexually oriented material available online. Not only had he put together a picture collection that rivaled Bob Guccione's (917,410 in all), but by tracking how many times each image had been retrieved by computer users (a total of 6.4 million downloads), he had obtained a measure of the consumer demand for different categories of sexual content, some of them, as a faculty adviser put it, "extremely rough." [The Rimm study has since been discredited by its numerous flaws in research methodology. Visit http://www2000.ogsm.vanderbilt.edu/cyberporn.debate.cgi for details.—PL]

The problem, from the Pittsburgh, Pennsylvania, university's point of view, was not that Rimm had found sexually explicit content on the computer networks; there is sex in every medium, from comic books to videotapes. Nor was it even that he had found some of it on CMU's own computers; every university connected to the Internet is a conduit, however unwitting, for gigabytes of salacious words and pictures. The immediate issue was that Rimm had brought it to the administration's attention, pointing out that some of the images on CMU's machines—digitized pictures of men and women having sex with animals, for example—had been declared obscene by a Tennessee court a few months before.

William Arms, vice president of CMU's computing services department, spent an hour reviewing the questionable material "with the law

of Pennsylvania in one hand and a mouse in the other" and decided that the university was in deep trouble. It is illegal in the state to knowingly distribute sexually explicit material to anyone under the age of 18—as many freshmen are—or to distribute obscene material at all, no matter what the consumer's age. Fearing that the university would be open to prosecution—and the worst kind of publicity—CMU's academic council hurriedly voted to shut down those areas of the computer system that carried discussions or depictions of sex. The plug was scheduled to be pulled last Tuesday.

Thus the lines were drawn for a battle over the preservation of free speech in the new interactive media—a battle that not only raised tricky questions about how to balance openness with good taste, but also managed, on a campus not noted for activism, to rouse something resembling a student protest movement. CMU casts a long shadow in cyberspace. It was one of the first universities to join the Arpanet (the precursor to the Internet) and the first to wire up its dorms. It even provides Internet access to some of its bathrooms. Using the computer networks to spread the word and muster support, the students quickly organized a "Protest for Freedom in Cyberspace" that drew 350 students and faculty members. (Pittsburgh in the 1990s, though, is hardly Berkeley in the '60s: the protesters last week politely applauded their opponents and then retired to a reception with cheese and fruit.)

At the core of the CMU dispute is a question that goes beyond the campus and could touch every media and entertainment company that wants to do business on the info highway: to what extent can the operators of interactive media be held responsible for the material that moves through their systems? Are they common carriers, like the phone companies, which must ignore the content of the messages? Are they like TV stations, whose broadcasts are monitored by the government for fairness and suitability? Or are they like bookstores, which the courts have ruled can't be expected to review the content of every title on their shelves? And what happens when that content hops over borders and lands in a different city—or country—whose laws and community standards may differ?

The last issue came to a head most dramatically last July, after a U.S. postal inspector, posing as a customer in Tennessee, downloaded X-rated pictures from an adult computer bulletin board in California. Though the

images might have been acceptable by California standards, they were judged obscene in the Bible Belt, and the owners of the bulletin board were convicted of transporting obscene material across state lines. Their appeal may be headed for the Supreme Court.

There's more to free speech than sexy words and pictures, of course. Publishers who venture onto international networks like the Internet are particularly concerned about libel and slander. The rules of libel in England, for example, are considerably more restrictive than those in the U.S.; what might be considered a fair crack at a public figure in New York City could be actionable in London. Conversely, the muzzles that are slapped on reporters covering trials in Commonwealth countries can't be placed so easily on writers living abroad, as Canadian officials learned to their dismay last year when foreign press reports of a particularly sensitive homicide case in Ontario began drifting back into Canada through the Internet.

All sorts of subversive materials have found their way onto the computer networks, from secret spy codes to instructions for making long-range rocket bombs. As if to provoke the authorities, some college students have posted collections of electronic pamphlets that include Suicide Methods, an instruction manual for self-destruction, and The School Stopper's Textbook, which tells students how to blow up toilets, short-circuit electrical wiring and "break into your school at night and burn it down."

High schools pose a special problem for administrators, who want to give students the benefits of computer networking without exposing minors to everything that washes up online. Many lower schools have adopted the CMU approach, cutting off access to the electronic discussion groups where the most offensive material is carried.

At CMU, the administration determined that its problem was centered in a collection of discussion groups, called Usenet newsgroups, with awkward but functional titles like alt.sex, rec.arts.erotica and alt.binaries.pictures.erotica. The "binary" groups are the most controversial, for they contain codes that savvy computer users can translate into pictures and movie clips. The university's initial decision was to pull the plug on all the major "sex" newsgroups and their subsidiary sections—more than 80 categories altogether.

That decision drew fire from all sides. The student council pointed out that the administration was restricting the reading matter of adults to

what was acceptable for children. The American Civil Liberties Union complained that the ban was overly broad and included discussions of sexual matters that were clearly protected speech. Mike Godwin, staff counsel for the Electronic Frontier Foundation, made a distinction between words and pictures, arguing that while images are still sometimes found obscene, words never are—a view confirmed by the Allegheny county assistant district attorney, who told *Time* there was "not a chance in a million" his office could win an obscenity case based on a written work.

But the central objection was more fundamental: that the university had ignored decades of constitutional law and abrogated its responsibility as a center for free inquiry. "I'm deeply ashamed that Carnegie Mellon capitulated so spinelessly," said one CMU student in a radio call-in debate. "Some lawyer told them they might someday be dragged into court, and they just decided, 'To hell with the First Amendment.'"

By midweek, the university had begun to back down. First it seized on Godwin's formula, banning the binaries and leaving the text in place—pending review by a student-faculty committee. Then, on Thursday night, the faculty senate voted to recommend restoration of all the newsgroups, including the binaries.

But the issue will not go away. There is material on the networks—child pornography, in particular—that has been targeted for prosecution by U.S. Attorney General Janet Reno. Unless computer users exercise some self-restraint, control could be imposed from the outside. If that happens, the next generation of interactive media may not have the freedom and openness that today's users value so highly.

Reported by John F. Dickerson/New York and Douglas Root/Pittsburgh

21

ACLU Letter to CMU on alt.sex Newsgroups

ACLU Urges Carnegie Mellon to Reverse Internet Censorship; Letter to University President Says Students Must Have Access to Information

For IMMEDIATE RELEASE

November 8, 1994

In a strongly worded letter to the President of Carnegie Mellon University, the American Civil Liberties Union today urged the university to reconsider and reverse its decision to prohibit student access to six network news groups that deal with sexual topics on the Internet.

"Carnegie Mellon has a well deserved reputation in higher education as a leader on technology issues," said Barry Steinhardt, Associate Director of the ACLU. "You have already recognized the extraordinary potential of networked communications to enhance and democratize speech.

"But if the full potential is to be reached, it is important that leaders like Carnegie Mellon stand strong for free and open access to information and that you resist the urge to censor," Steinhardt concluded.

In its letter, the ACLU said that Carnegie Mellon officials based their decision to remove the news groups on a broad misreading of Pennsylvania obscenity law. The vast majority of information on these news groups has never been challenged as obscene, the ACLU said, nor could the university be held liable for distributing this material through the Internet.

"Your policy sweeps far too broadly," the letter said. "Out of fear that your students may be exposed to a few unprotected works, you have cut off access to a large volume of protected ideas and information."

A copy of the letter to Carnegie Mellon is attached.

November 8, 1994

Dr. Robert Mehrabian
President
Carnegie Mellon University
5000 Forbes Ave. Warner Hall
Pittsburgh, PA 15213

Dear President Mehrabian:

We write on behalf of the American Civil Liberties Union (ACLU) to urge Carnegie Mellon University to reconsider and reverse the decision to prohibit student access to six network news groups which deal with sexual topics. We believe that the University's plan is inconsistent with the principles of academic freedom and free speech, which a great University must defend, and is based on a serious misreading of relevant laws.

Carnegie Mellon's decision to offer its students broad access to the Internet and its thousands of news groups was a farsighted recognition that networked communications will increasingly provide the means for academic research and the forum for the free exchange of ideas. Like your decision to establish and nurture a library, it was a decision to give your students access to the widest variety of information and to allow them to make their own judgements about their worth.

Your decision to revoke access to the six news groups deprives your students of the opportunity to judge for themselves the value of these groups.

As we understand the rationale for your decision, the University is concerned about its potential liability under Pennsylvania's obscenity law and particularly its provisions relating to minors. We believe that the conclusions you have drawn are mistaken and misperceive your role in providing student access to the Internet. This also seems surprising since there has apparently been no effort by any governmental entity to assert the existence of any such liability.

To begin with, these news groups are not "obscene" merely because they contain sexually explicit materials. Most sexually explicit speech is protected by the First Amendment and only a small portion of sexually explicit materials can constitutionally be deemed obscene. A jury applying the three-part test of Miller v. California must decide on a case by

case basis whether a particular work is obscene. There are literally thousands of postings to these news groups every year and only a small fraction have ever been challenged as obscene.

Your policy sweeps far too broadly. Out of fear that your students may be exposed to a few works that a court might ultimately find unprotected, you have cut off access to a large volume of protected ideas and information.

In fact, you have cut off access to information that is clearly of significant value to your students and deals with serious societal issues. For example, in barring access to the alt.sex news group, you have deprived students of access to all of its branches, including alt.sex.safe, which discusses responsible sexual behavior.

The University's conclusion that you must cut off access to these news groups because "it is a criminal offense to knowingly disseminate sexually explicit material to minors . . ." is equally troubling.

First, it is not illegal to distribute any "sexually explicit" material to minors. The state can only ban the distribution to minors of materials that satisfies the classic three part Miller test, as modified for minors. A careful reading of Pennsylvania law makes this clear. 18 Pa. C.S.A. 5903 applies only to those types of sexually explicit materials which satisfy the modified test and are "harmful to minors."

Secondly, even assuming that some of the material posted to these news groups might properly be restricted to adults, it is a well established principle that obscenity policies cannot reduce adults to reading only that which is fit for children. But that is precisely the effect of the new policy.

The vast majority of your students are undoubtedly over the age of 18 and legally adults. The policy effectively treats them as children and limits their access to materials deemed suitable for children.

That is not what the law requires and we do not think that you would draw this conclusion in other contexts. Surely, you don't believe that Pennsylvania law requires that the University library and bookstore, or for that matter your literature classes, must be purged of all sexually explicit material because there may be minors on campus or minors may access the imagery.

Finally, by its terms, the Pennsylvania obscenity law cannot be applied to the University's provision of Internet access; nor should it be. The

Pennsylvania statute provides that a party can only be held liable for the content of material . . . which is reasonably susceptible of examination by the defendant." (Sec 5903 (b))

By its very nature, the Internet is vast and chaotic. There are millions of speakers and countless speeches. The Internet is so large and disorderly that the University could never reasonably be expected to examine the content of every message available to your students.

Nor could you reasonably be expected to be able to control all access to any particular group of messages. One of the unique features of the Internet is that there are many paths to the same destination.

CMU's students are smart and technologically savvy. They will quickly learn how to use the access provided by the University to subscribe to Internet "mailing lists," to download archival files or to link with other networks to obtain exactly the same material that is available in the banned news groups.

The new policy assumes that the University has an obligation to prevent its students from obtaining access to any possibly illegal materials about which the University has knowledge. If that assumption is correct, then you have not gone far enough in simply blocking access to a few news groups. To meet such a heavy burden you would need to take the draconian steps of either monitoring all student communications or cutting off all access to the Internet.

Fortunately, the law does not impose that burden. The law properly holds speakers and publishers liable for the content of their messages. By offering your students access to thousands of news groups and hundreds of millions of postings, you are neither a speaker, nor a publisher. At most, you are acting as a distributor providing access to information, exactly in the same way that you provide access to library books, and could not and should not be held liable, for example, if a 17-year-old freshman happens to check out a book with "sexually explicit" content.

In fact, Subsection (j) of the Pennsylvania obscenity law explicitly exempts "any library of any school, college or university . . ." from its reach. The Legislature recognized that universities and libraries have special protections as access providers to knowledge and that you should not be chilled in your mission by the specter of criminal or civil prosecution of an obscenity law that you cannot be reasonably expected to enforce.

While your connection to the Internet may not be housed in your library building, it is no less deserving of the protection offered by Subsection (j). By providing wide access to the Internet, you are, in effect, functioning as electronic librarians and the exemption should apply. Furthermore, as a well established principle of constitutional due process, any doubts about applicability of Pennsylvania obscenity laws must be resolved in your favor.

Indeed, we would think and hope that the University would want to lay vigorous claim to this exemption in order to protect future concepts of academic freedom. A library free from government control is an essential component of a vibrant university. As technology changes the ways in which we store and access information, it seems beyond dispute that the digital library of the next century will bear far greater resemblance to the Internet than to today's brick and mortar constructs. As the technology changes, it is essential that we not lose sight of core principles of academic freedom.

Carnegie Mellon has a well-deserved reputation in higher education as a leader on technology issues. You have already recognized the extraordinary potential of networked communications to enhance and democratize speech. But if the full potential is to be reached, it is important that leaders like CMU stand strong for free and open access to information and that you resist the urge to censor.

We strongly urge you to reconsider the new policy. It is our understanding that the University's decision was made without consultation with your counsel. We hope that you will now take that step. We would be happy to provide your attorneys with additional citations and materials to facilitate a reconsideration.

Thank you for your consideration of our views.

Sincerely,

Barry Steinhardt
Associate Director, ACLU

Marjorie Heins
Director, ACLU Arts Censorship Project
Witold Walczak
Executive Director, ACLU Greater Pittsburgh Chapter

22

Virtual Community Standards: BBS Obscenity Case Raises New Legal Issues

Mike Godwin

At first glance, the obscenity prosecution of Robert and Carleen Thomas of Milpitas seemed little different from the average obscenity prosecution. Sure, this case involves a computer bulletin board system (BBS), but there's nothing new about prosecuting pornography distributors in conservative states like Tennessee, is there?

Except that this BBS wasn't in Tennessee. It was in California. But that didn't stop Tennessee prosecutors from going after it. Because of the way BBSs normally operate, a conservative jurisdiction like Memphis may be in a position to dictate what's allowable on BBSs all over the country, from New York City to San Francisco. For this reason, the prosecution of the Thomases and their "Amateur Action BBS" calls into question the continuing validity of the Supreme Court's obscenity decision in Miller v. California, now more than 20 years old. That case, which was designed to allow communities to set their own standards of what is acceptable and what is obscene, has now been used for just the opposite purpose—it has allowed a Memphis prosecutor to dictate the content of a computer system in California.

Memphis Reaches Out to Touch Someone

The facts of the case are straightforward. The Thomases are the system operators (sysops) of an adults-only sexually oriented BBS in Milpitas, California. The operator of a BBS typically dedicates a computer and one or more phone lines at his home or business for the use of a "virtual community" of users. Each user calls up the BBS (using a modem con-

nected to his or her telephone) and leaves public messages that can be read by all other users and/or private mail that can be read by a particular user. BBSs become forums—digital nightclubs, salons, and Hyde Park corners—for their users, and users with similar interests can associate with one another without being hindered by the accidents of geography. A BBS also can be used to trade in computer files, programs, and digital images, including sexually graphic images.

A Tennessee postal inspector, working closely with an assistant U.S. attorney in Memphis, became a member of the Thomases' BBS. Once he had become a member, he did three things: he downloaded sexually oriented images, ordered a videotape (which was delivered via UPS), and sent an unsolicited child-porn video to the Thomases. This led to a federal indictment with a dozen obscenity counts, most based on the downloading of the computer images.

The indictment also included one child-pornography count, based on the unsolicited video. At trial, the Memphis jury convicted the Thomases on all the obscenity counts, but acquitted them on the child-porn count. (A reporter at the scene who interviewed jurors said they believed the child-porn count smacked of entrapment.) The Thomases now face sentencing on the 11 obscenity convictions, each carrying a maximum sentence of five years in prison and $250,000 in fines.

The Thomases' lawyer says they will appeal, based at least in part on a claim that the jury instructions as to "community standards" were incorrect. "This case would never have gone to trial in California," he has said.

Community Standards and BBSs

It has long been held that obscenity is not protected by the First Amendment, but what qualifies as "obscenity" has not always been clear. After Miller v. California, a 1973 Supreme Court case, there has been no national standard as to what is obscene. In that case, the Court stated that material is "obscene" (and therefore not protected by the First Amendment) if 1) the average person, applying contemporary community standards, would find the materials, taken as a whole, arouse immoral lustful desire (or, in the Court's language, appeals to the "prurient interest"), 2) the materials depict or describe, in a patently offensive way, sexual conduct specifically prohibited by applicable state law, and 3) the work, taken as a whole, lacks serious literary, artistic, political or scientific value.

To put it in layman's terms, the trial court would ask something like these four questions:

1. Is it designed to be sexually arousing?
2. Is it arousing in a way that one's local community would consider unhealthy or immoral?
3. Does it depict acts whose depictions are specifically prohibited by state law?
4. Does the work, when taken as a whole, lack significant literary, artistic, scientific, or social value?

If the answer to all four questions is "yes," the material will be judged obscene, and it will be constitutional to prosecute someone for distributing it. (It should be noted in passing that pictures of the "hardness" of *Playboy* and *Penthouse* photography are never found to be obscene—their appearance in digital form on Usenet sites may create copyright problems, but they won't create obscenity problems. Remember also that "pornography" and "obscenity" are not identical categories—much pornography is not legally obscene.)

Normally, an appeal on the issue of obscenity will focus on one or more of the answers to the four questions. If, for example, a Robert Mapplethorpe photo is found obscene at a trial, defense on appeal might argue that, even if the photo is sexually arousing in a way that violates community standards and state law, the work's social value renders it protected by the First Amendment. In hardcore porn cases, the defense might argue that, in fact, the community is highly tolerant of such images (in adult bookstores, films, and the like).

It has long been held to be constitutional to prosecute any porn vendors located in more liberal jurisdictions who have knowingly or intentionally distributed obscenity into conservative jurisdictions. Many large-scale commercial porn vendors have made deliberate decisions not to distribute their materials into jurisdictions likely to prosecute—postal inspectors frequently engage in "sting" operations in order to test whether a vendor will send obscene material into their states.

This case is different, however. Consider: a seller of adult magazines normally makes a conscious decision to send his product into the jurisdiction in which he's prosecuted, thus establishing criminal intent for the purpose of an obscenity-distribution prosecution. In contrast, a BBS operator may be wholly unaware of the distribution—it may occur overnight, for example—due to the automatic operation of his software.

What's more, even if the Thomases were to attempt to screen their users on a state-by-state basis, there's no guarantee that this attempt would protect them—a user could simply lie about which state he is calling from, or he could obtain a membership while living in California yet maintain it after he moved to Tennessee. Since a BBS operator cannot block out calls from conservative jurisdictions, there is inherent vulnerability for a BBS operator that exceeds that for traditional pornography distributors.

While the Thomases' conviction with regard to the UPS-delivered video is likely to stand on traditional grounds, their convictions with regard to the downloaded images raise a number of critical issues. For example, does it make sense for a court to infer a defendant's criminal intent to distribute obscenity into Tennessee merely because neither he nor his BBS can ensure that someone cannot download that material into the state?

More importantly, the case turns the whole community-standards doctrine on its head. The Supreme Court was attempting, in Miller v. California, to prevent the standards of acceptability in New York City or San Francisco from dictating the standards of Kansas City or Norman, Oklahoma. Yet if it's wrong for New York City to set the standards for Norman, it's surely just as wrong for Memphis to set the standards for Milpitas.

Finally, the case raises the question of whether it makes sense to define "community standards" solely in terms of geographic communities. Now that an increasing number of Americans find themselves participating in "virtual communities" on services such as America Online, CompuServe, Prodigy, and the WELL, does it make sense to have what those citizens are allowed to bring into their own homes be dictated by the arbitrary fact of where their physical homes happen to be?

It's time for the courts to revisit the Miller obscenity standard. In the face of changes in communications media and the evolving nature of "community," the courts should modify the application of the Miller standard to prevent this kind of prosecutorial overreaching. Failing that, the courts should abandon the "community standards" approach altogether.

Until these issues are addressed, this case will create a "chilling effect" all over the country, as BBSs either censor themselves or cease operations in order to avoid prosecution. The case sends a frightening message to virtual communities: "It doesn't matter if you're abiding by your own community's standards—you have to abide by Memphis's as well."

24

Sex and the Single Sysadmin: The Risks of Carrying Graphic Sexual Materials

Mike Godwin

It's the kind of nightmare that will cause any sysadmin to bolt upright in bed, shaking, gripping the sheets with white-knuckled fingers.

In this nightmare scenario, the facts are simple: you hear a knock at the door, you answer to discover grim-faced law-enforcement agents holding a search warrant, and you are forced to stand by helplessly while they seize your system to search it for obscene or child-pornographic images.

In some versions of the nightmare, you may not even have known your hard disk contained such images; in others, your lack of knowledge may prove to be no defense in a criminal prosecution for possession of child pornography.

A Wave of Concern about Porn

In recent months, the Legal Services Department here at EFF (the Electronic Frontier Foundation) has faced a wave of concern in the United States about the legal issues raised by online obscenity and child pornography. Most recently, a nationwide federal investigation into the importing of child-pornographic computer files led first to several well-publicized searches and seizures of computers and bulletin-board systems (BBSs) and later to a number of indictments of computer users on charges relating to possession or distribution of this material. One result has been that a large number of BBS operators and network site administrators have contacted EFF with questions and concerns about their potential liability under obscenity and child-pornography laws.

Why so much concern? Partly, it's that, thanks to the availability of cheap image scanners, fast modems, and capacious hard disks, a large number of this country's BBSs and network sites carry GIF (Graphic Interchange Format) files or other kinds of graphic images with sexual content. These images can range from centerfold-type nudes to "hard-core" pornography. (For the sake of simplicity, I will refer to all graphic-image files as GIFs, although there are a number of other formats commonly available.)

Just as the growth of the consumer VCR market was linked to a growth in the market for adult videos, the increasing availability of certain kinds of consumer computer technology has led to a rapid increase in GIF-file traffic. System operators who might never consider opening an adult book or video store have either allowed or encouraged sexually oriented images to be exchanged on their systems. To understand this difference in attitudes one has to understand how online conferencing systems are generally run—as forums for their users to talk to each other, and to trade computer programs and files with each other.

How Porn Gets Online

Although these problems pervade the world of the Internet, the easiest case to understand is the microcomputer-based BBS. The operator of a BBS typically dedicates a computer and one or more phone lines at her home or business for the use of a "virtual community" of users. Each user calls up the BBS and leaves public messages (or, in many cases, GIFs) that can be read by all other users or private mail (which may include GIFs) that can be read by a particular user or both. BBSs become forums—digital public houses, salons, and Hyde Park corners—for their users, and users with similar interests can associate with one another without being hindered by the accidents of geography. By some estimates, there are currently in excess of 40,000 BBSs throughout North America, ranging from low-end free-access BBSs with only one or two phone lines to BBSs run by companies, government agencies, user groups, and other organizations.

A step up from the BBS in complexity is the conferencing system or information service. These systems differ in capacity from BBSs: they

have the capability of serving dozens, or hundreds, of users at the same time. But they're like BBSs in that uploaded files can be found at a fixed geographic location. A further step up are entities like Fidonet and Usenet, which, because they're highly distributed, decentralized conferencing systems, add complications to the legal issues raised by the computerization of sexual images.

Internet nodes and the systems that connect to them, for example, may carry such images unwittingly, either through uuencoded mail or through uninspected Usenet newsgroups. The store-and-forward nature of message distribution on these systems means that such traffic may exist on a system at some point in time even though it did not originate there, and even though it won't ultimately end up there. What's more, even if a sysadmin refuses to carry the distributed forums most likely to carry graphic images, she may discover that sexually graphic images have been distributed through a newsgroup that's not obviously sexually oriented.

Depending on the type of system he or she runs, a system operator may not know (and may not be able to know) much about the system's GIF-file traffic, especially if his or her system allows GIFs to be traded in private mail. Other operators may devote all or part of their systems to adult-oriented content, including image files.

Regardless of how their systems are run, though, operators often create risks for themselves under the mistaken assumption that a) since this kind of material is commonplace, it must be legal, and b) even if it's illegal, they can't be prosecuted for something they don't know about. EFF's Legal Services Department has been working actively to educate system operators about the risks of making these assumptions.

What Counts as "Obscene"?

First of all, we've explained that the fact that graphic sexual material is common on BBSs doesn't mean that it's not legally obscene and illegal in their jurisdiction.

As Judge Richard Posner comments in the October 18, 1993, issue of *The New Republic,* "Most "hard-core" pornography—approximately, the photographic depiction of actual sex acts or of an erect penis—is illegal," even though it is also widely available. (Let me emphasize the

word "approximately"—Posner knows that there are countless exceptions to this general rule.) That is, distribution of most of this material is prohibited under state or federal anti-obscenity law because it probably would meet the Supreme Court's test for defining obscenity.

But what precisely is the Court's definition of obscenity? In Miller v. California (1973), the Court stated that material is "obscene" (and therefore not protected by the First Amendment) if 1) the average person, applying contemporary community standards, would find the materials, taken as a whole, arouse immoral lustful desire (or, in the Court's language, appeals to the "prurient interest"), 2) the materials depict or describe, in a patently offensive way, sexual conduct specifically prohibited by applicable state law, and 3) the work, taken as a whole, lacks serious literary, artistic, political or scientific value.

This is a fairly complex test, but most laymen remember only the "community standards" part of it, which is why some system operators are under the mistaken impression that if the material is common and available, "community standards" and the law must allow it.

The Perils of Online Obscenity

In theory, most "hardcore" pornography qualifies as "obscenity" under the Supreme Court's test. Yet theoretically obscene material is commonly available in many urban areas—this signifies, perhaps, that the relevant laws, when they do exist, are underenforced. At EFF, however, we have been telling system operators that there is no *legal* basis for their assuming that the laws will remain underenforced when it comes to online forums.

For one thing, most of this country's law-enforcement organizations have only recently become aware of the extent that such material is traded and distributed online—now that they're aware of it, they're aware of the potential for prosecution. In a recent case, an Oklahoma system operator was charged under state law for distribution of obscene materials, based on a CD-ROM of sexual images that he'd purchased through a mainstream BBS trade magazine. He was startled to find out that something he'd purchased through normal commercial channels had the potential of leading to serious criminal liability.

Still another issue, closely related to obscenity law, is whether an online system creates a risk that children will have access to adult materials. States in general have a special interest in the welfare of children, and they may choose to prohibit the exposure of children to adult materials, even when such materials are not legally obscene. (Such materials are often termed "indecent"—that is, they violate some standard of "decency," but nevertheless are constitutionally protected. If this category seems vague, that's because it is.) In Ginsberg v. State of New York (1968), the Supreme Court held a state statute of this sort to be constitutional.

Although there is no general standard of care for system operators who want to prevent children from having such access, it seems clear that, for a system in a state with such a statute, an operator must make a serious effort to bar minors from access to online adult materials. (A common measure—soliciting a photocopy of a driver's license—is inadequate in my opinion. There's no reason to think a child would be unable to send in a photocopy of a parent's driver's license.)

It's worth noting that, in addition to the risk, there are also some protections for system operators who are concerned about obscene materials. For example, the system operator who merely possesses, but does not distribute, obscene materials cannot constitutionally be prosecuted—in the 1969 case Stanley v. Georgia, the Supreme Court held the right to possess such materials in one's own home is constitutionally protected. Thus, even if you had obscene materials on the Internet node you run out of your house, you're on safe ground so long as they're not accessible by outsiders who log into your system.

And, in the 1959 case Smith v. California, the Court held that criminal obscenity statutes, like the great majority of all criminal laws, must require the government to prove "scienter" (essentially, "guilty knowledge" on the defendant's part) before that defendant can be found guilty. So, if the government can't prove beyond a reasonable doubt that a system operator knew or should have known about the obscene material on the system, the operator cannot be held liable for an obscenity crime.

In short, you can't constitutionally be convicted merely for possessing obscene material, or for distributing obscene material you didn't know about.

Child Pornography—Visual Images That Use Children

When the issue is child pornography, however, the rules change. Here's one of the federal child-porn statutes:

18 USC 2252: Certain activities relating to material involving the sexual exploitation of minors.

(a) Any person who—

(1) knowingly transports or ships in interstate or foreign commerce by any means including by computer or mails, any visual depiction, if—

(A) the producing of such visual depiction involves the use of a minor engaging in sexually explicit conduct; and

(B) such visual depiction is of such conduct; or

(2) knowingly receives, or distributes, any visual depiction that has been transported or shipped in interstate or foreign commerce by any means including by computer or mailed or knowingly reproduces any visual depiction for distribution in interstate or foreign commerce by any means including by computer or through the mails if—

(A) the producing of such visual depiction involves the use of a minor engaging in sexually explicit conduct; and

(B) such visual depiction is of such conduct;

shall be punished as provided in subsection (b) of this section.

(b) Any individual who violates this section shall be fined not more than $100,000, or imprisoned not more than 10 years, or both, but, if such individual has a prior conviction under this section, such individual shall be fined not more than $200,000, or imprisoned not less than five years nor more than 15 years, or both. Any organization which violates this section shall be fined not more than $250,000.

(N.B. For the purposes of federal law, "minor" means "under age 18"—it does not refer to the age of consent in a particular state.)

This statute illustrates some of the differences between the world of obscenity law and that of child-pornography law. For one thing, the statute does not address the issue of whether the material in question is "obscene." There's no issue of community standards or of "serious" artistic value. For all practical purposes, the law of child pornography is wholly separate from the law of obscenity.

Here's the reason for the separation: "obscenity" laws are aimed at forbidden expression—they assume that some things are socially harmful by virtue of being expressed or depicted. Child-porn laws, in contrast, are not aimed at *expression* at all—instead, they're designed to promote the protection of children by trying to destroy a market for materials the production of which requires the sexual use of children.

This rationale for the child-pornography laws has a number of legal consequences. First of all, under the federal statute, material that depicts child sex, but in which a child has not been used, does not qualify as child pornography. Such material would include all textual depictions of such activity, from Nabokov's novel *Lolita* to the rankest, most offensive newsgroups on Usenet, all of which are protected by the First Amendment (assuming that, in addition to not being child pornography, they're also not obscene).

Secondly, the federal child-porn statute is limited to visual depictions (this is not true for all state statutes), but does not apply to *all* visual depictions: computer-generated or -altered material that *appears* to be child pornography, but which did not in fact involve the sexual use of a real child, would not be punishable under the federal statute cited above. This makes sense in light of the policy—if real children aren't being sexually abused, the conduct these statutes are trying to prevent has not occurred. Although prosecutors have had little trouble up to now in proving at trial that actual children have been used to create the child-porn GIF images at issue, we can anticipate that, as computer-graphics tools grow increasingly powerful, a defendant will someday argue that a particular image was created by computer rather than scanned from a child-porn photograph.

Third, since the laws are aimed at destroying the market for child pornography, and since the state has a very powerful interest in the safety of children, even the mere possession of child porn can be punished. (Compare: mere possession of obscene materials is constitutionally protected.)

The fourth consequence of the child-protection policy that underlies child-porn statutes is that the federal law, as interpreted by most federal courts, does not require that the defendant be proved to have known that a

"model" is a minor. In most jurisdictions, a defendant can be convicted for possession of child porn even if he can prove that he believed the model was an adult. If you can prove that you did not even know you possessed the image at all, you should be safe. If your knowledge falls somewhere in between—you knew you had the image, but did not know what it depicted, or that it was sexual in its content—the law is less clear. (In other words, it's not yet clear whether it is a defense for a system administrator to claim he didn't even know he possessed the image, either because it had been uploaded by a user without his knowledge, or because it had appeared in "pass-through" mail or through a Usenet newsfeed.)

In sum, then, the child-porn statutes create additional problems for the system administrator who wants to avoid criminal liability and minimize the risk of a disruptive search and seizure.

What You Can Do

The first thing to do is not to overreact at this discussion of the risks. It would amount to a serious "chilling effect" on freedom of expression if a sysadmin—in order to eliminate the risk of prosecution for distribution of obscenity, or for possession or distribution of child pornography—decided to eliminate all newsgroups with sexual content. The textual content of such newsgroups is constitutionally protected, as is much of the GIF content.

What's worse is that the tactic wouldn't eliminate the risks—it's always possible for someone to post illegal material to an innocuous newsgroup, like sci.astro or rec.arts.books, so that it would get to your system anyway. Similarly, an illegal image might be uuencoded and included in e-mail, which, if you're a system covered by the Electronic Communications Privacy Act, you're not allowed to read.

You should begin with the knowledge that nothing you can do as a sysadmin will eliminate altogether the risks of prosecution or of a disruptive search and seizure. But a few sensible measures can reduce the risks of a search or an arrest, and at the same time preserve the freedom of expression of your users and of those users who transmit material through your system.

- If you plan to carry graphic sexual material, look up your state's obscenity laws. A lawyer or librarian can help you find the relevant state statutes. Find out what, specifically, your state tries to prohibit. (If the state statute seems inconsistent with what I've written here, consider seeking legal advice—it may be that the statute predates the Supreme Court's decisions on obscenity and child pornography but has not yet been challenged.) You may also want to consult local adult bookstores—they often have clear, practical information about avoiding obscenity prosecutions.
- If you're running an online forum local to your system, and that forum has an upload/download area, prescreen graphic images before making them publicly available for downloads. While "calendar" and "foldout" images are constitutionally protected, you may want to consider deleting "hardcore" images that might be found "obscene" in your community. You also want to delete anything that looks like child pornography.
- If you're running a Usenet node, and you are informed by users that an obscene or child-porn image has been posted to a newsgroup you carry, examine it and consider deleting it. If there's any ambiguity, err on the conservative side—remember, if you guess wrong about the age of the model, you can be convicted anyway.
- Take pains on your system to limit childrens' access to adult material, even if that material is not legally obscene (it may still be "indecent"). This includes textual material dealing with adult topics. Hint: asking for a photocopied driver's license in the mail is probably not an adequate safeguard—too easy for industrious minors to circumvent. A good set of rules to follow is spelled out in an FCC regulation applicable to phone-sex providers—47 CFR 64.201. The easiest FCC suggestions for a for-pay BBS, online service, or Internet access provider is to require payment by credit card; the easiest for a nonpay system is have an application process that reasonably ascertains whether an applicant for access is an adult, and to have a procedure whereby one can instantly cut off that access when informed that a user is in fact a minor.
- Don't delete discussions of sexual topics—they're constitutionally protected. And even though the Supreme Court has not limited the definition of "obscenity" to visual depictions, as a practical matter, there is little legal risk in carrying textual narratives ("stories") on sexual themes.
- Don't inspect individuals' e-mail without their consent—unless they're employees of your company, their mail is probably protected by the Electronic Communications Privacy Act.

• If you're a university site, or if you're simply interested in the law of freedom of speech, consult the Computers and Academic Freedom (CAF) archive, which is part of the EFF archive at ftp.eff.org. If you have gopher, the archive is at gopher.eff.org; if you are limited to e-mail access, send e-mail to archive-server@eff.org, and include the line

send acad-freedom/law

The CAF archive has a number of instructional materials that deal with obscenity and child-pornography law.

These measures won't guarantee that you'll never have legal troubles—nothing can guarantee that. (And if you have particular legal worries, you should consult a lawyer in your jurisdiction.) But they can reduce the risks you face as a system administrator and as a carrier and distributor of information. At the same time, they'll minimize the extent to which you interfere with your users' freedom to communicate—which is, after all, one of the chief reasons they're online in the first place.

25

Computer and Academic Freedom News's List of Banned Computer Material on College Campuses

Inspired by Banned Book Week '92, this is a list of computer material that was banned or challenged in academia in 1992. Iowa State University has the dubious distinction of being listed most often (three times).

The list proper starts after a list of the academic institutions where bans or challenges have occurred. The list proper is followed by instructions on how to get more information about specific incidents and then by instructions on how to get general information about computers and academic freedom.

Please send reports, corrections, and updates to either caf-talk@eff.org (a public mailing list) or kadie@eff.org (private).

Carl Kadie, kadie@eff.org,
co-editor, *Computer and Academic Freedom News*
Disclaimer: I do not represent EFF; this is just me.
version: 1.09

Academic Institutions

USA

Ball State University
Boston University (i)
Carnegie Mellon University
Iowa State University (i)
North Dakota State University
Princeton

23

Public Networks and Censorship

Jeffrey Shallit

Good afternoon. Thank you very much for the opportunity to speak to the Ontario Library Association on the subject of public networks and censorship.

1 Librarians and Computers

I had planned to start off with something sententious such as, "We stand today at an information delivery crossroads," but the truth is, that we have already passed this crossroads and are heading into the information age at very high speed. The crossroads, I think, was traversed back in 1989—when, for the first time, the number of videotapes rented exceeded the number of books checked out of public libraries.

The old concept of the library, as we have known and loved it, is dying. Now I'm not saying that books will cease to be published, or that traditional library concerns such as shelf space and book theft will disappear tomorrow. But I *am* saying that there is an enormous flood of information and communication that is about to be unleashed, that is already being unleashed, and that librarians and the principles they have developed and fought hard for, are desperately needed in the new world as "information intermediaries."

The librarians of yesterday were valued by the general public for, among other things, their abilities to determine just *where* in that intimidating building full of books, magazines, newspapers, and scholarly journals the particular piece of desired information resided. The librarians of tomorrow will be equally valued, but now much of the information lies

in cyberspace. Yesterday, the *Reader's Guide to Periodical Literature* and *Ulrich's;* today, the Lexis/Nexis search service; tomorrow . . . ?

The librarians of yesterday were also known as guardians of intellectual freedom and the freedom to read. The principles of their profession can be found in statements produced by such groups as the American Library Association (for example, Library Bill of Rights, the Freedom to Read Statement, and the Intellectual Freedom Statement); the Canadian Library Association (Statement on Intellectual Freedom); and the Canadian Association of Research Libraries.

Let's take a look at just one of those statements, the Intellectual Freedom Statement of the Canadian Library Association [15]:

> All persons in Canada have the fundamental right, as embodied in the nation's Bill of Rights and the Canadian Charter of Rights and Freedoms, to have access to all expressions of knowledge, creativity and intellectual activity, and to express their thoughts publicly. . . .
>
> Libraries have a basic responsibility for the development and maintenance of intellectual freedom.
>
> It is the responsibility of libraries to guarantee and facilitate access to all expressions of knowledge and intellectual activity, including those which some elements of society may consider to be unconventional, unpopular or unacceptable. To this end, libraries shall acquire and make available the widest variety of materials. . . .
>
> Libraries should resist all efforts to limit the exercise of these responsibilities while recognizing the right of criticism by individuals and groups. . . .

I find those words very inspiring, and I hope you do, too. The question I would like to pose to you today is: as the libraries of yesterday are transformed into the libraries of tomorrow, will these principles govern electronic communication technologies such as the Internet?

2 Shallit's Three Laws

Before we begin discussion of fundamental freedoms on computer networks and the challenges to those freedoms, I'd like to tell you about what I modestly call Shallit's Three Laws of New Media. Shallit's first Law is the following:

Every new medium of expression will be used for sex.

Now you might say that I'm overstating my case, but think about it for a moment: some of the very earliest sculptures we know about are fertility symbols, such as the Venus of Laussel (c. 20,000 BC). One of the earliest books printed after Gutenberg invented the printing press was Bocaccio's erotic classic, *The Decameron.* Shortly after the introduction of photography, there was a thriving trade in pornographic pictures. And some anthropologists have even claimed that speech evolved so quickly in humans because it facilitated seduction! And this brings me to Shallit's Second Law:

Every new medium of expression will come under attack, usually because of Shallit's First Law.

Before I get to Shallit's Third Law of New Media, I'd like to tell you a story from a really terrific book, Carolyn Marvin's *When Old Technologies Were New: Thinking About Electric Communication in the Late Nineteenth Century.* Marvin's book is largely concerned with the societal impact of the telegraph and telephone, and, as we will see, neither was exempt from Shallit's Three Laws.

As Marvin observes

> New forms of communication created unprecedented opportunities not only for courting and infidelity, but for romancing unacceptable persons outside one's own class, and even one's own race, in circumstances that went unobserved by the regular community. The potential for illicit sexual behaviour had obvious and disquieting power to undermine accustomed centers of moral authority and social order. [7, p. 70]

Now here's that story I promised: in the summer of 1886, in New Jersey,

> a "nice young man" from the city met "one of the rustic beauties of the place" and they fell in love. They corresponded, and she invited him to visit. One day a telegram appeared with news of his impending arrival.
>
> Somehow—nobody ever will know just how—fifteen minutes after the message clicked into the [telegraph] office every person in town knew that young Blake was coming to see Miss Trevette. Every young lady of the town made up her mind to catch a glimpse of this rash young man who sent telegrams, and every man determined to be there to see that everything went smoothly.
>
> When young Blake alighted from his carriage . . . an audience of 499 villagers had gathered to watch. They observed while he paid the driver, studied him as he asked directions to the young lady's house, and followed his progress up the hill.

Panicked by the approaching procession, Miss Trevette sent word of her absence, halting the romance at a blow. [7, pp. 70–71]

An amusing story—but with a cautionary moral. Today's new communications technology—electronic mail—does not yet enjoy any of the legal or societal protections we associate with communication by more traditional means. While employers would think twice before opening an employee's mail delivered by Canada Post, e-mail is another matter. For example, Nissan Corporation dismissed a man for "inappropriate jokes and language" found in his e-mail. Epson, a computer company, dismissed a woman after she reported on a co-worker who was reading another employee's e-mail—apparently with the blessing of management. [8]

And this brings up Shallit's Third Law of New Media:

Protection afforded for democratic rights and freedoms in traditional media will rarely be understood to apply to new media.

Shallit's Third Law can be rephrased as the fallacy of focusing on the medium and not the message. A good illustration is the regulation of radio and television broadcasting. We tolerate content restrictions on television, for example, that would be intolerable if they were applied to print media [9]. When asked why, most people cite the supposed scarcity of the airwaves as a justification for government regulation of content. The truth is that this scarcity itself is a product of government intervention. The technology now exists to make possible hundreds or even thousands of broadcast stations in any metropolitan area. You don't hear much about this, because broadcasters are understandably less than enthusiastic about new competition, and the CRTC doesn't wish to relinquish its control on content. As Jonathan Emord shows, in his excellent book *Freedom, Technology, and The First Amendment*, regulations on broadcasting were historically enacted with little understanding of the technology and its capabilities [3].

3 Threats and Challenges to Freedom

We see that traditional democratic freedoms, such as freedom of expression and privacy, are under threat when these freedoms are asserted electronically.

And make no mistake, there is indeed a threat. One danger is that the new medium will be regulated to death before it is firmly established. For example, in November 1994, Reform MP Myron Thompson issued a press release alleging "highly pornographic, illegal stories available on Internet . . . that are reaching our children" and saying, "this smut must be stopped." (Shallit's Second Law again!)

Also, in a report recently presented to the Canadian Parliament, the Justice Committee recommended changes to the legal definition of "obscenity" to include "undue exploitation or glorification of horror, cruelty, or violence." In addition to cards and games, the report names "music, videos, comics, posters, and computer bulletin boards" as forms of communication that need to be controlled by the government. Communication that falls within this expanded definition and has "no redeeming cultural or social value" would be prohibited. The Internet is at risk, but books are safe . . . at least for the time being.

One reason for this difference in legal protection is that the print medium has existed for more than five hundred years, and libraries have existed for thousands of years. During that time, librarians have earned a good reputation for their craft, and have developed intellectual freedom principles that are well-respected. In contrast, electronic computers have existed for barely fifty years, and computer networks for barely twenty years. Computer system administrators have their own conferences and their own journals, but to my knowledge, they have no statement of duties, responsibilities, or ethics even remotely like the ALA's *Intellectual Freedom Manual* [2].

Within the next ten years, I predict that the power of many computer system administrators to regulate content on the machines they administer will wane. They will still be needed to help plan day-to-day use, install new software, and fix bugs, but the responsibility for such public forums such as Usenet news, etc., will move to people trained in principles of acquisition and intellectual freedom.

It may be that in the near future, the sheer volume of information flow will make selection much more necessary than it is today. When this happens, shouldn't the decisions on what electronic materials to subscribe to be based on the acquisition principles that librarians have worked so hard to enunciate? I hope so.

4 Censorship Incidents

As I pointed out, freedom of expression is at risk on the Internet. I think it's worthwhile to make this concrete by examining some censorship incidents in detail. Since I am most familiar with one Canadian institution, the University of Waterloo, I will focus on that university.

First, a little background. Due primarily to historical accident, universities are currently one of the principal locations where people have free and unlimited access to the Internet, one part of the so-called Information Highway. The Internet is also one of the principal places where Usenet news may be accessed.

Usenet consists of thousands of bulletin boards called "newsgroups," on a variety of topics—a kind of shared electronic mailbox. Users may read messages that have been posted on a particular topic, reply to those messages (by sending electronic mail directly to the poster), or "follow-up" (post a reply to the newsgroup itself). Usenet has existed for about fifteen years, and readership estimates for some newsgroups are in the millions or hundreds of thousands.

Usenet censorship can take place in a variety of ways, some more subtle than others. For example, it is possible for a local system administrator to expurgate a news feed, so that only certain newsgroups get through, and others are blocked. When this is done, the user is typically not informed. It is also possible to block certain postings locally from certain newsgroups, as has recently been done at the University of Kentucky [6]. Finally, messages do not stay forever on the bulletin boards they are posted to: something called an "expire time" governs how long they are available to the public. By differentially setting the expire times, it is possible to control locally which newsgroups actually get read.

The first censorship incident at Waterloo took place in 1988. Brad Templeton, a UW alumnus and operator of a Waterloo-area computer company, moderated a newsgroup called rec.humor.funny, a bulletin board devoted to jokes. People from all over the world sent him jokes; he chose the best ones, and posted them to the Internet. When an ethnic joke offended a student at MIT, he complained to the local newspaper, the Kitchener-Waterloo Record, and the Waterloo administration responded by banning the newsgroup. Ironically, after the ban, compilations of the jokes from the newsgroup could still be found for sale in Waterloo's own bookstore. (Shallit's Third Law!)

More recently, the University administration discovered that some of those thousands of newsgroups dealt with sex. In today's climate—as Trent University professor John Fekete calls it, an atmosphere of "moral panic" [4]—such a thing has become unacceptable.

To give you some idea of what we're dealing with, here are some of the newsgroups you can find on the Internet:

alt.sex.bestiality
alt.sex.bondage
alt.sex.stories
alt.sex.stories.d [d = discussion]
alt.tasteless
rec.arts.erotica
alt.sex.anal
alt.sex.breast
alt.sex.exhibitionism
alt.sex.fetish.feet
alt.sex.fetish.tickling
alt.sex.intergen
alt.sex.masturbation
alt.sex.pedophilia
alt.sex.safe
alt.sex.services
alt.sex.pictures
alt.sex.spanking
alt.binaries.pictures.erotica
alt.binaries.pictures.erotica.fetish
alt.binaries.pictures.tasteless
alt.binaries.multimedia.erotica
ont.personals.whips.and.rubber.chickens

For reasons known only to that arcane bureaucracy known as a University administration, all these newsgroups are currently available at the University of Waterloo, except for the first five. I should point out that all five groups are groups in which text, not pictures, is primarily distrib-

uted. The newsgroups in which pictures are distributed are not yet banned at Waterloo.

How did this censorship happen at Waterloo and other Ontario universities, and why is it being tolerated? I believe (although I cannot prove it) that it started with this September 1992 memo from Bernard Shapiro, Deputy Minister for Colleges and Universities from the Ontario Ministry of Education [19]:

> It has recently come to my attention that computer systems at Ontario's colleges and universities, normally used for the exchange of information between academics and scientific researchers, may be providing access to pornographic and/or racist material through international computer networks.
>
> It is the ministry's position that publicly-funded postsecondary institutions in Ontario should have appropriate policies and procedures in place to discourage the use of their computing systems for purposes of accessing or sending racist or pornographic materials. Furthermore, offensive material should be removed when it is identified, and appropriate sanctions should be in place to deal with offences. . . .
>
> . . . I do not believe that publicly-funded institutions should be seen to support either access to, or distribution of offensive material. . . .

I find this memo bizarre for a number of reasons. First of all, it exhibits no comprehension of the current purpose or use of the Internet. The Internet is not simply used for the "exchange of information between academics and scientific researchers."

Second, the memo exhibits the fallacy of "the medium, not the message." Pornography—a word that is often used pejoratively, but should not be—just means material that is intended to cause an erotic response in the viewer. Pornography is not, per se, illegal in Canada. Many pornographic materials in the print medium are freely available in many Ontario libraries. For example, the University of Waterloo library carries a subscription to *Playboy*, and the University of Waterloo bookstore carries a book called *Women's Erotic Dreams* [16]. Where is the concern and outrage over these materials?

Third, the memo asks for the suppression of *offensive materials* at Ontario universities. I was under the impression (in Clark Kerr's words) that the purpose of a University was to make students safe for ideas, not to make ideas safe for students. If you haven't been offended by *some* idea put forward at a university, then you haven't been paying attention.

Again, my university contains books in its library that are patently offensive to many, including *The Protocols of the Learned Elders of Zion*, Bret Easton Ellis' *American Psycho*, and Arthur Butz's *The Hoax of the Twentieth Century*, a book that claims that the Holocaust is a massive Jewish hoax. Butz's book is banned from importation into Canada, but it is nevertheless freely available in the Waterloo library.

It was not long after the Shapiro memo that action began to happen at Ontario universities. At Waterloo, the University Ethics Committee was empowered to investigate the Internet and decide what material might possibly break Canadian obscenity laws. In February, 1994, based on an opinion from the Ethics Committee, the University administration banned the five newsgroups previously listed. Here is part of the memo from the President of the University, James Downey [17]:

> Last fall I became aware that certain newsgroups on the Internet carried material which was almost certainly obscene and therefore contrary to the Criminal Code. Advice from the University solicitor was unequivocal: under the Criminal Code it is an offence for anyone to publish or distribute obscene material, and the University is running a risk of prosecution if it knowingly receives and distributes obscene material. In these circumstances I felt the University had to act to protect itself. . . .
>
> I am aware, of course, that this is a sensitive area: there is no precise and agreed-on measurement of where on the scale of human taste pornography begins. . . .
>
> I am now authorizing implementation of the following process:
>
> Complaints concerning newsgroups which contain material considered to be obscene are to be referred to the Ethics Committee. The Ethics Committee, with advice from legal counsel as appropriate, will make a recommendation to the Vice-President, Academic & Provost for the removal of any newsgroups it judges to be carrying obscene material. . . .

This memo also troubles me. First, the muddled conflation of "pornography" with "obscenity." Again, pornography is not illegal in Canada—only certain kinds of pornography are illegal. Second, a quick glance at the Criminal Code informs you that one cannot be convicted under obscenity law if "the public good was served by the acts" [18]. Surely guaranteeing free expression at a university is a case of the public good. Third, notice that the stated goal is simply to avoid legal liability. This would be a reasonable objective for a business or corporation, but

not for a university, whose hallmark is the guarantee of freedom of expression.

Finally, obscenity law is traditionally among the most vexing and difficult to interpret of all the criminal laws, even with the recent *Butler* decision to give guidance. As Ontario Judge Stephen Borins once remarked, "Judge or jurors lacking experience in the field of pornography and the attitudes of others toward it face a substantial challenge in making the findings demanded by the law." [14]

Because of this difficulty, the American Library Association offered the following interpretation of its Challenged Materials policy [2]:

> Particularly when sexually explicit materials are the object of censorship efforts, librarians and boards of trustees are often unaware of the legal procedures required to effect the removal of such items. Many attorneys, even when employed by state and local governing bodies, are not aware of the procedures to determine whether or not a work is obscene under the law. According to U.S. Supreme Court decisions, a work is not obscene until found to be so by a court of law, and only after an adversary hearing to determine the question of obscenity. Until a work is specifically found to be unprotected by the First Amendment, the title remains a legal library acquisition and need not be removed.

Although this policy is written for the United States, its principles are equally valid in Canada. Material in Canada is not obscene until declared so by a court; until then it enjoys the protection of the Charter of Rights and Freedoms.

This point was driven home by Canadian Supreme Court Justice John Sopinka, in a November 26, 1994 speech at the University of Waterloo. Mr. Justice Sopinka, author of the *Butler* decision, said:

> Difficult issues also arise in the context of universities which take action to ban certain communications found to be offensive and undesirable. First, one must ask whether it is not preferable to permit the expression and allow the criminal or civil law to deal with the individual who publishes obscene, defamatory or hateful messages rather than prevent speech before it can be expressed. Otherwise, individuals may be putting themselves in the positions of courts to determine what is obscene and what is acceptable. [10]

Isn't this precisely what happened at Waterloo? No Internet newsgroup or message has ever been declared obscene by a court of law. Nevertheless, five newsgroups were banned from the campus.

There is an interesting historical parallel. Back in 1961, four copies of Henry Miller's *Tropic of Cancer* were acquired from Grove Press by the Toronto Public Library. The Department of National Revenue, having declared the book obscene and unfit for importation into Canada, demanded that the Toronto Public Library hand over all copies of the book.

But chief librarian Henry C. Campbell refused. [11] As he pointed out, no Canadian court had declared the book obscene. The *Toronto Star* editorialized, "If the authorities deem *Tropic of Cancer* pornographic, they should test that belief in court. . . . Censorship guided by open court hearings, even on the basis of imperfect law, is preferable to any attempt at censorship by official decree." [12]

Unfortunately, Campbell's principled refusal to turn the book over to the censors at National Revenue was later overruled by Toronto Public Library Board Chair W. Harold Male. But the inner workings of the censorious mind may be judged by the following: Male huffed that "any self-respecting public library shouldn't have it on its shelves," and then was forced to admit that he had never even read the Miller novel. [13]

The sad conclusion: librarians understand the principles of intellectual freedom better than some university administrators.

5 A Simple Principle

We have seen that, true to Shallit's Third Law, the current public perception is that communication on the Internet does not merit protection under the Charter of Rights and Freedoms.

In the meantime, what are we to do? One possibility is to establish and debate fundamental principles on which policy can be based. To that end, I would like to bring your attention to a principle of intellectual freedom for electronic bulletin boards, as enunciated by Carl Kadie. The principles of intellectual freedom developed by libraries should be applied to the administration of information material on computers. [5]

Let us try to apply this principle to two specific cases, and see what results.

First, the case of access to the Internet by minors. As we have seen, people like Reform MP Myron Thomson are worried that children might gain access to pornographic material. Now, as I have pointed out, many public and university libraries in Canada already contain pornographic

materials. For example, the Cambridge Public Library purchased two copies of Madonna's recent book, *Sex*. Following Kadie's principle, we must ask, what special actions have been taken by librarians to restrict access by minors to this kind of pornography?

The answer is, nothing. For example, the American Library Association has a policy on access to library material by minors that reads, in part,

> Library policies and procedures which effectively deny minors equal access to all library resources available to other users violate the LIBRARY BILL OF RIGHTS. The American Library Association opposes all attempts to restrict access to library services, materials, and facilities based on the age of library users. . . .
>
> Every restriction on access to, and use of, library resources, based solely on the chronological age, educational level, or legal emancipation of users violates Article V. . . .
>
> The selection and development of library resources should not be diluted because of minors having the same access to library resources as adult users. Institutional self-censorship diminishes the credibility of the library in the community, and restricts access for all library users. [1]

Although this is an American policy, it is generally adhered to by Ontario libraries. Most Ontario public libraries, including the Cambridge Public Library, have ended their two-tier library card system and now only offer a single library card. Madonna's *Sex* is now freely available to any child with a library card in Cambridge (but they'll have to wait in line to see it, since there is currently a waiting list of 100 people). If parents are worried about the kinds of materials their child might borrow, they are free to refuse permission for their child to obtain a library card. Ontario librarians recognize the right of parents to control their children's reading, but they refuse to act in loco parentis.

In the same way, schools and libraries that provide Internet access should refuse to provide a two-tier service in which some newsgroups are censored or suppressed for children. Should parents worry about the kinds of material their children might encounter on the Internet, they are free to deny access entirely for their children; for example, by not telling them the password.

Let us now examine another problem, that of requesting new newsgroups. In some systems, users are forced to make their request for new newsgroups in public—at the University of Waterloo, for example, some newsgroups are automatically subscribed to, but as of this writing others must be requested by posting to a newsgroup called uw.newsgroups. The result is that some newsgroups—particularly those dealing with sexual topics—may end up not being subscribed to because users are too embarrassed to make their request in front of everyone.

If we apply the intellectual freedom principles enunciated by libraries, however, we see that some other method for requesting newsgroups should be provided. For example, Article III of the ALA's "Librarian's Code of Ethics" states [2]: Librarians must protect each user's right to privacy with respect to information sought or received and materials consulted, borrowed, or acquired.

I believe that the principles librarians have developed for traditional media are a good basis for the protection of the new electronic media.

6 Why EFC?

The Internet and related communications technologies are going to change the way we communicate and research in the 21st century. Rules will be needed to make sure that everyone has a chance to participate, and to prevent abuse of the technology. But those rules should be made with careful thought, by people informed about the possibilities, limitations, and dangers of the technology. It is with this goal in mind that the Electronic Frontier Foundation was founded in the United States in July 1990. But until recently, there was no similar organization in Canada. Professor David Jones (then of McGill University and now of McMaster University) and I founded Electronic Frontier Canada in January 1994. Here is our raison d'etre (based on a similar statement from the Electronic Frontier Foundation): Electronic Frontier Canada (EFC) was founded to ensure that the principles embodied in the Canadian Charter of Rights and Freedoms are protected as new computing, communications, and information technologies emerge.

EFC is working to shape Canada's computing and communications infrastructure and the policies that govern it, in order to maintain privacy, freedom of speech, and other democratic values. Our work focuses on the establishment of:

- clear institutional policies and new laws that guarantee citizens' basic rights and freedoms on the electronic frontier;

- a policy of common carriage requirements for all network providers so that all forms of speech and expression, no matter how controversial, will be carried without discrimination;
- a diverse electronic community that enables all citizens to have a voice in the information age.

I hope that EFC will become a voice for reason and education as the electronic frontier becomes more civilized. And I also hope that librarians and their understanding of intellectual freedom principles will be at the forefront of the civilizing process. We need you.

References

1. American Library Association, "Free Access to Libraries for Minors: An Interpretation of the Library Bill of Rights," July 3, 1991. (Available by gopher or anonymous FTP to gopher.eff.org.)

2. American Library Association, *Intellectual Freedom Manual*, 3rd edition, 1989. (Sections also available by gopher or anonymous FTP to gopher.eff.org.)

3. Jonathan Emord, *Freedom, Technology, and the First Amendment*, Pacific Research Institute for Public Policy, 1991.

4. John Fekete, *Moral Panic: Biopolitics Rising*, Robert Davies Publishing, 1994.

5. Carl Kadie, "Content: The Academic Freedom Model," paper delivered at the *Third Conference on Computers, Freedom, and Privacy*, Burlingame, California, March 1993. Full text available at ftp://ftp.eff.org/pub/CAF/statements/cfp93.kadie.

6. Carl Kadie, "Applying Library Intellectual Freedom Principles to Public and Academic Computers," paper delivered at the *Fourth Conference on Computers, Freedom, and Privacy*, March 1994. Full text available at http://www.eff.org/CAF/cfp94.kadie.html.

7. Carolyn Marvin, *When Old Technologies Were New: Thinking About Electric Communication in the Late Nineteenth Century*, Oxford University Press, 1988.

8. Corey L. Nelson and Bonnie Brown, "Is E-mail Private or Public?" *Computerworld*, June 27, 1994, pp. 135–137.

9. Ithiel de Sola Pool, *Technologies of Freedom*, Belknap Press of Harvard University Press, 1983.

10. John Sopinka, "Freedom of Speech and Privacy in the Information Age," text of speech delivered at the University of Waterloo, November 26, 1994. Text available at gopher://insight.mcmaster.ca/00/org/efc/doc/sfsp/sopinka.

11. "Librarian Refuses to Give Banned Novel to Customs," *Toronto Globe & Mail*, October 30, 1961, p. 5.

12. "Censorship by Decree," *Toronto Star*, October 31, 1961, p. 6.

13. "Banned Book," *Toronto Globe & Mail*, November 27, 1961, p. 6.

14. Quoted in Lynn King, "Censorship and Law Reform" in *Women Against Censorship*, Varda Burstyn, ed., Douglas & McIntyre, 1985, p. 86.

15. Canadian Library Association, Intellectual Freedom Statement. Full text available at gopher://insight.mcmaster.ca/00/org/efc/library/library-cla-policy.

16. Celeste T. Paul, *Women's Erotic Dreams (and What They Mean)*, Grafton Books, London, 1988.

17. Memo from University of Waterloo President James Downey, January 31, 1994. Full text available at gopher://insight.mcmaster.ca/00/org/efc/univ/waterloo/uw.memo.netnews.31jan94.

18. Criminal Code of Canada, Section 163 (3). Full text available from http://insight.mcmaster.ca/org/efc/pages/law/cc/cc.163.html.

19. Memo from Bernard Shapiro, September 1992. Full text available from gopher://insight.mcmaster.ca:70/00/org/efc/univ/ontario.univ-ministry.memo.