

## *Notes and Communications*

### *Archives, Automation, and Access: The Vancouver Island Project Revisited*

by **TERRY COOK**

One of the most interesting recent developments on the Canadian archival scene is the Vancouver Island Project.<sup>1</sup> It addresses directly or reflects such thorny issues as networking, provenance *versus* content or subject indexing, the role of automation in archival work, the nature of the “new” historians and the “new” archivists as both professions mature, the recent revival of concern for local records, the impact of the SSHRCC Research Tools Programme that funds this project (and many others), and the implications of such “independent” undertakings as the Vancouver Island Project *vis-à-vis* “established” provincial or other repositories. The principal investigators of the VIP, historians Peter Baskerville and Chad Gaffield of the Department of History of the University of Victoria, have not rushed pell-mell into their project, but rather have read widely in archival, library, and historical literature and methodology and sought through numerous articles to engage archivists in particular in a theoretical debate concerning the assumptions and implications of the Vancouver Island Project.<sup>2</sup> Their most recent endeavour in this regard was sponsoring “Archives, Automation and Access: An Interdisciplinary Conference for Archivists, Librarians, Records Managers, Information Scientists, and Researchers in the Humanities and Social Sciences,” which was held at the University of Victoria on 1-2 March 1985 and attracted about 140 registrants. Although many of the papers will eventually be published separately, a summary of the conference may be of interest here.

---

1 See Peter A. Baskerville and Chad M. Gaffield, “The Vancouver Island Project: Historical Research and Archival Practice,” *Archivaria* 17 (Winter 1983-84), pp. 173-87. The scope and range of the VIP are outlined there, and so not repeated here.

2 *Ibid.*, p. 187; plus Baskerville and Gaffield’s letter to the editor in *Archivaria* 19 (Winter 1984-85), p. 7. See also their “The Crisis in Urban Documentation: ‘The Shame of the Cities’ Revisited,” *Urban History Review* 13 (1984); and “The Automated Archivist, ‘Interdisciplinarity’ and the Process of Historical Research,” *Social Science History*, forthcoming.

Too often conferences, articles, or books on automation and archives, on big databases as the VIP obviously is, degenerate into the nitty-gritty of software and hardware comparisons, or worse into preachy pronouncements that archivists must wake up and smell the roses, for the age of the computer is HERE, and we must do SOMETHING ABOUT IT! Yet computer or other technology solves nothing that is ill-founded conceptually; too often the system analysts come in to computerize a manual descriptive system that never made any or very much sense. The result is the computer truism: garbage in equals garbage out. As Nancy Sahli aptly remarked in 1981:

In a sense it is the same situation that a carpenter would encounter in trying to construct a house, with laser-beam drills, automatic plumb-line adjusters, synchronized electrical wiring articulars, but, alas, no architectural plans ... it would lead to the same skewed, wasteful results.<sup>3</sup>

David Bearman has likewise noted that the automation issues facing the landmark National Information Systems Task Force and the archival profession generally “are not essentially technical. They are, however, of considerable complexity and will require archivists to think carefully about the premises of archival practice and the underlying tenets of archival theory.”<sup>4</sup>

In opening the conference proceedings with a stimulating keynote address which had a laudable focus on concepts, principles, theories, and paradigms, Baskerville and Gaffield obviously support this approach. They argued that a “conceptual confusion” is both obscuring a needed transformation in the way archivists and users are beginning to think and act, and thus retarding the new descriptive ideals that should animate this new information age. Until this conceptual confusion is cleared away, until, as they say at another point, the basic professional “priorities are examined in a systematic way,” debates about software, computer equipment, and system development must be rather uncritical and, in their words, “less important than is implied in the literature.”

Part of this confusion, in their view, rests with a general failure to appreciate the “shifting paradigms ... in the modern research world,” to cite their paper’s title. In the first place, the old pattern, the crumbling paradigm, was one of archivists primarily serving the creators or producers of records and benignly allowing users, following the Great Detective syndrome, to fend for themselves, which they were pleased enough to do. As noted Canadian historian Chris Armstrong joked — but only half in jest at last year’s ACA conference — who wants to read a four-hundred-page inventory which seems mainly written for archivists; just get me the records or allow me to poke around for myself!

Under this old paradigm, provenance indexing within archives, with its institutional and hierarchical emphases, was welcomed by the producers of records, given their natural and narrow institutional bent, while most researchers of the old paradigm did not care one way or the other. Indeed, most archival guides or inventories were ignored by researchers — notice Chris Armstrong again — who preferred to browse to find great

---

3 Nancy Sahli, “Finding Aids: A Multi-Media, Systems Perspective,” *The American Archivist* 44 (Winter 1981), p. 20.

4 David Bearman, “Towards National Information Systems for Archives and Manuscript Repositories,” *The American Archivist* 45 (Winter 1982), pp. 55-56.

treasures, using finding aids and inventories that adhered to the provenance tradition only to “infer” where such treasures might be, or directly only if they were doing old-fashioned institutional or biographical studies. Provenance in short served creators of records and archivists; it was an organizing or arrangement tool and a descriptive concept, but hardly an effective or often employed means of retrieval for most users.

The third stage of the author’s argument is to posit a revolution among all members of the information continuum or triumvirate; there are new actors on the stage. In the first place, historians and other users no longer browse in a leisurely way; their interests have broadened in content and methodology far past the clues available in current inferential provenance-based reference tools; they are interested in new records described in new ways; their work is interdisciplinary and collaborative, and searches out the “ordinary, routine, and everyday” in the authors’ words rather than the “exceptional, spectacular, and precedent setting” which, apparently, is the most telling blow against provenance-based archival guides and inventories. The producers or creators of records have also changed; they are no longer primarily chasing after the isolated fact or the clear reference — although that will always go on — as they are taking a broader view of their archives as the collective corporate memory of the government or corporation or union or university — the fertile font for precedents, trends, policy development, organizational studies, promotional activity, and so on. Producers, the authors assert, are accepting (or soon will be) the insights of the “public history” movement into their mental framework; history and archives are not old stuff kept about purely for cultural reasons, but a vital part of current, daily operations. From this perspective, the older, narrow, institutionally-based provenance approach of archives also no longer suits the new producer or creator. And archivists in the middle share this ferment: they are reaching out through NISTF in the United States, through the new Working Group on Descriptive Standards in Canada, and through many local and regional initiatives, of which the VIP is obviously one of the most advanced, to break down the old institutional and media barriers; they are actively pursuing the local and case file records that support the new “everyday and ordinary” research perspectives, including developing some very sophisticated statistically valid sampling strategies to cope with mountains of such records which before were routinely thrown away; and they are waking up in a scholarly way, as advocated by Tom Nesmith, myself, and others, to study the historical intricacies of the actual records in their care. Such concerns for the “history of the record” are seen by Baskerville and Gaffield, and some of the newer social historians they cite, as essential to the flowering of the new research and historical paradigm. Much of this activity or new perspective among producers, archivists, and researchers has happened in utter isolation of one from the other, and the authors preach (and practise) that much fruitful interchange and interaction among all three could occur if such isolation ended and better communication characterized the records continuum.

This broad conceptual sweep is not often recognized or acknowledged by archivists when discussing automation and researcher access. The dynamic imperatives of our age are such that archivists want to jump in feet first to embrace the new technology — to solve this small problem, to control that function, to index this medium, to sort that collection. Many such applications — laudable as they may be in their way — are ultimately piecemeal and do not have the breadth of vision advocated by the VIP investigators. They do not ask the large theoretical and conceptual questions raised or implied in Baskerville and Gaffield’s paper:

- does the alleged new paradigm of producers following the “public history” model *and* of users being interested in different subjects, approaches, and records require that “traditional” archival theory be changed?
- does provenance — and by extension the whole approach to archives that describes records first and foremost for their context, their functional links with their creator and creating activity, their interrelationships, and their evolution over time, rather than for their specific subject content — does this approach no longer serve the new researcher delineated by Baskerville and Gaffield? Or can it be expanded, adapted, and redefined to so serve?
- what are the implications — for scheduling, description, records analysis and understanding, current policy development, historical and other scholarship, and indeed automation proposals — behind Baskerville and Gaffield’s interesting insight that the well-known life cycle of *records* (active-dormant-archival) is matched by a *human* one of creators, archivists, and researchers who, through records management and public history at the front end, analysis and scholarly study of records per se in the middle, and research use in a myriad of ways at the end, are all organically and inextricably interconnected?
- should archivists (with producers and users) not approach records description and thus eventually archival automation in a more global and holistic fashion, as the VIP bears witness? Should not our “systems” try to transcend the narrow boundaries of collection, medium, and institution? Indeed, to fail to do so, as I have argued elsewhere, is to diminish or destroy that central key element of provenance on which so much knowledge needed by producers, archivists, and researchers ultimately rests.<sup>5</sup>
- will a new kind of itinerant archivist become a significant factor in our profession in the next two decades? The local school boards, churches, regional districts, and villages whose records are now well described in the VIP computer database could never afford a full — or even part-time archivist, but they could perhaps hire one for a day or two a month to arrange, further describe, and conserve these newly identified records, records which under the new social science research paradigm outlined by Baskerville and Gaffield will be attracting increasing attention from creators and researchers. Will roving bands of archivists, each responsible for perhaps ten institutions in a hundred-mile radius, be the wave of the future, the saviours who preserve the myriads of local, community records which, for whatever reasons, fall below the collecting activities of larger established archives? If so, what does this mean for archival education, for developing and maintaining acceptable archival standards, for the emerging network of the Canadian Archival System?
- how do the professional associations of archivists — provincial and national — respond to these initiatives? If the political realities of SSHRCC financing mean that such massive undertakings as the VIP will likely continue to be undertaken by universities and individual academics rather than by provincial archives, then should not the professional associations seize the very innovative work of the VIP, study and critique it as necessary, and issue standards for such work with local records and

---

5 See my “The Tyranny of the Medium: A Comment on ‘Total Archives,’ ” *Archivaria* 9 (Winter 1979-80), pp. 141-49; “Media Myopia,” *Archivaria* 12 (Summer 1981), pp. 146-57; and “Archival Networks and Congresses,” *Archivaria* 17 (Winter 1983-84), pp. 13-17.

automated description? Or must the wheel be reinvented — at great cost — each time one of these so-called megaprojects is launched? In this regard, as Terry Eastwood told the conference, the new Working Group on Descriptive Standards will offer some guidance, although its initial recommendations will probably be too general to answer the specific points raised above.<sup>6</sup>

- and what happens when these megaprojects wind down, as the VIP will this summer? Who is to take over, maintain, and administer — and one hopes update — the innovative, impressive database that has been created? The provincial archives is the obvious answer, but what if it cannot or will not? Needless to say, it would be a colossal waste of skill and effort, to say nothing of taxpayers' money, if upon completion such projects atrophy on a shelf somewhere.

No small number of the concerns raised by Baskerville and Gaffield in their paper, and by the implications of the VIP itself (on which there were workshops, demonstrations, and printouts available), were addressed by other speakers at the conference. David Bearman of the Smithsonian Institution and late Director of NISTF spoke of *the* central issue: “establishing intellectual control standards” for archival materials. The emerging computerized information systems in archives and the increasing information sharing among repositories will be hollow shells indeed if the controlled vocabularies governing intellectual access to archives are not legitimized. Indeed, far more seriously, archivists can rarely decide which vocabularies to describe what are even needed! In a quite revolutionary proposition, Bearman suggested that such intellectual control might not be based on a description of records per se — the archivist could put a paper bag over the records and ignore the contents — but rather on the mission, mandates, organization, functions, and history of the body which created the records. Undertaking such research and presenting the results in controlled vocabulary form would provide subject access points to modern archives which no amount of records description per se would do, especially in the burgeoning information world we live in. Bearman’s approach describes records by revealing how, when, where and why they were created, but from the top down, rather than attempting to sift through the miles of aisles of archival boxes at the bottom. It is subject indexing, but within the framework of provenance, even if a much wider definition of provenance than the older, narrower, purely hierarchical and structural model that Baskerville and Gaffield find wanting.<sup>7</sup> The computer printouts for the Vancouver Island Project indicate that Bearman’s suggestions have been followed (or anticipated), for the records of a local municipality for example are described in terms of

---

6 In this regard, attention should be drawn to the Vancouver Island Project’s publication, *Field Definitions and Data Entry Guide for Archival Material: A Working Document* (Victoria, 1985), which was jointly done by Baskerville, Gaffield, Catherine Panter, and Pauline Shepherd and which is available from the VIP. It proceeds from the assumption that “a rigorous system of field definitions is the first step in gaining the administrative and intellectual control necessary for automation in archives.” Using the insights of NISTF and work on the MARC format, the authors attempt to outline all the data fields, with a careful breakdown of each, complete with definitions and glossary, that are necessary to automate an archives. This is germane not just to the VIP or similar surveys of regional records or automated networks, but equally useful to gaining intellectual control over records generally and to setting up automated access systems to such records in any archives. The manual is neither hardware nor software specific. The volume should serve as a good discussion point or base as the archival community in Canada moves towards descriptive standards in a more concerted way.

7 This approach of Bearman’s is exactly what I have recently argued should animate archival work; see “From Information to Knowledge: An Intellectual Paradigm for Archives,” *Archivaria* 19 (Winter 1984-85), pp. 40-49.

the mandate, organization, functions, and areas of activity of that municipality. This gives many subject points of access, but is quite radically different from the classic content indexing model of opening up each file and indexing its contents by subject — a quite impossible task in today's archival world.<sup>8</sup>

Terry Eastwood, who directs the Master of Archival Studies Programme at UBC, outlined areas where descriptive standards are needed and which the Working Group which he heads is investigating: context, which is the classic provenance concern of archivists everywhere; composition, which involves the peculiarities of medium (map, photograph) and record type (journal, diary, letterbook); content, which goes beyond Bearman's modified provenance model where it will not suffice (photographs, for example, whose subject content may be unrelated to the photographer or his sponsor); and user needs and behaviour which, while raising the danger of the tail wagging the dog, cannot be ignored either. While the Working Group hopes to point out appropriate directions and processes to achieve standardized archival description *vis-à-vis* these four factors, its concrete accomplishment may take a generation. Hugo Stibbe of the National Map Collection of the Public Archives of Canada demonstrated how the pioneering work there combines discrete item indexing with series or group and medium or form description, all with controlled vocabulary, and thus how some of the classic theoretical distinctions can be rendered somewhat artificial.

If intellectual access to archives, with all its attendant theoretical challenges, was the focus of part of the conference and of the VIP itself, then the new records unearthed by the VIP in particular and this generation of social scientists generally was the other. Susan Falb, former history professor and NARA staff member and now involved with records appraisal and retention at the Federal Bureau of Investigation, offered a sophisticated treatment of statistically valid sampling methodologies for archivists. Voluminous case files contain that very everyday information about ordinary people that is essential to new social science research. Archivists often resist collecting such records, she found. Archivists' bias is institutional; they collect largely the records of institutions and arrange and describe them according to provenance to reflect those creating organizations. Most social scientists, however, are interested in people, not institutions per se. It is clear that the "keep it all" *versus* "pitch it all" dichotomies will no longer serve the "new" paradigm of producers, archivists, and users. Sampling is the intelligent middle ground, and Falb reviewed several recent cases where large-scale sampling projects had been carried out successfully. She noted, too, that David Bearman's "bag over the records" syndrome might not suffice for case file series, either for appraisal or description; such files sometimes contain significant local or personal information that might not fall under any reasonable definition of the creating agency's mission or mandate. Some investigation or spot-checking of the actual records would be needed as well.

Besides case files and similar local, individual records, the other "new" source for tomorrow's social scientist is obviously machine readable records, and Sue Gavrel of the Machine Readable Archives at the PAC addressed several of the issues that complicate user access to such records, including overcoming privacy restrictions, creating public use files, developing union lists, and deciding upon descriptive standards for this new medium.

<sup>8</sup> The classic exposition of provenance (P) *versus* content or subject indexing (CI) is Richard H. Lytle, "Intellectual Access to Archives: Provenance and Content Indexing Methods of Subject Retrieval," *The American Archivist* 43 (Winter 1980), pp. 64-75.

Other speakers, commentators, or workshop presenters at the conference looked at hardware and software configurations that might meet some of these challenges, outlined the significance of on-line databases in the era of microcomputers and telephone modems, sketched in the details of the administrative control system (SAPHIR) adopted by the Archives nationales in Quebec, or revealed the difficulties and possibilities of establishing word processing/microcomputer facilities in medium-sized repositories such as the Provincial Archives of British Columbia, and the kinds of routine archival tasks made both faster and better as a result.

Ted Dürr of the Baltimore Region Institutional Studies Centre reminded everyone in his paper "Zap Archives!" and in several discussion sessions that an affordable computer system can be in every but the very smallest archives in the next few years. The era of million dollar mainframes and complicated MARC formats is over. Most archives will be able to send the index to a collection, its control data, *and* the collection itself on a microcomputer disk to any researcher (or other archives) anywhere, whether by popping it in an envelope or firing in down the telephone line. In a variation of the chicken-in-every-pot theme, one will be able to zap archives into every citizen's basement, if they have a microcomputer. This is not science fiction, but feasible now and commercially so by 1988, through laser disks replacing floppy magnetic ones and thus expanding microcomputer memory exponentially and by the development of totally interchangeable hardware and software. This will be as revolutionary to information control and sharing, Dürr asserted, as the imaginary situation of every make of automobile requiring a separate brand of gas, and then someone suddenly inventing an engine or gasoline that made them all compatible, running on a single fuel. By the 1990s the microcomputer will be as accepted as the telephone. What does this mean to archives? Simply that the issues raised by Baskerville and Gaffield, by the Vancouver Island Project, and by this conference are ones of immediate concern to the archival profession, not esoteric musings or project specific. The new technology means that archivists will be tied to creators and users' information networks whether they like it or not. As Dürr said, archivists need to think more clearly about the information matrix in which they sit. Archival automated systems are usually developed with separate foci: to provide subject access and other services to users, to administer or manage the archives' internal operations, to link into creators or producers of records and their automated systems and EDP records, and to share information with other repositories and users through networks. The information needs and thus automated systems of producers, archivists, and users have their discrete concerns, of course, but there are important areas where any two overlap, and a crucial central area — as Baskerville and Gaffield also stress with their new paradigm — where all three intersect. That should be the locus of archival activity in the first instance as it relates to automation and access.

The key to successful automation, as this conference highlighted repeatedly, is having a very clear idea of what to automate, and why, and for whom? Such questions perforce drive one back to examining the basic concepts and theories of archival arrangement and description. That work in some ways is only beginning, and the complexity of research needed to do it properly, as Bearman outlines, is often not even recognized in some archival circles. Well, the needs of producers and users are here today and the technology will be easily available tomorrow. Will archivists be ready to use it intelligently and consistently to fulfil their role in the shifting information paradigms?