Bibliographies/archival inventories

BRACERS: the Bertrand Russell Archives catalogue entry & retrieval system (I)

by Kenneth Blackwell

HOW MANY LETTERS by Russell do we have? What correspondence survives from 1 May 1918 to 14 September the same year? Where was Russell during May 1917? A letter in the 1950s concerns Korzybski—which one? All of these questions, and many more, cannot be answered (if at all) in a reasonable period of time by conventional means.[1] It is high time we were able to answer efficiently these and a host of other definite, research-supportive questions.

For three years the staff of the Russell Archives have been creating, testing and running a computerized cataloguing project with the goal of one record for every piece of correspondence under our control. The basis of the project is the archival inventory begun for the other collections at McMaster, to which the names of the Russell correspondents were to be added. The extension to the Russell Archives took place in January 1988. Charlotte Stewart-Murphy, Director of the Ready Division of Archives and Research Collections, gave me the delightful opportunity to add the fields I considered necessary to construct not just an index but a catalogue of Russell's correspondence. All this was to take place using software and a mainframe computer under the guidance of McMaster's Computing and Information Services ("CIS").

This was not the first time an electronic catalogue had been projected. As soon as the Archives arrived, in 1968, the University Librarian, William Ready,

1 Though we try. For finding-aids, see my "Finding-Aids of the Russell Archives", *Russell*, n.s. 5 (1985): 66–71, and Albert C. Lewis, "Some Computer-Produced Finding-Aids for the Russell Archives", *ibid.*, 7 (1987): 167–70.

was keen to enter the data for each letter into a giant computer. Some staff from CIS even worked on a prototype. Partly to this end the classification system I devised for the papers was numerical, with a detachable, unique number for each document.[2] Dr. Ready and I used to mention the project in interviews, and at least one newspaper reported that we were "attempting to programme the papers for a computer: so a researcher could ask the machine a question and get a full list of Russell views on any subject—religion, sex, nuclear disarmament, mathematics, birth control, education, marriage, child-raising, passive protest."[3] A guide to an interim catalogue of the Archives referred to "the immense task of microfilming and computer indexing them *in toto*." The microfilming task was completed, but all that survives of any actual computer product is single sheet of input and another of output. I worked during the next year on an analysis of the relevant (and irrelevant) fields as an academic project. In those days microfilming specialists and data-retrieval experts seemed to be working together, but the practical difficulties of uniting pictures of the documents with retrieval of their texts must have been one source of failure. Nowadays it is technically unexciting to combine graphic images with character-based retrieval, as in police data banks. One can imagine colour-scanning all the documents and then subjecting them to character-recognition software. Alas, the funds and space to create, manage, staff and store such a project are not available. The first project was abandoned because of the size of the task, the lack of precise objectives, and underestimation of the technical difficulties. My own lack of database training must have been a factor, and has led me to believe that archivists *need* that sort of training.[4] Then there would doubtless have been financial problems due to the coming provincial underfunding of universities, had all that been understood and the task properly shaped and focused. The present project is utterly different in these regards from both the first and its imagined replacement.

Computer-based projects attract acronyms. No one is especially keen about

2 As recorded in an internal memorandum, "The Russell Archives: Document Numeration". It contains the caution: "It may be several months before the computer catalogue is out...."

3 John Doig, "A Young Man and World's Oldest 'Peacenik' ", *Toronto Daily Star*, 2 Oct. 1968, p. 69. The first issue of *Russell* pledged that "a computer-based catalogue of all the material will eventually appear—when the problems of subject-indexing have been solved" (Spring 1971, p. 2).

4 The availability of personal computer programmes like Ashton-Tate's dBASE has assisted me, as it must many archivists, in doing their jobs better. We make use of a growing number of in-house databases. I wish the contents of the *Bibliography* were in a database. (Thanks to hardware provided by my SSHRCC grant, Arlene Hill's downloading, and Dr. Lewis's programming, the contents are electronically indexed.) Perhaps they can be extracted for BRACERS, which needn't be limited to cataloguing letters. The same structure could be adapted for manuscripts and printed documents.

"BRACERS", but it does have a ring of Monday-morning reality to it. The point of the "ERS" is to stress the two main parts to the electronic process: extracting and *entering* the data, and querying the database to *retrieve* it. Because the parts interact, and we adjust our cataloguing rules and recommend system enhancements with great care, it's a *system.*

BRACERS is independent of massive funding, local hardware requirements (and the obsolescence thereof) and, at the input level, even special-purpose staff in the Russell Archives. In some form, it's likely to survive cutbacks. All the data resides in mainframe storage in another building—equipment about which we need know little and can do even less. The staff, including myself, totals two fulltime positions. The input is done by the job-sharing Archival Assistants for 1.5 hours *every* day (all of it instantly useful),[5] the contract position of Archives Cataloguer/Bibliographical Assistant (due to expire in June) prepares and proofreads output daily and maintains the manual, and the Archivist decides on priority and quantity of input, devises software queries,[6] works with the Archives Cataloguer on enhancements, and fills in where necessary to keep BRACERS advancing daily. The Cataloguer, Sheila Turcon, adds a good deal of quality from her dedication and background in the Archives, but the work could continue, at a reduced level, without her position. We have no hardware problems to solve, because there are none that *we* can solve; we have no software choices to make, because that's not up to us. We are simply in the enviable position in the automated world of having no greater duty than to maximize what we have to work with. Having to contend with desktop typesetting for this journal, I know whereof I speak. One often has to choose between technical self-sufficiency and efficiency in other things. This is not to ignore flaws in our system, e.g. its lack of textual sophistication.

BRACERS began, then, as a structural copy of the data-entry system devised for the University Library's other archival collections, with which it remains related. The original system was devised to capture details about individual documents, which suited us fine. There are, however, no recognized descriptive standards for archives at the item level. The fields in place were the collection code, box and file numbers, a pair of "to" names and a pair of "from" names, the date, form (whether autograph letter signed, and so on), and notes (used for the number of pieces in a document). We found a need to add fields for our three-digit classification number, the six-to-eight digit document number, the source in the case of recent acquisitions, notes on enclosures, an

5 They are Judith Bourke and Jeanette Wilkinson. Cheryl Walker did the input for 2.5 years, but the greater part of it awaits retrospective checking by the Archives Cataloguer.

6 There are two query languages: SQL and QBE. The ready-made queries are in SQL. Their structure is a model for new queries, and CIS's Sonya Martin helps me revise them.

address code, notes on publication of the item, and topics.

The topic field was used for an experiment in subject indexing. Abstracting a letter's contents—even the proper names and referents of various denoting phrases—takes a great deal of specialist time. A controlled vocabulary is required as well as what librarians call an authority file for the standardized form of personal names.[7] The software analysts found the topic field the most difficult of all to get working the way we needed it.[8] In the end we did find it possible to download our vocabulary and alphabetize it in a wordprocessing programme. That was the minimum requirement for thesaurus control. We also had to standardize the abbreviations used in the name fields.

Meanwhile the University Library had been funded to install and convert to an on-line catalogue of McMaster's book and serial holdings. A demonstration of the software, NOTIS, showed a keyword search—in combinations of which George Boole surely never dreamt—of all the many fields of the many thousands of records in a particular library system dialled into for the purpose. Textual retrieval like this is what we need, I thought to myself. At the same time conversations with colleagues at the Russell Editorial Project on free-form textual databases led to exploring other indexing programmes. The result was a revised system with a number of ready-made queries that search certain combinations of fields for textual "strings" of our choice and that relate BRACERS to the on-line catalogue of the other Ready Division collections. For instance, I wanted to find the correspondence concerning the edition of N.I. Lobatchewsky on whose editorial board Russell served in the 1920s. The string "%MOSCOW%" (the percent signs represent any number of characters in the string) in the name query found it, although that would be a futile point of entry for *A Detailed Catalogue of the Archives of Bertrand Russell,* where the entry is "Mathematics Institute, Moscow University".[9] This search concerned a standard, fixed-length field. Of equal interest is the capacity to search a variable text field. Our software analyst devised a 32-line Topic_Text

7 An example of this approach is detailed in Ronald J. Zboray, "Microfilm Editions of Personal Papers and Microcomputers: Indexing the Emma Goldman Papers", *International Journal of Micrographics and Video Technology,* 5 (1986): 213–21.

8 Our chief software analyst is Roland Carter. The programmes he incorporated in BRACERS are DB2, ADF (for entry), and QMF (for querying). He provided us with a manual, *Archives and Research Collections System.* All programmes run on an IBM mainframe. Data entry is via a 9600-baud ROLM dataline connected to an IBM PC/XT (although slow by today's standards, it responds in terminal mode as fast as the dataline) on which time is made available by the Russell Editorial Project; my own querying is by modem.

9 Ed. Barry Feinberg (London: Continuum 1, 1967). Although the index to *A Detailed Catalogue of the Second Archives of Bertrand Russell* (not yet published) is much fuller, it doesn't pick up "Moscow" either, because a different corporate name is on the later letters and was used in the main entry.

field into which we can put lengthy notes or even the text of a letter itself. We lack the staff to input text on a large scale. The task seems as formidable as subject indexing but requires less specialist time; for now we must be satisfied with having installed the potential for it. We are just completing the cataloguing of Russell's letters to Lady Ottoline Morrell. We have entered, for precise identification, either the first sentence of each letter Russell wrote her or my extensive notes on the letters when one still had to go to Austin, Texas, to read them. The Topic_Text field contains some of Sheila Turcon's notes on the physical features of the letters. For another section, it contains my contents notes on Russell's dictation from 1952 to 1958. That makes possible a quick, reliable answer to the Korzybski query at the beginning of this paper.

We are nearing the point where it will be more efficient to consult BRACERS than any other tool. There has had to be some priority of input. All of the Feinberg catalogue listings of the personal correspondence (class 710) are in, as is the listing of personal correspondence in Archives II. These entries are on the so-called file level rather than the document level. They summarize groups of letters rather than individual ones. Also in BRACERS at the file level are the Publishing Correspondence (*410) and the early Bertrand Russell Peace Foundation correspondence (*313). The Dictation (*750) through April 1958 and the Ninetieth-Birthday Cards (*724)[10] are in at the document level. The file-level entries will eventually be replaced by document-level entries. File entry is desirable in order to gain quick control over the personal correspondence and notes, especially the many revisions made by Ms. Turcon to the Feinberg catalogue. The number of records was 18,981 at the end of 1990. Only the unwary would forecast a completion date now.

The most interesting group of letters to catalogue has been those to Lady Ottoline. The copies from Texas are among the most read in the Archives. Their dating is notoriously bad, and although numbered in Lady Ottoline's day there are gaps. It would be very useful for researchers to list them all and improve their order. Such improvement would take place at the logical level in the computer, not at the physical level of documents. At the same time we could make the first extensive use of the address field. And we could get such mundane information as an accurate count of the letters and their enclosures. It had always seemed that 2,000 was a good estimate; now I stand corrected. We shall be telling visitors (provisionally, pending confirmation at Texas) that a mere 1,870 letters survive—plus 57 other letters, not between the pair, for a total of 1,927. Before starting on this correspondence the number of address

10 The birthday cards include a lot of messages—almost letters—from friends and other important people. Only a few were published in *Into the Tenth Decade: Tribute to Bertrand Russell* ([London: Printed by the Malvern Press, 1962]).

codes we had was far less than the present 127. The current London section of the codes makes an interesting illustration. Russell's residences are indentified.

LAC	London: Athenaeum Club, Pall Mall, S.W.
LAG	London: 153, Ashley Gardens (Dora Russell's father's flat).
LBM	London: British Museum.
LBP	London: Brixton Prison, S.W.2. (BR's residence in 1918.)
LCI	London: 4, Clement's Inn.
LCK	London: Cranston's Kenilworth Hotel.
LCS	London: 17, Carlyle Square (the Whiteheads).
LEG	London: 47, Emperor's Gate, S.W.7. (BR's residence in mid-1930s.)
LFS	London: 6, Fitzroy St. (The "studio" at #5 belonged to BR & Colette, 1917–18.)
LGB	London: 46, Gordon Square (Clive Bell).
LGH	London: Grosvenor Hotel.
LGS	London: 57, Gordon Square (Frank Russell). (BR's residence, 1916–18.)
LHS	London: 43, Hasker St. (BR's residence in late 1950s and 1960s.)
LIH	London: Ivanhoe Hotel.
LLH	London: Liverpool Street Hotel.
LMG	London: More's Garden, Cheyne Walk, Chelsea (a block of flats).
LOV	London: 70, Overstrand Mans., Prince of Wales Rd., Battersea. (BR, 1919–20.)
LPH	London: Paddington Hotel
LRC	London: 34, Russell Chambers, Bury Street, W.C. (BR's residence, 1911–16.)
LRI	London: 41, Queen's Road, Richmond, Surrey. (BR's residence, 1950–55.)
LRP	London: Richmond Park.
LSY	London: 31, Sydney Street, S.W.3. (BR's residence during 1920s.)

How many people write from 127 places in their lives? Russell's mobility, and evident desire to move around, impress all of us. And we have far to go before completion. The address field helps in dating letters and in biographical research. For example, Ms. Turcon has identified many of the obscure addresses in Russell's pocket diaries as the homes of friends where he intended to spend the night.[11] The accumulation of this data, essentially a tool, is new knowledge for Russell studies. Collecting the data has to go by certain rules: we don't infer addresses where Russell gives none, but we do accept a postmarked address or a printed letterhead address if it doesn't conflict with other information. The assignment of dates thus reflects old-fashioned human intelligence. The next scheduled group of documents, the "802" letters to Lady Constance Malleson, will multiply those address codes again. We will then have an near-daily address register for long stretches between 1911 and 1970.

The retrieval system will be very good at answering queries like the first three at the beginning of this article. It's too early to answer the first question.

11 Fred Keay, our free-lance researcher in London, looked up several of the addresses in relevant street directories. One frequently visited address was that of Russell's dentist. He wrote many letters from trains and on board ship, but none in the dentist's chair.

The second concerns Russell's first imprisonment, during which he wrote many excellent letters. Although we have put together a file of photocopies of these letters, we cannot be sure of having found them all. BRACERS will do that for us. Russell's whereabouts during a given period is of intense interest to editors and biographers. Now we can track him better than ever before.

The computer facilitates the discovery of our own errors, such as inconsistency of input and false addresses based on misdating. As the catalogue grows, we anticipate more positive discoveries. One will be the roster of those to whom he wrote on a given day. A remark that has intrigued me is to Lady Ottoline on 26 May 1911, when Russell told her that he had "cleaned off a host of letters". Which were they? The computer can't tell us yet, but some day it will, if only the letters are extant. We'll see them catalogued on the print-outs, which now list only the letters to Lady Ottoline. And if staffing should become sufficient to input and check the text, researchers will be able to request reports of the letters that use this or that terminology. The future of such requests does seem to lie in a textual approach, rather than in subject indexing. Researchers now examine letters for concepts that are expressed or alluded to in a variety of terms, using their knowledge of Russell's life and thought and their own academic backgrounds. Our database will function for them as an electronic extension of their own searching techniques.

Just as MORRIS, McMaster's on-line book catalogue, is publicly available via modem off the campus and, therefore, around the world, I expect that some day BRACERS will be, too—perhaps as part of MORRIS. If the letter texts are there, BRACERS will constitute a kind of substitute edition of the Archives microfilm or even of Russell's *Collected Letters*, suitable for research if not for enjoyable reading at a researcher's home base. If the text is not available, BRACERS will still be a catalogue of unparalleled flexibility and completeness for Russell studies. It will relieve the staff's burdened memories to concentrate on connections that are less susceptible to being pinned down in electronic form. BRACERS could be published in a paper format, but it would be difficult to regard that, even in the extreme case of publishing the complete data tables, as more than a one-sided view of a multifaceted object. Indeed, its true utility lies in its very unadaptability to paper form. The most common form of "publication"—outside of the possibility of CD-ROM—will probably be as single-copy reports in answer to this or that researcher query or, at the reference library level, as extracts of the data in certain name fields. Paper catalogues of growing collections are soon outdated. Since new documents arrive almost monthly in the Russell Archives, the practical way of keeping our catalogue up to date is to maintain it in electronic form.

The Bertrand Russell Archives
McMaster University Library