

Meeting Scholarly Information Needs in an Automated Environment: A Humanist's Perspective

David Crawford

Probably most humanistic academicians today use word processing, but this does not assure that we will quickly move on to higher forms of automation. Humanists who care about language can be repulsed by awkward, poorly organized, and jargon-ridden discourse and documentation; they will also expect automated access to information to seem as simple as earlier methods of inquiry. Nevertheless, many private databases—surely in the thousands—in the humanities are now being developed. The most pressing need confronting those scholars is to create a means of communication that effectively transcends disciplinary boundaries.



any of the current questions about evolving technologies and future library services emerge because technological progress can bring with it a complex of nontechnological problems. Perhaps the proper task for a humanist here is to focus on one of those inconvenient nontechnological problems—some of the people who will use the machines. As you adopt more technology, you will encounter the need to introduce users to the new processes. Your role as curators of knowledge seems to be taking on a new dimension. Now part of that knowledge, rightly or wrongly, is technology itself, and you will find yourselves teaching others how to use it. Effective teaching includes considering the attitudes and skills of those who must learn, so my first step will be to discuss some humanists' perceptions of evolving technologies. Then I will illus-

trate some of their applications of automation and, finally, offer some of my views regarding technologically enhanced library services.

Humanistic academicians easily recall the euphoric literature of the 1960s. Our thinking of that decade evidently oversimplified the degree of intelligence that lurked behind humanistic expression. Contrary to some expectations at that time, such operations as translating languages, writing good poetry, or composing symphonies in the style of Beethoven have continued to confound the programmer's art. In the course of more modest triumphs we have also seen colleagues suffer through some ambitious and expensive failures. As a consequence, many of us have become more cautious over the past twenty years. Before investing energy and resources into an evolving technology, we want to be convinced that

David Crawford is a professor in the School of Music at the University of Michigan, Ann Arbor, Michigan 48109-2085. This article was presented in its original form at a seminar sponsored by the Research Libraries Group in November 1985 at Conoco's Purple Sage Ranch in Bandera, Texas, through the hospitality of Conoco Inc.

the investment will be repaid, preferably with interest.

Surely a large majority of humanistic professors now use a computer for word processing, and we all know of some whose computer applications are truly impressive. But of those who work with computers, only a few, perhaps 20 percent, use their machines for anything more than typewriters. How quickly, or slowly, academic humanists will respond to the potentials of other new tools is difficult to predict. Answers will be related to their perceptions of the new processes. How much time must we invest in order to learn them? Will they really make our work better or more efficient?

We want to think that successes in word processing will encourage humanists to explore other technological applications. However, even word processing can lead to mixed perceptions because, to many humanists, the coin of the realm is language. Humanists who are drawn to word processing because it facilitates their use of language are also the people likely to be offended by the abused language they encounter. After plowing through some manuals on hardware, the documentation of software, and dialogues with vendors and other advisors, a humanist can logically conclude that:

[What is needed is] a kind of discourse accessible to many humanists that combines the authority of tone and clarity of expression that will catch his or her imagination. To a certain kind of humanist, jargon of any kind is a signal that the writer does not command sufficient precision of expression and depth of understanding to go beyond that rather pat phrase—or other phrases that seem indifferent to shades of meaning and implication in language, reducing expression instead to pre-packaged formulas. . . .

I sense some humanists shrinking from technology because, at least in part, they don't trust the sensibilities of those who are urging them to jump in the fray. Now of course, this is not necessarily the cheerleader's fault. Many humanists are timid by nature when it comes to technology, intimidated by gadgets, afraid of looking silly, not ready to endure the ego-buffeting that taking ten steps backward can bring, and perhaps a bit snobbish about old-fashioned ways. But their hesitancy, I think, has a core of common sense about it, too. It's

also a challenge to the technologists to assimilate their skills and to translate them convincingly into an acceptable form of humanistic discourse. The appearance of computer jargon shows, to some people, that the message has not really been assimilated and the common ground does not yet exist.¹

To put it another way, a humanist who interfaces with cybernetic processing is accessing a miniworld containing parameters of linguistic inelegance that possess the undeniable potential of arousing various negative responses ranging from a certain degree of suspicion to outright anger.

Humanists might be addressing this issue more effectively than we now do. Paradoxically, disorganized and poorly written documentation is commonplace at a time when universities graduate scores of trained and talented writers unable to find employment. Some restructuring of curricula for composition majors might help supply better technical writing. One of the issues you face now is the idea of integrating the library and the computing center, and I urge that writers of professional quality be added to the picture. One of the positive by-products of new technologies is that they can encourage us to reconsider the institutional barriers that separate minds from each other.

DATABASES MADE BY INDIVIDUALS

For some scholars, the thrust of humanistic research has been changing over the past two decades. Interest in heroes and geniuses is still in place, but now we are also asking about the daily lives of more common people. Consequently, information needs now differ from what they were a generation ago. This change brings us closer to the social sciences, quantitative methodology, and the computer. If computers have disappointed some humanists because of what they were unable to do, we, like librarians, have also learned some important lessons about what computers can do well. Although only a minority of us is active in building databases, such projects are now rapidly increasing in number. Database management opens the doors for many different

types of humanistic inquiry, including the ability to analyze the happenings within a full community or even a civilization. Some humanists' databases are bibliographic projects to assist in searches; others are conceived to analyze the correlations between variables in the data. Following are three illustrations of databases being developed now by individual musicologists:

1. I am making what appears at first to be a bibliographic database for searches, *Renaissance Liturgical Imprint: A Census (RELICS)*. My initial purpose was to locate the sources of the raw materials (chants and texts) that composers used in sacred music. RELICS now contains about twenty-four hundred and fifty records and will probably grow to about seventy-five hundred. Each title is described by up to twenty-one fields so that I can search for generic titles, ranges of publication dates, printers, cities where printed, countries of intended users, dioceses, rites, and monastic orders. Boolean operators permit extracting any combination of these variables. RELICS also informs me where to find a given book locally or in other libraries. Future work, in addition to adding more titles, should enter the usual bibliographic citations, formal physical descriptions, and descriptions of manuscript addenda.

Although conceived as a tool for bibliographic searches, when RELICS's sampling more fully represents the subject, the database can provide useful correlations between data. For example, from what foreign printers did the clergy of Spain, England, or Yugoslavia get their books? How do individual printers or cities compare in book production? What are the chronological fluctuations? Which dioceses or monastic orders took kindly to the early printed book, and which did not? Once number-crunching of this kind is done, we may have clues that lead to more secure interpretations.

2. My colleague, Dale Monson, is building an Eighteenth-Century Opera Database (ECOD) from information in printed opera libretti (booklets of opera texts) of the period. ECOD now includes about twelve hundred *libretti* and could grow to

as many as sixty thousand. Although Monson can produce a bibliography of libretti, his real purpose is to reshuffle the data. ECOD's structure is complex. Each libretto is described by over one hundred fields, and this would normally impede data entry. Monson solved the problem by writing a program for a microcomputer. The data are entered at the micro, which provides friendly prompts, and restructures the data so they can be telephoned to ECOD's big appetite. ECOD's power is illustrated by its ability to answer questions such as the following:

What is the performance history of a given theater? Were certain singers often cast together? When, where, and in which roles? What set designers worked in Venice? When and for which operas? When and where did the known performances of a given opera occur? Where can I now find all known libretti printed in Naples? When and where did performances of libretti by Metastasio take place?

3. A different kind of database is being developed by John W. Hill at the University of Illinois. Turning his attention to compositions from 1600 to 1800 that warrant a modern edition, Hill encodes the music in the PLATO system. The music can then be displayed on the screen, printed by a good-quality dot matrix printer, or heard through a synthesizer. If users wish to perform the music, individual performers' parts are printed automatically. Unlike many databases that permit the storage and retrieval of information in alphanumeric code, this output is graphics or sound. Distribution is available on PLATO terminals linked to the University of Illinois Computing Center. Communications costs are expected to fall within a few years, because PLATO is installing microwave satellite communications that will connect to existing cable television systems. Hill's project offers some significant advantages over more conventional means of music publishing: (1) purchasing prices may be reduced, (2) no library shelf space is encumbered, (3) the music can be heard as well as seen, (4) error-free performers' parts are available, and (5) delivery time is a matter of seconds rather than weeks or months.

With these three examples in mind, we

can now turn to the issue of standardization. Monson's project employs a locally developed utility system on one of our campus mainframes,² but his data are easily rolled over into formats acceptable to other systems such as dBase III. Therefore the information can be processed on some personal computers as well as numerous mainframes connected to a telephone. RELICS is driven by SPIRES, a mainframe system widely used for bibliographic projects.

Monson and I are using orthodox methods to answer orthodox questions, but Hill's project requires users to employ unique hardware in order to enjoy unique services. The problem of standardization is to try to strike a balance between what might appear to be two forms of intellectual censorship. In the case of Hill's project, the freedom to invoke atypical methods is justified because it delivers unique and useful services. On the other hand, many other technological incompatibilities are unnecessary; my inability to read an incompatible database impairs my freedom of inquiry. If we are to succeed in distributing information to humanists electronically, unnecessary inconsistencies must be discouraged.

These three illustrations have not been formally announced in musicological literature, so they bring to mind another problem that might be solved in a more specific manner. No one can guess how many other private databases in humanities are now being built in the country today, but the total is surely in the thousands. Even though the journal *Computers and the Humanities* is helpful, there exists today no truly effective means whereby humanists can inform each other of their work with databases. Even the communication within a discipline can be faulty,³ and communication between members of different disciplines is more problematic. How can I make RELICS compatible with the research of someone else working on early printed books? How can Monson synchronize his efforts with someone working on Italian literature?

The most pressing need involving humanists and technology today is to develop an effective flow of information be-

tween scholars. Such traditional media as the journal and the monograph are declining in esteem for several reasons. First, these forms of publishing involve time delays that technology has now rendered unnecessary. Such delays are particularly troubling because research seems to be executed faster now.

Thanks to word processing, scholars spend more time in research and less in clerical toil. One fast-moving piece of research ideally influences someone else's work, but that ideal can be thwarted by the bottleneck at the publisher. Secondly, database management is changing some perceptions about publishing. We used to embark upon a project and then eventually reach the fateful day when we had to declare the work finished and ready to be frozen forever on the printed page. Now, however, a database, even one that offers a fragmentary sampling of its subject, can be distributed online or on a print-to-order basis, thereby aiding other scholars. More complete databases can continue to enjoy refinements or expansions into related subjects. Many of us regard various types of humanistic research as ongoing processes and attitudes that resist the artificial categorizations of "in process" or "completed." The new medium, electronic distribution, can help us transmit information in accordance with that attitude.

Humanists need an umbrella organization, perhaps supported by NEH or ACLS, and it might be implemented through the information networks being put in place by RLG. Something like a national database of databases, distributed through RLIN, would be a helpful first step.⁴ If this could be developed, several further possibilities emerge.

1. Scholars who happen to be building compatible databases will know of each other's existence and can exchange data on disk or through telecommunications. Scholars with incompatible systems can at least exchange hard copy.

2. Needs might be defined to locate facilities for translating data from one format to another.

3. We will be in a more knowledgeable position to consider standardizing data formats and search protocols.

4. If a degree of standardization can be achieved, some scholars might be willing to release their databases to a central depository where they can be distributed as public domain electronic data. ECOD and RELICS, for example, will eventually be valuable enough to warrant online distribution. This could be done at our computing center, but many users across the country are accustomed to the protocols and accounting methods at other installations. Such complications should be unnecessary; we have returned to the call for standardization.

PUBLIC DATABASES AND SERVICES

The emerging online catalog will alter the work patterns of scholars who have electronic access from remote sites. Such users will welcome the full range of services that can be provided: learning the actual location, charge status, or acquisition status of a book; putting a hold on books charged to other people; linking to catalogs at other institutions; filing interlibrary loan requests; and initiating the paging of books from distant storage locations. Also, the physical delivery of volumes to someone's remote location, a service being developed in some libraries, is conceptually related to an ability to request the item while at a remote site.

But not all is rosy, because of user-oriented issues that jeopardize the returns on your frightfully large investment. For many humanists, the online catalog will be the first hands-on experience with telecommunications. They will not want to spend time studying documentation, and they will not want to learn new protocols when their screen lands them in a different library's catalog or when they travel to another library. In other words, they will want to be convinced that the online catalog represents an improvement over the well-known and standardized card catalog. A consortium of research libraries developing online catalogs is in a good position to insist upon friendly programs; to develop clear, well-written documentation (or, better yet, menu-driven programs); and to promote standardization protocols. These are prerequisites for

comforting the scholar who is devoted to subjects other than computing for its own sake. Our present failures to achieve these goals are understandable consequences of technology's recent history. Future failures, however, may reflect a shortfall in thinking and organization, not in technology. I fear that the more we automate, the more expensive the process of standardization will be, whenever it finally does emerge.

When turning to another kind of public database, the discipline-oriented bibliographic one, surely it is unnecessary to repeat the plea for standardized protocols with these projects. As we move more and more toward putting the user in the hands-on position, some faculty and students may express a need for formal training or counseling. If such services seem necessary, they may be treating the symptom rather than the disease. A humanist, no matter how uninitiated, is likely to believe that the logical operations involved in reading a bibliographic database should appear simple, resembling those required for reading a card catalog. If our tools obscure that simplicity, the design of the tools should be reconsidered.

Another cautionary note needs to be kept in mind with regard to the discipline-oriented bibliographic database: not all disciplines are similarly well served by the databases that presumably support them.⁵ Books, editions, and journals are dutifully reviewed in a discipline's literature, but public databases seem to be subjected to less critical review, if any at all. If a given discipline neglects to review its bibliographic databases, perhaps such evaluations could be undertaken jointly by librarians and individual representatives of the discipline in question. Qualitative evaluation is a daily problem in book selection departments, and we can expect it to extend to databases as well. The issue may become more acute as these databases become more integrated; a user familiar with one discipline may be more frequently vulnerable to assumptions about the integrity of databases in other disciplines.

Many of us regard bibliographic databases as providing screen displays or hard

copy, but we can expect a growing interest in writing the output into mainframe files or on floppy disks. This interest will be stimulated by the increasing access from remote stations, because scholars will wish to build their own bibliographies more efficiently and accurately. RELICS, for example, would have benefitted from the data already machine-readable in such databases as RLIN and OCLC.

Although the relationship between information needs and library services largely involves database resources, some attention might also be directed toward electronic mail. Academicians often communicate electronically and so do librarians, but I wonder how often the two parties talk to each other through their modems. The need to appear physically at a librarian's office for routine dialogue results in undue congestion; it is inconvenient, and it is often badly timed for at least one of the parties involved. Some users would appreciate being able to use modems to consult reference and circulation desks as well as officers such as those in charge of book selection, cataloging, and public databases. Also, the ability of library staff to handle routine exchanges with users during off-peak hours may contribute to using staff more efficiently.

Promising though local electronic mail is, the possibilities increase when we add existing national and international net-

works to our thinking. While we wait for the expansion and refinement of online catalogs, we should keep in mind that it is economically feasible for me to be online to virtually any librarian in the country. The international dimension is especially valuable to humanists, for many of our important sources are preserved abroad. Not only would this service be useful to many humanists, it would also provide a simple and instantly rewarding exercise in the world of evolving technology.

CONCLUSION

Evolving technology holds great promise for helping meet the information needs of humanists, and in only a few exceptional instances, if any, can we believe that its potentials are fully realized. In many cases, the technological advances of recent years have outstripped our ability to install comfortable handshakes between the machine and the user. Research libraries, being drawn inevitably into technology, need users who can invoke the new tools. People will accept these changes more gracefully if they are greeted by friendly programs, lucid documentation, standardized protocols, and compatible databases. We will want to see that the former logical and simple methods of locating information are replaced by new tools that appear equally logical and simple.

REFERENCES AND NOTES

1. I quote a letter of October 9, 1985, from my colleague, Profesor Richard Crawford, who was responding to a preliminary version of this paper.
2. The TAXIR system was developed by Bob Brill of the U/M Computing Center. We gratefully acknowledge the center's assistance with hardware and software and also the graduate research assistantships from the School of Music.
3. For example, in 1982 some members of a national learned society in humanities requested the formation of a committee to serve as a clearinghouse for information about computer-assisted research. Although the proposal was simple and almost without cost, the board of directors denied it after two years of discussion.
4. What is suggested here at the national level has already been discussed at a local level. See the promising *Final Report* (April 4, 1985) of Columbia University's Data Base Policy Committee.
5. For example, I recently asked my colleagues in Michigan's musicology department about the online database in music. Of the ten of us, two had seen it demonstrated, but no one had actually used it for his or her work, and no one had advised students to consult it. Since this database now has a performance record of more than a dozen years, its promise for the future seems cloudy.