Unlocking Global Memory

James Hagan
University of Wollongong, jhagan@uow.edu.au

Andrew Wells
University of Wollongong, awells@uow.edu.au

Gavan McCarthy

Bruce Smith

Publication Details
Theory and method courses in the humanities and social sciences have for many years stressed the importance of the primary source. The best evidence is the original evidence, and oral sources apart, it is documentary. Some of it finds its way into archives where it may become a check, via footnotes, on the ways in which scholars have interpreted events and evolved causal explanations. In its own right it may, with the aid of a scholar’s imagination, pose questions, suggest causes and lead to the evolution of theories.

But constraints of time and finance have meant that traditionally researchers in the social sciences have had to limit their inquiry to what is readily available. In practice, this has meant that they have depended on a few, perhaps only one readily accessible archive. What they have found there has become the foundation of their argument and explanation. That means that their findings have been based on information that has come their way by chance. They have—for good reasons—remained ignorant of information that might have been relevant, simply because it was elsewhere.

Our own research on the ARC project ‘Selling Their Souls’ offers a clear example. This is a study of the use of native labour by private companies and their relationships with governments in recruiting and regulating it. These companies and governments operated internationally. Even with generous funding and four of us combining our research, there is no way in which we can visit and properly exploit the various archives that contain material relevant to our work within the currency of an ARC grant. We need a system of gathering data that will allow us access to more of this information than we can gather by personal visits.

The techniques necessary to establish such a system have already been tested. In 2002 we developed a project called ‘Unlocking Regional Memory’. This was a pilot project that made available some of the records at the archives of four regional universities in New South Wales to researchers at a distance. We created a gateway through which the archival items could be identified and then ordered for display on the researchers’ own screens for a small fee. (A visit to www.nswera.net.au will explain the process in detail.) Our pilot project demonstrated that there were no longer any technical reasons why the whole range of holdings within these regional archives could not be made available cheaply to anyone, anywhere.

This became possible when we found a software solution that was easy to use, was not expensive and had been used successfully in building other internet gateways to archival information. The search for this solution brought the researchers to the Australian Science and
Technology Heritage Centre (Austehc), at the University of Melbourne. The Centre had over a number of years developed a software tool to capture information about archives and their creators and to publish that information in a consistent manner on the World Wide Web. This tool, known as the Online Heritage Resource Manager or OHRM, met the initial needs of the project in that it had a proven record and facilitated the construction of an internet gateway to the holdings of the participating Regional Archives.

The OHRM had been developed by Austehc from a relational database that had been built to suit in-house needs to publish a register of people involved in the development of science, technology, engineering and medicine in Australia, and list their archives. The first iteration of the database was geared to print publication (McCarthy 1991) but the data was also able to be reused to produce a MARC online register available through the National Library of Australia. In 1994 a second iteration of the database enabled the data to be exported in a form suitable for web publication and this saw the birth of Bright Sparcs (Australian Science and Technology Heritage Centre 2004–2005). Demand from other practitioners led to the third iteration of the database as a generalised software tool and in 1999 the first version of the OHRM was released (Australian Science and Technology Heritage Centre 2003). The OHRM, now at version five, has been further developed through utilisation in a variety of projects and would now be better described as a contextual information manager. A strength is that it can map a variety of entity types, note their relationships and register references to archival materials, bibliographic resources and digital objects.

As Gavan McCarthy (Director of Austehc) noted in 1999, initiatives were underway in the heritage industry to capture, structure and use information about ‘context entities’, formerly known as archival authority records: that is, the actors and agents that have played roles, both significant and ordinary, in the evolution of our society. In Australia, work on the HTML encoding of context entities for the World Wide Web has been underway since 1994. Research is now being conducted into the SCML/XML encoding of context entities to enhance the utility of these objects in building bridges between the disparate sectors of the heritage industry, not just in Australia but worldwide (McCarthy 1999). In 2004 the International Council on Archives published the second edition of its International Standard on Archival Authority Records for Corporate Bodies, Persons and Families and this standard reflects the innovative work undertaken by Austehc in the development of the OHRM (ISAAR(CPF) 2004).

The OHRM was originally developed to meet an in-house need. Since 1999 it has been available to the wider archival community as a tool to capture information about record creators and the records they created, and to publish that information on the web. The great strength of this tool is that, unlike traditional finding aids, it permits the expression of a range of contextual relationships between the different record
creators that populate the OHRM. It is the representation of these relationships, the so-called contextual framework, through hyperlinks that gives the subsequent website a value beyond traditional finding aids. It is not so much that the OHRM is a solution in itself, but rather that the continuing development of the OHRM facilitates the expression and/or representation of concepts and ideas about records, record creators and associated relationships that are emerging within the worldwide archival community.

The technical implementation of the OHRM was not problematic and it was clear from the outset that this project would play a major role in driving the evolution of the software. Issues of data quality and the properties and structure of legacy data from the partners played a much bigger role in shaping the early work of the project and continues to be a factor limiting network development. However, once the data about record creators and their records was captured into the database it was a straightforward task to generate the website. During the initial years on the project, 2002–2003, the OHRM itself went from version two to version four, with each version marking improvements in the informatic structure and functional features.

The second part of the project—that is, linking digitised copies of records to information about the records—required an innovative approach. The funding provided meant that the project was limited to creating a pilot that demonstrated that it was possible to make digital copies of archival records available to researchers over the web. The solution needed to be simple to operate, use existing technology and be inexpensive. At the time the project commenced, the National Archives of Australia had begun a programme of providing access to digital copies of archives from their holdings. Prior to this the Archives had undertaken extensive research on the use of digital cameras and the rendering of images to the web (Ling 2002).

Following an exploratory visit to the National Archives of Australia to discover more about the research that had been undertaken and to view the digitising programme first-hand, it was decided to build image viewing and processing software. Existing research provided the basis for this development. It was necessary to develop a solution that permitted computer control of image capture and production of web quality images files, and that enabled the rotation of images captured.

Two software tools resulted which have collectively become known as AIFA (Archival Imaging for Access). The first one processes the images, including rotation and resizing as required, the production of a small image for quick previewing and a corresponding larger image for full viewing. The second tool permits viewing of those images when linked to descriptions of archives and their creators. These tools are platform independent and have been written using open source applications. The cost of producing images is dependent on the nature of the materials but is as a rule less than $1.00 per image, often much less.
The next stage of the project involves upgrading the project OHRM databases to version five and implementing the new functions. Funding is available to work on the legacy data, which will add to the quality and usability of the output. The AIFA tools have subsequently been used on other projects and are ready for another iteration of development and this will be undertaken as part of this project. The imaging project has created much interest in the local archival community and it is hoped that the knowledge and experience gained can be shared with partners overseas.

One significant aspect of the project is the open networking of archival resources based on a contextual information architecture. The archival community has endorsed the standards to support such a network and this project is attempting to implement such connections, especially with major resources in the United Kingdom and in the Asia–Pacific region.

Once we have linked Australian entities in the Regional Archives collections to the same entities in Australian web-based resources, we will link them to any related entities in the Historical Manuscripts Commission in Great Britain. We will then move to a second level that will allow for the linking of citations containing richer information. This level will use data harvesting techniques and provide for automatic updating of information. We will be working mainly at the first level, but exploring the possibilities of the second as we do so.

‘Digitising Global Memory’ will do more than put immense amounts of material at the disposal of individual researchers; it will make feasible comparative studies of much greater validity than those so far achieved. Comparative studies by their very nature require the collation and analysis of data from many sources, but so far logistical difficulties have impeded progress. A developed system of ‘digitising global memory’ will bring information from international sources to researchers’ desktops. Any material they can locate through finding aids—not just that already digitised—will be at their disposal for a small, and probably diminishing, fee. With it they will be able to generate many more contexts and comparisons. From there it should be possible to build generalisations of much greater power.

This potential is apparent in our study of colonial governments and companies. So far, writings on imperialism have been either of the ‘big picture’ variety—applications of a political theory without due empirical regard—or specialist case studies with limited general application. We aim to combine both approaches, but to do so we must research archives in several metropolitan colonial countries as well as the ex-colonial countries themselves. As global memory is digitised, many more, and better, studies will be able to revise and reshape pro- and anti-imperialist theories.

But the sheer mass of information that will become available to researchers will create problems of its own. How are researchers to cope? How will they know where to look, how will they know what
questions to ask and how will they know what tests of relevance to apply? Ironically, the amassing of empirical data will require researchers to develop their theoretical skills well beyond the level now regarded as satisfactory at graduation from our universities.

In undertaking this research and in similar projects, it becomes increasingly clear that the idea of discrete ‘national histories’ as the natural unit of historical analysis is problematic. What we might justifiably call global processes—whether conceived as circuits, cycles, or chains of activity—become more useful ways of thinking. These processes, depending on what they are, may operate nationally, regionally or globally, even though their effects may have a precise physical location. We have occasion to reflect on this in many parts of our work: an understanding of labour conditions amongst Aboriginal cattle workers in the Australian North cannot be disconnected from the Argentine meat trade, nor the investment opportunities calculated in Britain.

The volume of material available through digitised records, the diversity of themes by which it might be searched and the principle of connectivity that can be mobilised, have interesting consequences for historical theory and method. Electronic archives require researchers to articulate a developed principle of thematic relevance, at the very beginning of the research process. What is worth knowing and how best to investigate that problem become pressing questions. Theoretical knowledge becomes an essential tool in using electronic data. At the same time, the geographical, political or ethnic elements that once limited knowledge become less important. National histories, written to assist in the creation of nation states and the construction of distinct ‘national identities’, can be seen to impede historical understanding. The result is the separation of national clichés from the specific geographical impact of global processes.

Thus the project of digitising memory has considerably more than technical and organisational possibilities. It contributes to historical inquiry in many more ways than just offering the technical convenience of access to sources. It demands greater theoretical sophistication in the historical enterprise itself and a break with national histories. These are not just an unintended, but also a very welcome consequence of the proliferation of electronic data.

Notes

1 We use ‘documentary’ here in its widest sense, to include not only text, but illustration, diagram, plan—in short, whatever an archive stores.

References
Books, articles and reports

McCarthy, Gavan (1991), Guide to the Archives of Science in Australia: Records of Individuals, DW Thorpe in conjunction with the Australian Science Archives Project and the National Centre for Australian Studies.

Additional website

http://nswera.net.au