Machines and Minds: Historians and the Emerging Collaboration

Chad Gaffield*

Theoretical debates about the appropriateness of computer-assisted research for historians have now lost relevance. While some researchers have remained unconvinced since the 1960s and others, early-converted, have now had crises of faith, the computerization of research and writing is no longer in question. Word processing has proven to be the point of entry (and often the extent of application) for many current users, but substantive computer-based analysis has now reached most departments of history. The creation and examination of machine-readable data have encompassed scholars of both words and numbers and, while controversy continues over modes of analysis, the utility of computers has gained a consensus of acceptance if not general enthusiasm.

Nonetheless, computerization has proceeded far more slowly than anticipated by optimistic promoters in the 1960s. The cultural cleavage between humanists and scientists remains quite evident in the arts, where many scholars, often harbouring grim memories of high school mathematics, still recoil from the binary underpinnings of even logging on. The transition from index cards, foolscap and fountain pens to machine-readable files, printouts, and never-as-compatible-or-portable-as-promised equipment has proven to be both emotionally and technically less than smooth.

While many factors explain the reluctance of historians to enter the Information Age enthusiastically, a rarely discussed phenomenon involves a general misconception about what computers can do for historians. The misconception is that computers can simply help historians do better (or to a greater extent) what they already do. In other words, computers are a tool by which historians can handle larger amounts of data, undertake more complex analyses, and write up their own research reports more efficiently. Computers are, therefore, non-revolutionary. They represent for historians basic continuity rather than substantive change.

This emphasis on computers as tools much like earlier technological developments such as microfilm readers and xerox machines may be reassuring to some historians, but it may also be very misleading. Most importantly, this notion may deny for us the truly exciting future which computerization is bringing to other fields. I am referring here to questions of disciplinary organization, epistemology and, ultimately, creativity. My argument is that computerization is offering historians not only a tool to do better what we have always done, but also the opportunity to redefine and reconstruct the process of historical inquiry. The character and dimensions of this opportunity remain very unclear, and all

---

* Professor of History, University of Ottawa.

A draft version of this paper was presented for the session, "Les bases de données historiques : l'expérience canadienne", at the annual meeting of the Canadian Historical Association, June 1988.

Histoire sociale — Social History, Vol. XXI, nº 42 (novembre-November 1988) : 312-317

predictions will inevitably miss the mark. Nonetheless, it seems worthwhile to explore the possibility that computerization is fueling a research paradigm shift of truly revolutionary proportions. Significantly, this possibility is far removed from the now outmoded quantitative-qualitative, humanist-social scientist controversies of the 1960s and 1970s. In fact, the current terms of debate do not even allow for such dichotomous conceptualizations.

This perspective on the relationship between historians and computers emerges from what began as a quite non-revolutionary (though innovative) undertaking: the creation of a machine-readable listing of archival sources held in repositories on Vancouver Island.[1] Our ambition (that of Alan Artibise, Peter Baskerville, and myself) was to write a comprehensive regional history. As one of our first steps, we sought to identifiy the available sources. This stage was viewed to be exceedingly important in the case of Vancouver Island, since a survey of bibliographies revealed that the secondary literature was quite limited. Moreover, our interest in social and economic history would necessitate, we knew, a substantial amount of new research on routinely-generated sources which, theretofore, had been neglected. Inspired by an attachment to systematic history, we were reluctant to follow the conventional strategy for identifying relevant historical sources which we viewed as "unhurried burrowing" in the tradition of "romantic research".[2] In this traditional strategy, researchers poke around in finding aids and inventories, follow serendipity into various record groups and, most important of all, make friends with the appropriate archivists. We did not reject this strategy, but rather deemed it incomplete, since the type of evidence we sought would often not have finding aids or even be in the custody of archivists. The important individual-level data of municipalities, churches and other institutions could simply not be discovered in the traditional ways. In addition, we sought the admittedly unattainable goal of systematic identification of all the relevant sources, just as we strove for the ever elusive achievement of clearly articulated concepts and rigorous methodologies in our own research.

In this way, we created somewhat unconsciously a small but complex world of archivists, computer system analysts and programmers, librarians, historians, and an array of other specialists. Slowly, we came to see the possiblities of an integrated research process characterized by a dynamic continuum beginning with the creation of records and extending to their ultimate use when they become redefined as historical sources. In this research paradigm, archival information is the "life-blood" of a system which links the producers, managers and users of that information. Since the system depends upon circulation (access) of that information, computerization becomes more than heuristic; it begins to take on a life of its own.

The small step which the Vancouver Island Project took as we groped toward this new vision of the research process involved the creation of a prototype "automated archivist".[3] Our aim was to provide researchers with a computer system that could be interrogated concerning the contents of certain archives. The first challenge was to develop a

1. Peter A. Baskerville and Chad Gaffield, "The Vancouver Island Project: Historical Research and Archival Practice", *Archivaria*, 17, 1983-1984, pp. 173-187.

2. Peter A. Baskerville and Chad Gaffield, "Shifting Paradigms and Emergent Technologies", Baskerville and Gaffield, eds., *Archives, Automation, and Access*, Victoria, 1985.

3. Chad Gaffield and Peter A. Baskerville, "The Automated Archivist: Interdisciplinarity and the Process of Historical Research", *Social Science History*, 9, Spring 1985, pp. 167-184.

method of data collection for the holdings of Vancouver Island's repositories. The controversial aspect of this work involved our determination to provide researchers with systematic subject access. Here, we found ourselves in the midst of heated debate among archivists about the meaning of provenance, intellectual control and administrative control. It was at this point that we saw the very negative implications of the isolation (and considerable alienation) that characterized the relationships among the professional groups dealing with information. Each of these groups was computerizing, but in ways that precluded integration. Computers were being used as a tool to do better what each group had always done. The result for historians was not helpful. Instead of facilitating access to information, computerization was raising to new heights the traditional walls separating researchers, librarians and archivists. We were particularly concerned about the development of systems such as SAPHIR at the Archives nationales du Québec. Elaborate, extensive and expensive, SAPHIR (in its initial form) was designed by archivists *for archivists* allowing them to gain better administrative control of their holdings throughout Quebec. Similar projects elsewhere convinced us that the true potential of computerization was not being fulfilled. The possibility that automation could encourage access, integration and systematic research was not being pursued.

Eventually, we developed a comprehensive survey form which reflected developments among archivists in the United States involved in the National Information Systems Task Force, and among librarians who were building on the Anglo-American Cataloguing Rules 2. This survey form includes a series of descriptive fields of information related to holdings at both the record group and series level. Our field definitions and data entry system structured the creation of approximately 5100 CMS files.[4]

The next challenge was to develop a way to provide access to any user. Our vision called for a system whereby an individual, on site or remotely, could easily discover what historical evidence existed for whatever purpose was in mind. Toward this end (though certainly not achieving our full ambition), the Vancouver Island Project System (VIPS) was developed as a menu-driven main frame system which provides structured access to the data by way of a Structures Query Language (SQL) database. The basic conceptual design is predicated on two distinct databases: the primary database being a collection of computer files which contain the data amassed by the project regarding the holdings of repositories on Vancouver Island; and the secondary database consisting of a series of indexes of the files in the primary database.[5]

VIPS assumes that the user wants to determine information (contained in the CMS files) about certain records which are described by the index attributes (stored in the SQL database). When the user accesses the system, a query is formulated in CMS which is sent to the index database to determine which files meet the query criteria. A list of the files is then transmitted back to the user in CMS, where he or she may determine what is to be done with the list (e.g. save, erase, print, etc.). A menu-driven generator assists the user in formulating a query. This formulation is based on standard boolean logic (AND, OR, NOT).

---

4. Chad Gaffield, Peter Baskerville, Catherine Panter, and Pauline Shepherd, *Field Definitions and Data Entry Guide for Archival Material*, Victoria, VIP, 1985.

5. *Vancouver Island Project System: Project Design Manual*, Victoria, Computing and Systems Services, University of Victoria, 1987; *VIPS: User Guide*, Victoria, Computing and Systems Services, University of Victoria, 1987.

In developing this sytem, we found ourselves reading journals we had never even heard of, learning vocabularies of the "foreign" languages of other disciplines, and thinking in ways far removed from our own training as historians. This project became the result of a collaborative, interdisciplinary effort in which we relied on the expertise of specialists. It was essential, however, that this reliance was not blind; we had to learn enough to communicate effectively. And it was in exploring the debates of other groups that we began to re-think the established intellectual structure and institutional organization of historical research. Increasingly, our conclusion was that computerization offers an opportunity not simply to do better the old ways, but to bring together the usually disparate elements and participants in the research process. By collaboratively developing single systems which meet quite different needs (such as those of administrative and intellectual control in archives), we would be encouraging a more dynamic research process in which the continuum along which information should flow would be uninterrupted.

But what are the real implications of this perspective? How revolutionary, in fact, can computerization be for historical research? One promising answer to these questions is what Ian Winchester and Jan Sundin have termed "intelligent databases".[6] In this view, a computer system would be developed as "a combination of historical archivist, systems analyst and programmer, cohistorian and friend." Winchester identifies five activities which would characterize the desired intelligent database.

1. An intelligent database would be able to converse with a researcher in a natural language concerning all of the plausible types of analysis in which it might be expected to be involved.
2. An intelligent database would have to take hints and suggestions from a researcher and offer plausible interpretations of these in terms of data lists, tabulations, or analyses.
3. An intelligent database would be capable of re-programming itself to perform new analytic tasks as required by the researcher.
4. An intelligent database would offer interpretations of the data it has already processed, and engage in discussion with the researcher about the interpretation of the data.
5. An intelligent database would actively aid in the co-authorship of papers.

Obviously, VIPS is a far cry from being an intelligent database as defined by these criteria. But the thinking behind the project is indeed consistent with this collaborative vision of the meaning of computers for historians.

The key conclusion involves the importance for historians of participation in the development of computer applications. Thus far, historians have tended to ask the question: "What software and hardware exist which we can use?" The answer for most historians (though certainly not for the other participants in this Round Table) has generally been limited to Database Management Systems which often do not like text, statistical packages for rectangular data files and word processing. Perhaps, the better question is: "What do we want to do?" Revealingly, it is this question which has been at the core of one of the most significant contributions of historical projects to computerization, the advances in nominal

6. Jan Sundin and Ian Winchester, "Towards Intelligent Databases: Or the Database as Historical Archivist", *Archivaria*, 14, Summer 1982.

record linkage. Historical researchers in Canada have developed strikingly original concepts and methods for systematic re-identification in wide-ranging documents. For the past twenty years, they have been at the forefront of an international scholarly debate which concerns many disciplines. Two essential points are that research on record linkage began with an internally-derived agenda, and that the most effective computer systems involve semi-automated record linkage in which humans and computers work together to re-identify individuals. In this example, historians turned to the computer with a new purpose in mind, with a hope to open doors which had previously remained shut; collaboratively, substantial success has been achieved.

It is this spirit of inquiry and active participation which is crucial to the continuing development of computerization for historians. Perhaps, as a priority, we should engage in current debates and research such as those related to expert systems, chaos theory, and creativity.[7] In each of these fields, scholars are pursuing the use of computers not simply as tools, but more importantly, as co-workers. The most difficult pre-condition to such work is an acceptance that basic concepts such as "intelligence", "order" and "explanation" are now undergoing complete reconsideration in many disciplines of the arts *and* sciences.[8] Such reconsideration raises fascinating possiblities for the character of historical research. Should historians, for example, continue to limit their data analysis to the hypothesis-testing model imported from now-outdated social science strategies? Should the examination of evidence be structured only within the realm of possible analyses proposed by particular human researchers? Or can semi-automated systems be developed in which underlying patterns (perhaps lurking within the apparent "chaos" of specifc sets of historical data) can be systematically identified, and thus exposed to further analysis? In a world of such systems (along with "automated archivists" and word processing systems which edit, revise and correct text), the distinction between researchers and their "equipment" would be significantly blurred.

The emerging collaboration of machines and minds is now evident in diverse fields of scholarly work. The once-considered "hard" sciences are softening in the face of unanticipated complexity; scholars believe less and less in the possibility of a systematic march toward complete explanation of scientific phenomena. At the same time, the so-called "soft" sciences have been re-thinking the assumed implications of humanistic attachment; certain poets, for example, are now using computer programmes to generate "raw poetic material".[8] Moreover, a great deal of attention is now being given to the epistemological similarities and differences between history and sciences such as biology and

7. As would be expected, controversy swirls around all these topics. In *The Fifth Generation*, Reading, Mass., Addison-Wesley, 1983, Edward Feigenbaum and Pamela McCorduck promote the potential of artificial intelligence in all fields; on the other side, scholars such as Joseph Weizenbaum see an unalterable distinction between humans and computers; *see*, for example, his review of *The Fifth Generation* in the *New York Review of Books*, vol. 30, n° 16, 27 October 1983. On chaos, the most accessible description is James Gleick's, *Chaos: Making a New Science*, New York, Viking, 1987. Examples of recent work related to creativity are provided in the special issue of *Interchange*, vol. 16, n° 1, 1985, entitled "Creativity, Education, and Thought."

8. Even the questionable field of "intelligence" testing is undergoing re-evaluation based on recognition of far great complexity in mental processes than previously believed. As a point of entry, *see* Robert J. Sternberg, *The Triarchic Mind: A New Theory of Human Intelligence*, New York, Penguin, 1989.

8. For a summary of recent work (with examples), *see* A.K. Dewdney, "A Potpourri of Programmed Prose and Prosody", *Scientific American*, June 1989.

physics.[10] Disciplinary boundaries are no longer simply assumed. In many fields, computerization is playing a central role in this reconsideration as a result of the active particpation of scholars in the research and development of applications. Inevitably, results come much more slowly than anticipated, but achievements, thus far, do indeed suggest the possibilities of a more general collaboration of machines and minds as part of a truly revolutionary paradigm shift. In principle, there seems to be no particular reason why historians should not also seek such collaboration in the continuing pursuit of a better understanding of the past.

10. For example, Daniel Scott Smith analysed comparatively history and biology in his presidential address to the Social Science History Association, Chicago, 1988, while scholars such as Ian Winchester and Hubert Watelet have explored the respective characters of history and physics; *see* Winchester, "History, Scientific History, and Physics", *Historical Methods*, 17, n° 3, Summer 1984; Watelet, "Réflexions sur le rapport de l'histoire et de la physique avec le réel", Paper presented at the Centre de recherche en études québécoises, Université du Québec à Trois-Rivières, 26 March 1988, and published by La Société de philosophie de l'Outaouais, *Carrefour*, vol. 8, n° 2 et vol. 9, n° 1, 1988, pp. 141-154.