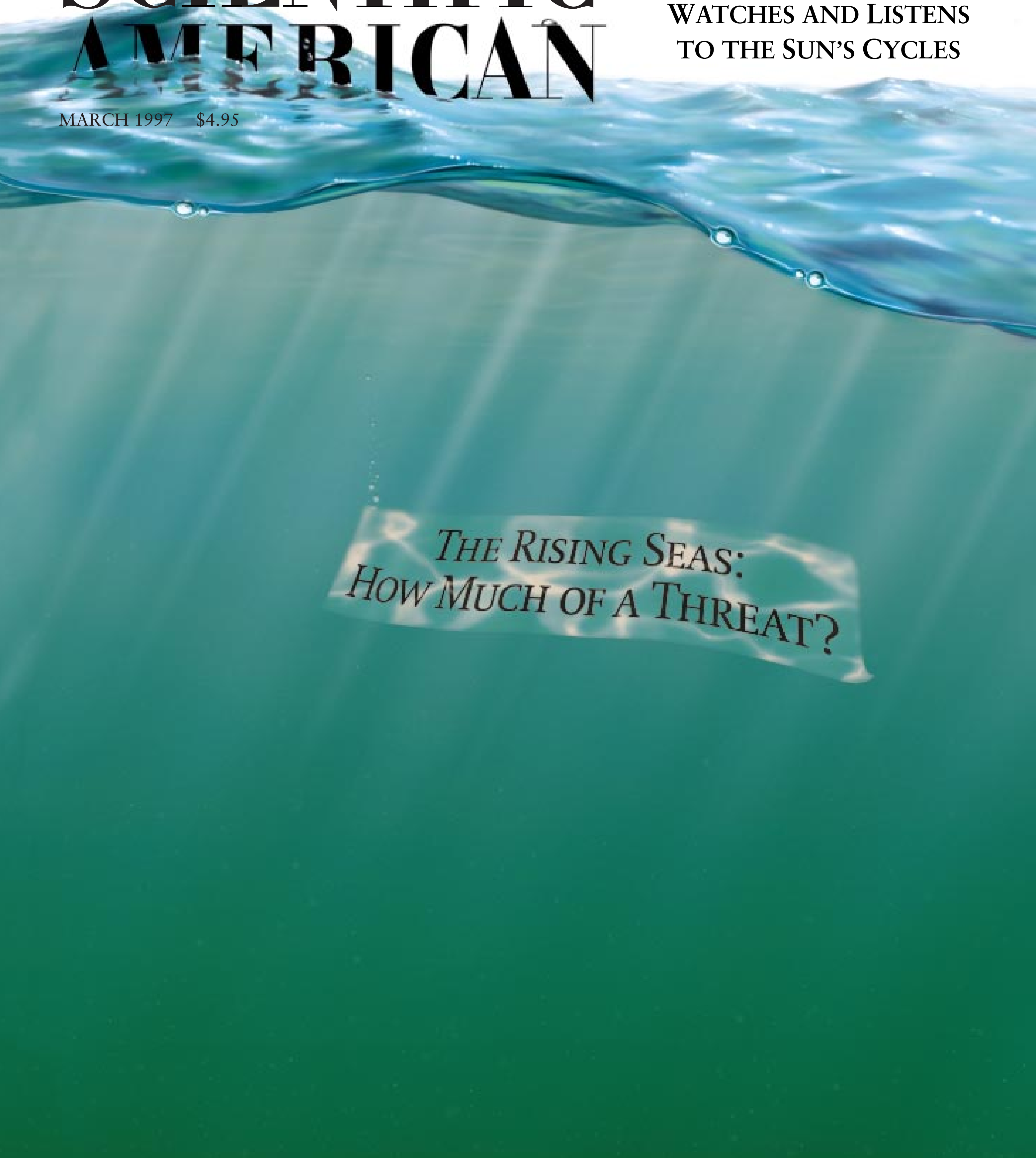
SPECIAL REPORT THE INTERNET: FULFILLING THE PROMISE SPECIAL REPORT

SCIENTIFIC AMERICAN

MARCH 1997 $4.95

SOLAR SECRETS

OBSERVATORY IN SPACE WATCHES AND LISTENS TO THE SUN'S CYCLES



THE RISING SEAS: HOW MUCH OF A THREAT?

Greek quotation in a Russian text that will be properly displayed on the reader's computer in South America. Software standards with this kind of capability are emerging. But the primary software producers, in their race to dominate the market, keep producing new versions, giving little chance to the usually small enterprises that develop multilingual products to keep up.

In real life, interpreters help to overcome language barriers. Human translators can also be employed on the Internet, but given the volume and variety of exchanges, they will play a limited role. Only machine-aided translation can bring us closer to a world, perhaps a utopia, where all the attendees at a virtual conference of the United Nations can each use his or her native language, which will be simultaneously translated into all other languages.

Research on machine-aided translation has been pursued over the past 50 years with somewhat mixed results. The systems actually in use are small in number and located mostly in Japan, Canada and Europe—the last of which faces the largest multilingual translation load. Electronic interpreters are usually just bilingual and need to be heavily specialized if they are to produce raw translations good enough to be revisable by human editors.

The first system available for general public use was Systran, which could translate 14 pairs of languages and was accessible as early as 1983 on the French Minitel network. Used by the European Commission, Systran now converts hundreds of thousands of pages a year. Another success story is the Meteo system, which translates Canadian meteorological bulletins between English and French. It handles 80,000 words (about 400 bulletins) every day, with only three to five human editing operations for every 100 words.

Multilingual translation will benefit from a two-step process now being developed by several groups. The text is first thoroughly analyzed into component parts (title, paragraph, sentence), clarified when possible by a dialogue with the author, then translated into an intermediate, abstract representation—which is used to generate translations in different languages. The effort is worth the expense when the text needs to be translated into more than 10 languages. The United Nations University in Tokyo has recently announced a 10-year collaborative project for implementing this two-stage scheme.

But a truly multilingual Internet will come to pass only with concerted international effort. Will we give it enough priority? The answer is not clear. It is so easy to let ourselves drift toward English as a unique common language.

BRUNO OUDET chairs the French chapter of the Internet Society. He holds both U.S. and French Ph.D.'s in economics. Oudet is a professor at Joseph Fourier University in Grenoble and a researcher at the Leibniz Laboratory of the IMAG Institute.

TRUSTED SYSTEMS

Devices that enforce machine-readable rights to use the work of a musician or author may create secure ways to publish over the Internet

by Mark Stefik

Everyday experience with computers has led many people to believe that anything digital is ripe for copying—computer programs, digital books, newspapers, music and video. Some digital-age pundits have gone so far as to proclaim that the ease of duplicating data heralds an end to copyright: information "wants to be free," they assert. It is impossible to thwart the spread of information, so the argument goes. Anything that can be reduced to bits can be copied.

This provocative notion undermines the dream behind the creation of the Internet: the possibility of universal access in a digital age—where any author's work could be available to anyone, anywhere, anytime. The experience of most people, however, is not that the Net contains great works and crucial research information. Instead most of what is there is perceived to be of low value.

JEFF BRICE

The root of the problem is that authors and publishers cannot make a living giving away their work. It now takes only a few keystrokes to copy a paragraph, an entire magazine, a book or even a life's work. Uncontrolled copying has shifted the balance in the social contract between creators and consumers of digital works to the extent that most publishers and authors do not release their best work in digital form.

Behind the scenes, however, technology is altering the balance again. Over the past few years, several companies, including Folio, IBM, Intertrust, NetRights, Xerox and Wave Systems, have developed software and hardware that enable a publisher to specify terms and conditions for digital works and to control how they can be used. Some legal scholars believe the change is so dramatic that publishers will be left with too much power, undercutting the rights and needs of consumers and librarians.

Yet consumers' needs can be served

even as this transformation progresses. As technology brings more security, better-quality works will reach the Net. Noted authors might be willing to publish directly on the World Wide Web. Although information might not be free, it will most likely cost less because of lower expenses to publishers for billing, distribution and printing. These savings could be passed on to consumers.

The key to this technological shift is the development of what computer scientists know as trusted systems: hardware and software that can be relied on to follow certain rules. Those rules, called usage rights, specify the cost and a series of terms and conditions under which a digital work can be used. A trusted computer, for instance, would refuse to make unauthorized copies or to play audio or video selections for a user who has not paid for them.

Trusted systems can take different forms, such as trusted readers for viewing digital books, trusted players for playing audio and video recordings, trusted printers for making copies that contain labels ("watermarks") that denote copyright status, and trusted servers that sell digital works on the Internet. Although the techniques that render a system trustworthy are complex, the result is simple. Publishers can distribute their work—in encrypted form—in such a way that it can be displayed or printed only by trusted machines. At first, trusted security features would be bundled into a printer or handheld digital reader at some additional cost to the consumer, because they would provide the ability to access material of much higher value. Eventually the costs would fall as the technology became widely implemented. Of course, a publisher could still opt to make some works available for free—and a trusted server would still allow anyone to download them.

How does a trusted system know what the rules are? At Xerox and elsewhere, researchers have attempted to express the fees and conditions associated with any particular work in a formal language that can be precisely interpreted by trusted systems. Such a usage-rights language is essential to electronic commerce: the range of things that people can or cannot do must be made explicit so that buyers and sellers can negotiate and come to agreements. Digital rights fall into several natural categories. Transport rights include permission to copy, transfer or loan. Render rights allow for playing and printing. Derivative-work rights include extracting and editing information and embedding it in other publications. Other rights govern the making and restoring of backup copies.

How Trusted Systems Work

Different intellectual works have different security requirements. But trusted systems allow publishers to specify the required security level to safeguard a document or video. The most valuable digital properties might be protected by systems that detect any tampering, set off alarms and erase the information inside. At an intermediate level, a trusted system would block a nonexpert attack with a simple password scheme. And at a lower security level, it would offer few obstacles to infringers but would mark digital works so that their source could be traced (such digital watermarking is now embedded in some image-manipulation software).

Most trusted computers have the capability to recognize another trusted system, to execute usage rights and to render works so that they either cannot be copied exactly or else carry with them a signature of their origin. For executing a highly secure transaction, two trusted systems exchange data over a communications channel, such as the Internet, providing assurances about their true identities. Managing communications over a secure channel can be accomplished with encryption and what are known as challenge-response protocols.

One example of the use of this protocol would be if computer A wishes to communicate with computer B. Computer A has to prove to B that it is a trusted system and that it is who it says it is. The interaction begins when A sends B a digital certificate confirming that it has entered its name with a registry of trusted systems. B decrypts the certificate. This action confirms that the certificate is genuine. But because a certificate can be copied, how does B know it is really in communication with A? To verify A's identity, B composes a random string of numbers called a nonce. It encrypts the nonce with a public software key that A has sent within the digital certificate. The public key allows B to send messages that only A will understand when it decrypts them with its own private key.

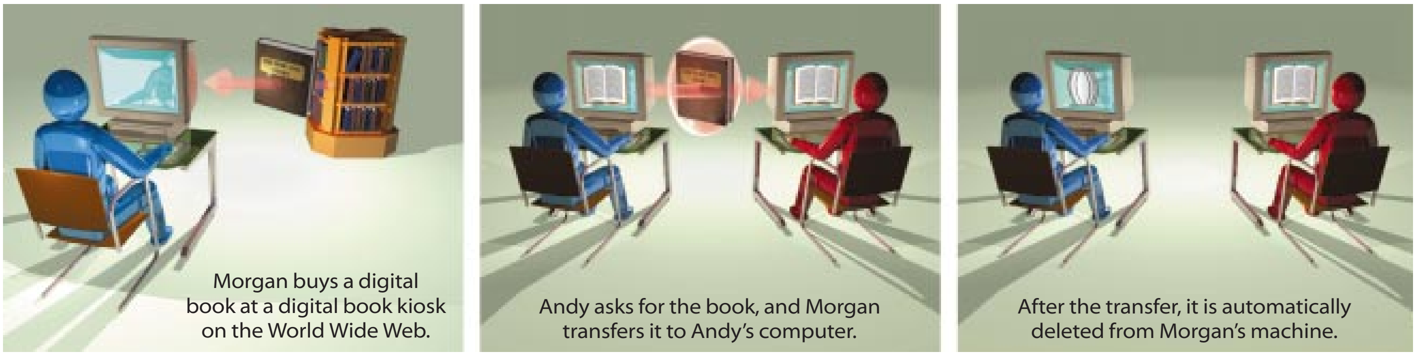
B sends the nonce to A, which decrypts it and returns an unencrypted message to B containing the numbers in the nonce. If the return message matches the one it first sent, B knows it is, in fact, communicating with A, because only A could have decrypted the message with its private key. In a few more steps, the two computers may be ready to transfer a book or carry out some other transaction.

COURTESY OF VIRGIN RECORDS; JASON GOLTZ

TRY BEFORE YOU BUY—a sales method for music stores—may endure with trusted systems that allow for sampling and buying over the Internet.

Although not all trusted systems use

TRANSFER


Morgan buys a digital book at a digital book kiosk on the World Wide Web.
Andy asks for the book, and Morgan transfers it to Andy's computer.
After the transfer, it is automatically deleted from Morgan's machine.

BRYAN CHRISTIE

TRANSFERRING A DIGITAL WORK from one trusted system to another resembles transferring money from a savings to a checking account. The money is in one account or the other. Digital-property rights distinguish the right to copy a digital work (which increases the number of copies) from the right to transfer a digital work (which preserves the number of copies).

a challenge-response protocol, most use encryption for exchanging digital works. They may also incorporate other security features. Some systems contain tamperproof clocks to prevent a user from exercising expired rights. Others have secure memories for recording billing transactions. Still others must be connected to an on-line financial clearinghouse during transactions.

Trusted systems can place identifying watermarks that make it possible to track down unauthorized duplications or alterations. Watermarks maintain a record of each work, the name of the purchaser and a code for the devices on which they are being played. This information can be hidden—in the white space and gray shades of a text image, for instance. As such, the identifying information would be essentially invisible to lawful consumers—and unremovable by would-be infringers.

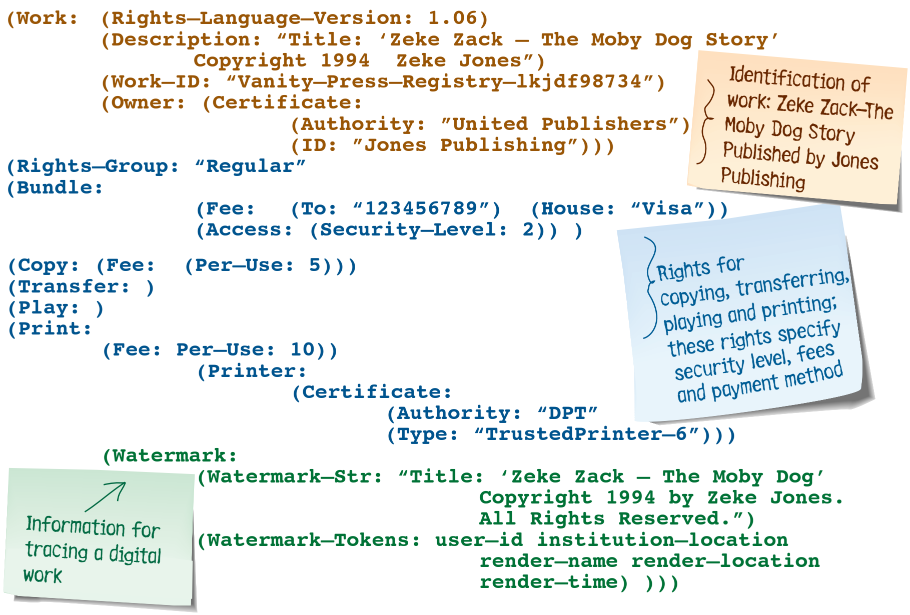
Publishers would still need to be watchful for unlicensed distribution of their property. A computer user can always print a digital page and then photocopy it. A digital-movie pirate can sit in front of the screen with a camcorder. What trusted systems prevent, however, is the wholesale copying and distribution of perfect digital originals. With appropriate watermarks, for instance, even pirated copies should still be traceable.

In digital publishing, trusted systems would allow commerce to proceed in a manner not unlike the way it is carried out in the distribution of paper copies. Suppose that Morgan wishes to buy a digital book on the Web [*see top illustration on these two pages*]. Copying the book initiates a transaction between the seller's system and Morgan's computer. At the end of the transaction, Morgan has used a credit card or digital cash to buy a copy of a book that can be read with a personal computer or some other digital reader. The entire transaction, moreover, is preceded by an exchange of information in which the seller ensures that Morgan's machine is a trusted system.

Exercising Usage Rights

As with a paper book, Morgan can give away his digital opus. If Morgan's friend Andy asks for it, Morgan can exercise a free-transfer right. At the end of the transaction, the book resides on Andy's reader and not on Morgan's. Andy can then read the book, but Morgan cannot. The transfer preserves the number of copies. Their computers, reading and interpreting the rights attached to the file containing the book, perform the transfer in this way, and neither Morgan nor Andy can command otherwise.

Morgan can also lend a book to a friend. If Ryan wants to borrow a book for a week, Morgan can transfer it to his computer, but while the digital book is on loan, Morgan cannot use it. When the week runs out, Ryan's system deactivates its copy, and Morgan's system marks its copy as usable again. Without any action by either of them, the digital book has been "returned" to its lender. The right to lend is crucial in enabling



```
(Work:  (Rights–Language–Version: 1.06)
        (Description: “Title: ‘Zeke Zack – The Moby Dog Story’
                Copyright 1994  Zeke Jones”)
        (Work–ID: “Vanity–Press–Registry–lkjdf98734”)
        (Owner: (Certificate:
                        (Authority: ”United Publishers”)
                        (ID: ”Jones Publishing”)))
(Rights–Group: “Regular”
(Bundle:
                (Fee:   (To: “123456789”)  (House: “Visa”))
                (Access: (Security–Level: 2)) )

(Copy: (Fee:  (Per–Use: 5)))
(Transfer: )
(Play: )
(Print:
        (Fee: Per–Use: 10))
                (Printer:
                        (Certificate:
                                (Authority: “DPT”
                                (Type: “TrustedPrinter–6”)))
        (Watermark:
                (Watermark–Str: “Title: ‘Zeke Zack – The Moby Dog’
                                        Copyright 1994 by Zeke Jones.
                                        All Rights Reserved.”)
                (Watermark–Tokens: user–id institution–location
                                        render–name render–location
                                        render–time) )))
```

Identification of work: Zeke Zack–The Moby Dog Story Published by Jones Publishing

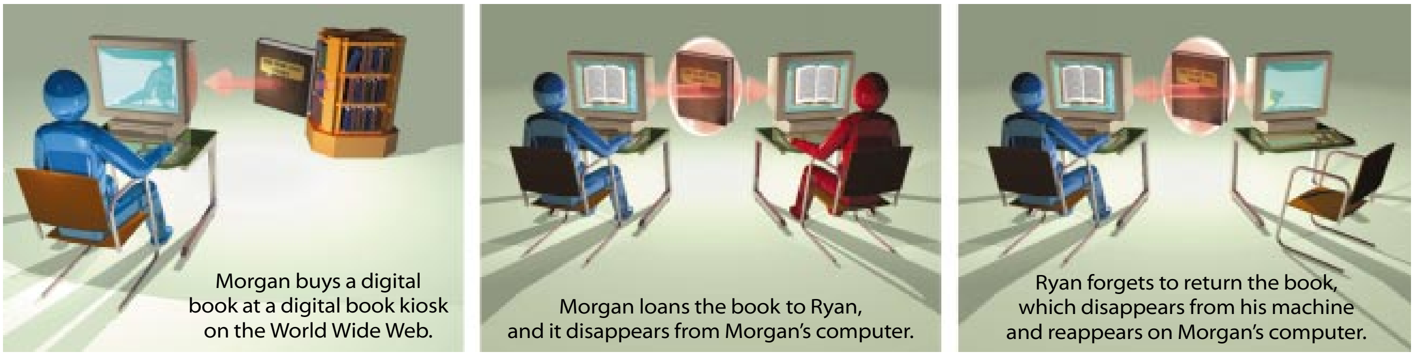
Rights for copying, transferring, playing and printing; these rights specify security level, fees and payment method

Information for tracing a digital work

JENNIFER C. CHRISTIANSEN

USAGE RIGHTS, the terms and conditions for a trusted digital book, called *Zeke Zack—The Moby Dog Story,* are written in a machine-interpretable language.

LOAN


Morgan buys a digital book at a digital book kiosk on the World Wide Web.

Morgan loans the book to Ryan, and it disappears from Morgan's computer.

Ryan forgets to return the book, which disappears from his machine and reappears on Morgan's computer.

Copying a work usually requires paying a fee, whereas transferring a work does not. Loaning a digital work is another transaction distinct from copying and transferring. The process is analogous to loaning a book in the sense that it temporarily grants use of the work to a second party, and the owner of the work cannot use it while it is loaned out.

the establishment of digital libraries.

Usage rights can be tailored for reading a work, printing it or creating derivative works. Depending on the publisher, particular rights can carry a fee or not. For some works, copying is free, but viewing them costs a fee. Fees can be billed for each use or by the hour; they may be billed when the user obtains the work or whenever a right is exercised. There can be discounts, sales and free trials. Distribution can be limited to people who can certify that they are members of a book club, a certain age group or citizens of a particular country.

Trusted systems can also respect the type of fair-use provisions that currently apply to libraries and some other institutions, allowing a reasonable number of free copies or quotations to be used. Members of the public with special needs—librarians, researchers and teachers—could receive licenses from an organization representing publishers that let them make a certain number of free or discounted copies of a work, if the rights of an author are understood. To balance against the risks of illegal copying, an insurance fund could be set up to protect against losses.

What's in all this for consumers? Why should they welcome an arrangement in which they have less than absolute control over the equipment and data in their possession? Why should they pay when they could get things for free? Because unless the intellectual-property rights of publishers are respected and enforced, many desirable items may never be made digitally available, free or at any price. Trusted systems address the lack of control in the digital free-for-all of the Internet. They make it possible not only for entire libraries to go on-line but also for bookstores, newsstands, movie theaters, record stores and other businesses that deal in wholly digital information to make their products available. They give incentives for 24-hour access to quality fiction, video and musical works, with immediate delivery anywhere in the world. In some cases, this technological approach to protecting authors and publishers may even avoid the need for heavy-handed regulations that could stifle digital publishing.

Fully realizing this vision will necessitate developments in both technology and the marketplace. Users will need routine access to more communications capacity. Publishers must institute measures to ensure the privacy of consumers who use trusted systems, although the same technology that guards the property rights of publishers could also protect personal details about consumers. Trusted systems also presume that direct sales, not advertising, will pay the costs of distributing digital works. Advertising will most likely prevail only for works with substantial mass-market appeal. By protecting authors' rights, trusted systems will enable specialized publishing to flourish: compare, for instance, the diverse collection of books in a library to the relative paucity of programs for television.

The dynamics of a competitive marketplace form the most imposing roadblock to fashioning protections for digital rights. Several companies have trusted systems and software in the early stages of testing. With some exceptions, though, the software is proprietary and incompatible. Whereas the technology could provide the infrastructure for digital commerce, the greatest benefits will accrue only if the various stakeholders, from buyers and sellers to librarians and lawmakers, work together. SA

The Author

MARK STEFIK is a principal scientist in the information sciences and technology laboratory at the Xerox Palo Alto Research Center. Stefik received a bachelor's degree in mathematics and a doctorate in computer science, both from Stanford University. In 1980 he joined Xerox PARC. There he has done research on expert systems, collaborative systems and technology for digital commerce. For several years, Stefik taught a graduate course on expert systems at Stanford. He wrote the textbook *Introduction to Knowledge Systems* (Morgan Kaufmann, 1995).

Further Reading

LEGALLY SPEAKING: REGULATION OF TECHNOLOGIES TO PROTECT COPYRIGHTED WORKS. Pamela Samuelson in *Communications of the ACM,* Vol. 39, No. 7, pages 17–22; July 1996.

FORUM ON TECHNOLOGY-BASED INTELLECTUAL PROPERTY MANAGEMENT: ELECTRONIC COMMERCE FOR CONTENT. Edited by Brian Kahin and Kate Arms. Special issue of the *Interactive Multimedia News,* Vol. 2; August 1996.

LETTING LOOSE THE LIGHT: IGNITING COMMERCE IN ELECTRONIC PUBLICATION. Mark Stefik in *Internet Dreams: Archetypes, Myths, and Metaphors.* Edited by Mark Stefik. MIT Press, 1996.