13

From Integration to Web Archiving

John Tuck

13.1 Introduction

The consumer review appendix of Review for JISC of Core Resource Discovery Services (Grayson and Gomersall, 2006, 52) shows that researchers want immediate and easy access to information.

“Many potential users have extremely limited awareness of the bibliographic services that are available to them... Reliance on Google is widespread... Almost all users have very limited searching skills... Even when using Google, many have never got beyond entering a single term... Many users give up easily... Academic users tend to expect immediate access to any useful documents that they find”.

This need for immediate information gratification was not the case when I first met Reg Carr in 1975. Googlization was far from my thoughts when, as a naive first-year research student, I entered the cataloguing room of the John Rylands University Library of Manchester, and asked to speak to the expert in French studies. This was Reg, who spent considerable time guiding me in the direction of relevant research tools for the study of French verse romances of the thirteenth century.

Three years later, this time in possession of a job as an assistant librarian, I walked into the same cataloguing room. Reg had moved on. I was able to build on his work in developing the collections and making them accessible to the next cohort of University of Manchester students of French, encouraging them in their early days of library catalogues and research. Tools of the trade at that time were not web browsers, pdfs or HTML but razor blades (used by assistant librarians to erase and allow corrections to be made to individual letters typed on to five- by three-inch catalogue cards), multilith masters and typewriter ribbons.

Twenty-eight years later, I was struck by an extract from a letter to the editor of the Oxford Magazine (Evans, 2006).
“The repeatedly-delayed Internal Audit Report on the Library Depository Project at Osney Mead: Review of the Business Case was made available on the internal web in August 2006, but no URL has been published in the Gazette, nor has the document itself. This does not really seem like reporting back to Congregation, as Congregation expected when it gave conditional permission for its land to be used for the new Depository under Statute XVI, A,4: “Recognizing the importance of the proposed book depository for the future delivery of library services in the University, and noting that there will be further reports to Congregation as later stages of the libraries’ estates strategy are planned in detail, Congregation approves the allocation of the site at Osney Mead for the purpose of developing an automated depository for Oxford University Library Services.” (http://www.ox.ac.uk/gazette/2005–6/weekly/220905/agen.htm#15Ref) (6 October 2005)”.

Beautifully crafted letters of this type, frequently published in the Oxford Magazine, and often relating to library provision, reminded me of the four and three quarter years (from 1998 to 2002) I spent supporting Reg, as he guided Oxford libraries through the choppy waters to integration and the establishment of Oxford University Library Services (OULS) in February 2000. For a full account of integration and the process leading to it, see Moran (2005, 262–94).

During this eventful period, Reg taught me much more than a grasp of research tools for French studies. He conveyed the essential qualities of patience, perseverance, determination and resilience in the face of any challenge. French literature was never far away, however. In amongst the papers and reports on the desk of Bodley’s Librarian were odd copies of Cahiers Octave Mirbeau and Reg Carr, Anarchism in France: the case of Octave Mirbeau. A year or two after leaving Oxford, I opened up a copy of Sac au dos, a short story by Mirbeau, which, through its title, took me back to an observation Reg made to me just before I left. “John, when you move on, you will put down one heavy rucksack, and pick up another”. And this is what happened as I moved from integration at Oxford to web archiving at the British Library.

13.2 Web Archiving

Wikipedia (15 November 2006) states that:

“Web archiving is the process of collecting the Web or particular portions of the Web and ensuring the collection is preserved in an archive, such as an archive site, for future researchers, historians, and the public. Due to the massive size of the Web, web archivists typically employ web crawlers for automated collection. The largest web archiving organization is the Internet Archive which strives to maintain an archive of the entire Web. National libraries, national archives and various consortia of organizations are also involved in archiving culturally important Web content”.

The history of web archiving in the British Library goes back to 2001 and a proof of concept pilot called domain.uk. Here an attempt was made to test the feasibility of harvesting a small number of UK websites (up to 100) on the basis of permissions having been granted by the website owners. The pilot proved successful and led to the establishment of a full web archiving programme, now firmly embedded in the British Library’s strategy.