

# Ethics of Digital Librarianship

Brewster Kahle

(brewster@archive.org)

Thinking Machines

February 1992

“As digital librarian, you should serve and protect each patron as if she is your only employer.”

As more of us become involved in serving information electronically to other users, we so-called “digital librarians” must become conscious of our ethical responsibilities to protect the privacy of our the users being served. Since computers are being used by many more people to find answers from diverse information sources, we librarians that operate these servers are coming exposed to the exact questions and interests of people we do not know. This information has power, a power that can be abused and thereby thwart the usefulness of the tools we promote. In this essay, I will use the Wide Area Information Server system as an example of a system of digital librarians to show what information is collected and used. With this example, I hope to illustrate some of the dangers and help list some of the rules of etiquette for this emerging class of information providers.

The Wide Area Information Server (WAIS) system is an electronic publishing system that allows end-users to ask questions of remote information sources. The system encourages people to ask questions in natural language so that the server system can try its best to find appropriate documents. Therefore the operator of the server can collect the questions, and importantly, collect what documents the users thought were worth looking at. This combines to portray exact interests of the users. While the identity of the user is not trivial to determine since only the machine that the query came from is accessible from the server logs, as personal computers become networked, the identity of the machine will approximate the identity of the user.

On the positive side, this means that the server operator (the “digital librarian”) can use that data to refine the database and the search techniques used in the system. On the negative side, this is exposing many remote operators to private information that may not be consciously given by the users.

This surrender of information is not new to librarians; and the responsibility is taken very seriously by the professionals in the field. Through training in library schools and by an intuitive sense of ethics, reference librarians do not betray their patron’s interests to others that are curious or devious. This ethical code is not coded in law as it is with psychiatrists, so these records can be extracted through subpoena, but this level of demand is usually required to pry the information from librarians. From the patron’s point of view, having a librarian know what she is interested in can be a great value because the librarian can help select and route useful information in the future.

The same type of information is available to the digital librarians of the WAIS system. I operate the directory of servers in the WAIS system, and as such, I know what users are requesting access to what type of servers. I know, for instance, every time Mitch Kapor uses the system, and what he asks for (he specifically allowed me to include his name here). At this point this is not a problem since few servers are of a personal nature yet, but as the system grows to include entertainment, employment, health and other servers, it is easy to imagine the types of information that will be accessible through operating such a server. Furthermore, I know when particular users are at their machines, and therefore know where they are and when.

The abuses possible with this information are often not as direct as other offenses, but should not be discounted. People will act differently if they think they are being watched. Most people will try not to look silly or ignorant in public, and therefore might be less willing to try something new, to learn about a subject that they know nothing about. If using a WAIS server feels like raising one’s hand in school, then people will craft their questions more carefully than if it felt more like browsing through a new book. Often people say “I have nothing to hide,” which may be true, but if a stranger approaches on the street and knows quite a bit of personal information, then the innocent will likely take that person more seriously than if a cold stranger approached. Even with nothing to hide, most people feel they should who knows what

about them. The personal nature of information access makes distributing collected questions a bit unnerving.

The information collected by the digital librarians have some different characteristics from physical librarians which can make abuse easier and more widespread: more people can be served, these people are often in other organizations, and the digital librarians rarely have personal contact with these users. Therefore, the patrons seem further away and therefore less real as human beings. Since the computer networks that are being used with WAIS span the globe and span company boundaries, the information collected can be useful in knowing what is important to a distant, and possibly competitive group. The lack of human contact can lead to the decay in social relations as has been documented in studies of electronic mail where the language and nature of relations tend to be stripped of grace, etiquette, and often respect [cite Sherry Turkle]. This detached nature of electronic interaction might lead librarians to not respect their patrons' interests where they would if they knew them personally.

On the other hand, the information collected from patrons can be very useful to the digital librarian to refine and enhance the server. An example of this is a reporter at a financial newspaper. She is in the business of collecting information from corporate contacts, finding the trends in that information, throwing out the proprietary details, and selling it back to that same population. If the reporter published too many details, then her contacts would not be forthcoming the next time, and if she sanitized the information to the point of uselessness, similarly, her contacts would not invest the time. Therefore, it is precisely the interaction with the users that builds the information that is sold. This example shows another facet, and that is value that the contacts invest in the reporter for their own benefit. The digital librarian is a less extreme case, but still she is being invested and entrusted with what the users want, and if this information is misused or not used, then the users will not be as well served as could be. Thus, the users will want to be able to be served better by the librarian through feedback on services rendered.

While there are some technological mechanisms to obscure the identity of the patron, such as encryption and redirection, hopefully these will only be used in extreme cases. Encryption can be used to protect packets in transmission and also be used to sign packets so that they can not be forged [cite Whitfield Diffie]. This can be useful in a system where the transport media is insecure, such as radio transmission. Redirection is a server forwarding technique that would concentrate all the requests from one trusted host so that the individual requesters are more difficult to determine. Combinations of these techniques have been contemplated to provably obscure requesters while still providing accountability for charges, but hopefully these techniques will not be the norm if most server operators will act in good faith towards their patrons.

To try to list a code of ethics for this field is difficult since the technology keeps changing, but I will offer a principle that can be used to test a code. As digital librarian, you should serve and protect each patron as if she is your only employer. Therefore each patron should be served and protected individually. In terms of WAIS, I feel it is safe to suggest:

- Don't give away user logs except for scholarly use. Consider sanitizing the records before any transfer is undertaken.
- Take the job of information serving seriously. This means to provide a consistent, reliable service and represent the service provided accurately.
- Count on wide use of the information served, for good uses and bad, so be proud of the information and the collection.
- Completeness is important. Users learn as much from a question that has no answer as from the ones with answers. This requires a complete and up-to-date collection.
- Assume that the patron will not know the your affiliations, and therefore do not tempt patrons to use a service they would regret if they new more about you.
- Respect your patrons. The opinion that users are "rocks with arms", as said by a colleague years ago, will not lead you to become a very helpful digital librarian.

In conclusion, the rewards from being a digital librarian are numerous and can be evident from notes from users from remote countries and companies. This electronic publishing revolution allows anyone with a personal computer and a modem to be a publisher will have far reaching effects on the structure of our society. Being a good digital librarian is a concrete way to create a future we all want to live in.