

Government information wants to be free.

by Glenn Tenney

An information revolution

Over the last few years we've witnessed an information revolution. It used to cost thousands of dollars to publish a book, yet now for a dollar or two we can produce a single CD-ROM containing the text of over two hundred books. As we cultivate this new digital frontier, everything we once understood about publishing, intellectual property, and knowledge itself needs to be rethought.

One example of the impact of digital information is that some people have said "information wants to be free." To most publishers and authors, this is a frightening thought since their very livelihood appears threatened by such a radical proposition. There is an aspect to this that must be better understood. Publishing information on a CD-ROM still requires selling something of substance to the customer. The digital publishing frontier is running headlong into the on-line frontier where nothing tangible is being sold, only the intangible access to the information.

Even though many of us do not agree that "all" information wants to be free, perhaps we could agree that some class of information does want, or needs, to be free. The very basis for our democratic government dictates that the people have the power to elect their own government, and that is closely coupled to the people having access to the data with which to make an informed choice. Our society, therefore, needs open access to the information produced by our government.

If we postulate that "all government information wants to be free" we can avoid the emotional backlash that copyright owners and publishers have expressed when they believed that their property rights were being attacked. However, all intellectual property owners will have to deal with the new realities of the emerging digital electronic frontier.

Today it costs a little more than a dollar to manufacture a CD-ROM that can then be sold for ten, hundreds, or thousands of dollars. That same information can also be placed on a computer network once and then it would be available to millions of people at no additional charge. The new era of digital publishing has already dawned, how we will adjust to it is the question.

Accessibility of government data

Earlier this year I wanted to know how my representative to the U. S. Congress voted. This sounds like a fairly simple request, but alas the answer was not so simple. My first try was to contact my representative's office. Their response was to point me to their library of the Congressional Record saying "Go look it up yourself, there are hundreds of votes each year."

Since the Congressional Record is maintained on a computer, there must be some way to obtain the voting record of a single member of Congress. Special interest groups keep track of votes of interest to them, and each member's own office must track the member's votes. I found that there is an on-line service available for \$6,000 per year. If a corporation needed the information, the fee would not be exorbitant. If the corporation were doing the research relative to a government contract, I'm also sure that the fee would be passed back on as a part of the "time and materials" or at least as "overhead" charges. But to the millions of average citizens, much of the government's data is being kept from them by such methods.

Let's say that a citizen wants to find some government information. First, one needs to find out if such information even exists in any form that's available. To make this example easy, we'll assume that the information is in a document sold by the NTIS. The citizen's task would then be to locate the document. The NTIS index is available on-line via Dialog which places it outside affordability for most people which means using either the paper or microfiche index. In the San Francisco bay area, my county's library system does not have a copy of the NTIS index. Instead you would be referred to a depository at a private university thirty minutes away. Even after finding a document, one would likely have to purchase it since it may not be available nearby, nor for interlibrary loan. This scenario would of course be much worse for people outside of a heavily populated area.

Even when the computer readable data can be found, there is an ever growing trend to keep the data itself unavailable: More and more agencies are packaging their data using proprietary software which makes the actual data completely inaccessible except via that proprietary software.

The Patent and Trademark Office's (PTO) CD-ROMs are one of many similar examples. The data on these discs are clearly public domain data available to anyone. Yet the data is encoded onto the CD-ROMs and the PTO can not disclose how to decode the data. If you don't use the hardware and software that was chosen, you can not access the data at all. You can only get to the data the way the that government allows you. That may be enough, but then again, you might need to search using some other criteria (perhaps searching a combination of fields not supported), or you might want to collect the results of your searches in some way that was not anticipated by the

government. Of course, conspiracy theorists might imagine even worse reasons for a proprietary search program.

It's clear that government data needs to be freely accessible. That should mean that it does not require a proof of one's financial worth to access the data, nor does it mean that we should be able to search the government's files only in the way that one agency wants. But such availability does not come for free. Government data being available to the population is to the benefit of the nation as a whole. The benefit accrues to all of us, not just to those who can afford it, or to those who have a certain brand of computer. Therefore, basic access to government data must be provided for as large a base as possible.

There is already in place a collection of groups that deal with such problems. The library system, although grossly under funded, is well suited to making information available to the citizenry. We should be finding ways to assist in getting the information out to as many people as possible, rather than preserving it for the special few.

Computer readable government data

If you knew that the information you wanted were available from the NTIS in computer readable form, your first attempt would be to contact the NTIS. But you would be shocked to learn that the catalog of computer products would cost you nearly seventy dollars. Imagine calling up a major mail-order company such as L. L. Bean and being told that their catalog, which you need, cost that much!

Various agencies make computer readable data available at extremely varying prices, from tens of dollars to thousands of dollars: a CD-ROM of national geographic and geologic data is available for around thirty dollars; a CD-ROM with census and map data for a few states per disc costs about two hundred-fifty dollars; the CD-ROM of the digital chart of the world costs several hundred dollars; the patent and trademark office's CD-ROMs with abstracts of patents over the last three years cost about eighty dollars each; the Library of Congress' CD-ROMs cost about a hundred dollars each; some CD-ROMs from NTIS cost about one thousand dollars each; a computer tape subscription to the Federal Register costs nearly \$40,000 per year and the Congressional Record costs nearly \$30,000.

In discussions with staff at various agencies producing these CD-ROMs, a trend is clear: these products are viewed as a revenue stream. The agencies maintain that the government is charging less than what commercial ventures have charged for similar data, but the government is charging "what the market will bear". These fees may be fine for government contractors who will add those costs to their overhead charges, but the average citizen is being forced into a position of subservience unable to get the information they need to make informed choices. Since many of these

CD-ROMs are produced to save time and money for the government itself, and to be made available to depository libraries, the discs should be made available at their true marginal cost -- a few dollars each.

But just producing less expensive CD-ROMs is not enough. Our library system is valuable in part because many people do not want to or are not able to purchase everything that they want to read. We need to make computer readable data available in a way similar to books in a library. There is already a network in place of computers at universities and companies around the world. The Internet was begun to aid research for governmental projects (eg. DARPA, NSF, DOE, etc.) but the network has grown to be the telecommunications infrastructure of the ninety's. This network will also be the on-line library of tomorrow.

Instead of just making the Congressional Record available on one computer for \$30,000, it could be made available to hundreds of thousands of computers around the country at the same time at very little extra cost. The private sector could then jump in and provide value added services such as indexing and searching, and charging for access to the Internet; and libraries could also offer access to the on-line documents.

Whither goes intellectual property in a digital on-line world?

More and more people are using on-line services. Many office computers are now connected by local area networks so that people can easily share information. It's become commonplace for one person in an office, or department, to find some interesting tidbit from some on-line database and to then post the results for everyone in that office. Thousands of people are doing this every day without regard to copyright restrictions.

Back in the mid 1970's when CONTU was looking at intellectual property rights of new technologies such as computer software, I commented that the record industry would react when records became digital -- each copy sold would be as good as the master and could replicate as many copies as desired and each copy would be as good as the first copy purchased. The recording industry DID react! That is why we now have special chips inside some digital players keeping people from making "politically incorrect" copies.

When data is available in digital form, every copy is a master, every copy is as good as the first copy. Each copy can be copied and passed around at no additional cost. People are mostly unfamiliar with copyright regulations and are already treating digital information as if it were popcorn to be shared with all of your friends.

Our intellectual property laws are based on the technologies and economic situations that are no longer true: printing presses used to be rare and expensive, today copying machines are commonplace and relatively inexpensive; books were expensive to print and produce, today a book can be

produced on your home computer; and today, you can send out a hundred thousand copies of your book for free. Some private sector companies charge a flat monthly fee for access to the Internet, whether you send one character or a hundred books, it costs the same. With the Internet, I can send a book once and make it available to a hundred thousand people at no additional charge than what I already spend each month for access to the network.

Since the basis for many of our intellectual property laws, specifically copyrights, has changed in ways that were never imagined, and since people are already ignoring the existing laws, something must be done. The concept of infringing someone's copyright was based on a time when it cost a lot to do so. Our country's founders could never have anticipated "shareware" where hundreds of thousands of copies of an author's works are given away hoping that some people will pay. It was unimaginable to think that publishing would not cost anything. We need to dramatically restructure intellectual property laws to adjust to the realities of the electronic frontier as well as to our constitutional mandate "To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries".

The challenge is: How does an author continue to be rewarded in an on-line environment that encourages wanton copying with no established mechanism for compensating the author?

About the author

Mr. Tenney has been designing software and hardware systems since the early 1960's. He is a senior member of the IEEE and has served on the IEEE USA's Intellectual Property Committee for many years. For years Mr. Tenney has been actively working, both professionally and as a volunteer with many organizations, on making information accessible. Mr. Tenney can be reached via email: tenney@well.sf.ca.us.

FIRST INTERNATIONAL SYMPOSIUM: NATIONAL SECURITY & NATIONAL COMPETITIVENESS: OPEN SOURCE SOLUTIONS Proceedings, Volume II - Link Page

[Previous](#) [Commercial Remote Sensing: Open Source Imagery intelligence](#)

[Next](#) [Ranked Retrieval and Extraction of Open Source Intelligence](#)

[Return to Electronic Index Page](#)