Windowing the Past: A Seventeenth Century Technological Archive and Its Electronic Exploitation

Alasdair Paterson
University of Sheffield
The management of manuscript collections - unique, irreplaceable, fragile - has historically struggled to reconcile preservation with access. Academic librarians know only too well the uneasy negotiation between the conflicting imperatives of security and availability, conservation and use, which comes to define a level of public service in this area.

It is therefore not surprising that here, as elsewhere, the potential of new electronic media is increasingly a matter for investigation. Manuscripts transcribed, keyboarded or scanned into microcomputer systems become "virtual" manuscripts, no longer unique or fragile, potentially unlimited as to time, place, frequency and volume of use, retrievable through a multiplicity of access points, manipulable into sets or views for specific research or teaching purposes.

This paper chronicles one such process of investigation and experiment, Sheffield University's Hartlib Papers Project, which has been set up to test the outcomes of the application of the latest in information technology to a collection of some 25,000 manuscript sheets.

The Archive

Samuel Hartlib, born around the year 1600, was a Prussian immigrant to England who corresponded indefatigably with the great names of his era - Boyle, Comenius, Cromwell, Descartes, Milton, Pascal, Wren - planned educational and social reforms, petitioned kings, acted as a clearing-house for scholarly communications, collected the papers of other savants and contracted out his services as an information searcher and provider. In the rhetoric of this conference, Hartlib was an evolutionary figure in the history of information science, whose fate it was to live in an era of revolutions and conflicts - the Scientific Revolution, the English Revolution and Civil War, the continuing effects of the Reformation in a divided Europe torn by the Thirty Years War - and it was their interplay which variously inspired, facilitated and impeded his efforts.

The correspondence contained in the archive in Sheffield University's Main Library demonstrates that Hartlib's contacts extended across Western and Eastern Europe to the Near East, to Japan and to the American colonies - all this the residue of an enormous effort in the gathering and exchange of information. It is sad to record that Hartlib's undertakings, which brought him a contemporary reputation as "the great intelligencer of Europe" and his appointment by the Cromwellian government as State Agent for Universal Learning, resulted in schemes left tantalizingly short of realization, death "in extreme poverty" and scant recognition from posterity. Even a likeness of his features has eluded us.

Through some miracle, however, a large portion of Hartlib's papers have survived; those in the Sheffield collection account for almost all those known to be extant, though more may yet be discovered. Topics covered in the archive range from military science to librarianship, from astrology to horticulture, from architecture to medicine, from silkworm-rearing to Pascal's computer from universal languages to the restorative nature of the King of Denmark's wine sauna - one could continue at length. All in all, the archive affords a fascinating and invaluable window on the intellectual concerns of Hartlib, his
circle and his times, with the most diverse material to be found in Hartlib's
day-book, the Ephemerides, in which he records and speculates on all things of
interest, creating as he goes a topical index in the margin and producing a
text whose chaotic penmanship has been a severe test for transcribers and an
impenetrable thicket for OCR programs.

So much of what those in the Hartlib circle discussed and passionately
believed in - the concept of the Invisible College, the idea of an Office of
Address as a kind of clearing-house for information, the desirability of free
and prompt exchange of discoveries, the notion of universal enlightenment -
would strike a chord with us, and this contemporaneity is part of the
excitement of the archive. On the other hand, it must be conceded that much of
what Hartlib and other Puritan intellectuals promoted was part of a rather
different agenda to our own - not only an intellectual cooperation world-wide
which would inaugurate a new era of peace and prosperity, but a preparation
for the millennium. Hartlib and friends were great readers of the Book of
Revelations, and believed with their mentor Francis Bacon that man, before the
Fall, had enjoyed a universal and perfect knowledge. By restoring knowledge,
by promoting international cooperation and building a single international
language - another concern of the circle - humanity was hastening the second
coming of Christ, the most desirable of future events. As Milton wrote in his
Tractate of Education, dedicated to Hartlib: "The end then of learning is to
repair the ruins of our first parents by regaining to know God aright."

The irrational and magical also recur throughout the archive, testifying to a
period in which alchemy is not yet disengaged from chemistry, nor astrology
from astronomy, nor numerology from mathematics. Thus, for example, papers
concerning ideas on a universal language include not only an early system of
signalling with lanterns which anticipates Morse Code, but also a set of
ciphers for communicating with the spirit world - perhaps the ultimate
ethernet!

It follows that what is unique about the archive is more than the individual
documents. The Hartlib Papers are in themselves a physical embodiment of a
scholarly network and of the philosophical, political and religious convic-
tions which drove it; they record the development of a scientific, empirical
approach to the seventeenth century world while embodying what was, in its
century, a dissident world view; they are both profoundly interconnected and
amazingly diverse; they are a treasure trove of handwriting, spelling, postal
arrangements, water-marks, publishing procedures, pseudo-science. And for some
years they had endured the normal fate of such collections in British academic
libraries - a passive management, kept secure, recorded in a handlist which
could not begin to engage with the volume of topics touched on (sometimes
within a single scribbled letter), little used by scholarship because not
truly usable.

THE PROJECT

Now, aided by grant assistance from the British Academy, the Leverhulme Trust,
the Training Agency and the British Library, a number of initiatives making
use of developing information technology have been set up under the umbrella
of the Hartlib Papers Project, jointly directed by the Library and the
Departments of English Literature and History.

Funded work on the archive began in 1987, when (following a major national
competition held by the British Academy and the Leverhulme Trust) the Hartlib
Papers Project was awarded a grant-in-aid over five years which constituted
one of the largest ever British research grants in the humanities.
The basic Hartlib Papers Project involves complete transcription and editing of the archive, using word processing and full-text retrieval software, working towards an electronic edition (online and on interactive optical disc) and a conventional publication of part or parts of the archive. Additionally, a grant from the Training Agency in 1989 has enabled the construction of a hypermedia demonstrator of a portion of the archive (of which more later) while a British Library grant in this present year has funded a project to build a multimedia demonstrator for use in library staff training programs.

In the main project the transcribers, a team of four graduates supervised by a research associate, initially create a document file for each item using the word-processing program Microsoft Word on a microcomputer network. Word has been adopted by the University as its word-processing standard; there was therefore much expertise on campus to draw on in the initial training and launch stages. Another strength of Word is the comparative ease of preparation of ASCII files for data transfer to other programs to be used in successive phases of the project. Transcription is carried out from photocopies, to save prolonged use of the originals, although each transcription is generally checked against the original to safeguard against the effects of imperfect photocopying, incorporation of showthrough from the reverse of a sheet, etc.

At this stage any necessary translations (mostly from Latin, still the language of scholarship in seventeenth century Europe, but also from German and Dutch) are input, and each document assigned a number of identifiers – document reference, date, names of correspondents, other individuals and places named, books referred to, and subjects. The transcriptions, together with these identifiers, are then stored within a textbase program, AskSam, chosen not simply because it does indeed enable us to ask "Mr Sam. Hartlib Esq." but also because it was in the early stages of the project the program which, through its relational database, free-text and rudimentary hypertext features, best allowed the building up of a transitional archive of transcribed material. Material stored in this manner could then be easily retrieved for editing and transcription comparison as well as for the support of a growing scholarly activity based on the papers. The program is however, really most suitable for those who know what they want; retrieval can be slow and limited, not extending for example to a statement of how many hits have been made, information crucial to query refinement.

The search is therefore underway for a more powerful and sophisticated system incorporating not only free-text retrieval for the expert but also thesaurus layers for less experienced users such as students (and for scholars finding themselves, as is characteristic with use of the archive, suddenly projected beyond their range of expertise), together with hyperlinking between documents. Such a system would also be central to planning the two electronic products to which the Project is committed – an online database of transcriptions, translations and interpretations available through British academic and thereby world-wide networks, incorporating bulletin boards and allowing a collaborative development of the edition, and an optical disk edition which can additionally carry scanned images of the original manuscripts, with potential for sound, graphics, video and animation.

As the main tasks of the Hartlib Papers Project proceed, a number of subsidiary projects involving project members have, as previously mentioned, got under way. The Hypermedia Demonstrator Project, funded by the governmental Training Agency, is an attempt to explore the effectiveness of hypertext and hypermedia in controlling the diversity and highlighting the interconnection of the archive, particularly in allowing the connections to emerge from the documents without imposing either 20th century notions of how subjects interrelate or a succession of views through the distorting glass of narrow academic specialisms. Using Hyperdoc software, whose powerful scripting language seemed best to accommodate(on PCs) image handling combined with scrolling through long text segments, researchers built an impressive stack
from the portion of the archive dealing with bees and beekeeping — full of symbolic and emblematic (as well as economic) possibilities for seventeenth century thinkers and the subject of one of Hartlib's published works. Images of manuscripts from the archive were scanned in, including the astronomer and future architect Christopher Wren's (impractical) design for a multistory beehive; windowing allows original and transcription to be compared, while images can be zoomed in on, bringing out a graphic detail or revealing what a scribbled word actually is.

The final report of the Hypermedia Project is likely to highlight both the difficulty of scripting, particularly in altering screen displays, and the high memory requirements of scanned images. There also emerges a distinction between the evident suitability of hypermedia for building a training module, with possibilities and outcomes fully envisaged and controlled, and its apparent inhospitality to the discovery process desirable in research applications. The required freedoms, in relation to the special characteristics of the Hartlib Papers instanced above, seem compromised by the predefinition of hypertext links and hierarchies.

A Multimedia Project funded by the British Library uses another topic prominent in the archive, gardening and husbandry, in the construction and evaluation of a demonstrator for library staff training — how to handle a reference query relating to a special collection, featuring the reference interview, interpersonal skills, literature searching, construction of bibliographies etc. Using video, sound and a variety of software (principally Supercard and Macromind Director) on an Apple Mac platform, the team will produce, as well as a demonstrator, a brochure on multimedia applications for library staff and a multimedia tool kit specifying key hardware/software configurations.

"What happens next" is not a problem for Project members; as well as working to a deadline for complete transcription in 1992, members are variously engaged in negotiations with book publishers, in a University-funded project to investigate modes of electronic publication, in the organization of a major conference in 1992 to celebrate the quartercentenial of Hartlib's great correspondent Comenius, in writing papers and organizing seminars. What was, from the Library's point of view, a problematic and little known collection of papers has been transformed utterly, while the opportunity to experiment with, and on occasion reject, a variety of novel approaches to text-handling will have a wider significance for the future development both of Library services and of approaches to campus teaching and research.

In addition, Project members are able to feel that, through the evolving technology, they are finally realizing Samuel Hartlib's goals of unhindered transmission of ideas and information between scholars. Coincidently, the opening-up of Eastern Europe has permitted a programme of visiting research fellows from Czechoslovakia, one of whose most celebrated historical figures, the educationalist Comenius, is extensively represented in the archive; Hartlib's own papers have thus played their part in re-establishing the kinds of scholarly contacts they chronicle.
REFERENCES


