Digitisation of Special Collections: National and Academic Libraries' Examples

Rochna Srivastava

Abstract

The paper describes a few major national and academic libraries of the world that have undertaken projects to exploit the potential of digital technology for displaying unique and fragile materials. It also highlights major goals of these digital projects. Selection criteria of special collections for digitisation, by these libraries, have also been discussed in the paper.

1. SPECIAL COLLECTION

Special collection forms the core of the primary research collection of a library. Whatever may be the format, the common features of special collection are that they are unique, rare, require special handling, and in limited availability or just something that cannot be located in the open stacks. Equally important is that the material is valued as an artifact. Special collection may use digitisation as a preservation technique but its real value is related to the increased access to special collections.

The selection of special collection material for digitisation is based upon the same criteria for placing material in special collections: uniqueness; artifactual value; importance to original research; etc. Almost all major libraries of the world are having digitisation programme to provide greater access to its special collections through the digitisation of material including photographs, manuscripts, maps, printed music, pictures, etc.

2. NATIONAL LIBRARIES

2.1 The British Library (BL)

There is a growing number of collaborative digitisation initiatives within the British Library (BL) and the library already has an estimated four terabytes of digitised information[1]. Special collections of the BL are digitised to maximise their use by facilitating a greater volume of networked access and by providing the enhanced functionality intrinsic to the digitised items. Materials covered under digitisation programme may be:

- Owned by the BL
- Deposited with the BL
- Owned by a collaborative partner of the BL
- Owned by a commercial partner of the BL[2].

The Early Printed Collections Department of the BL provides wider access to important items by digitising them so as to preserve fragile originals for future generations of researchers. A wide range of fine and historic bindings found in the BL are being digitised. Well represented bindings are mainly British and Continental.

Database of Bookbindings[3] is an image database of bookbinding compiled since 1997 covering Western European printed books. Evanion collection of Henry Evans (1832-1905) containing around 5,000 items relating to Victorian entertainment and everyday life has been digitised by the BL[4]. 'Burney Collection of Early English Newspapers' is trying to make the text of newspapers and newsbooks of pre1801 fully

searchable via a website. Individual plays of Shakespeare, referred to as *quartos,* are completely converted into digital copies under the project 'Shakespeare Quartos'. About 200 books describing public displays at royal marriages and funerals, stately entries into cities, and other similar events have completely digitised under the project 'Renaissance Festival Books'.

The digitisation project 'The Gutenberg Bible' aims to provide wider access to this rare material printed in Mainz around 1454-55[5].

Another project 'Canterbury Tales' has first two editions of Geoffrey Chaucer's 'The Canterbury Tales' as printed by William Caxton in 1476 and 1483 in complete digital form. In order to increase access to Beowulf, the oldest surviving epic in British literature has been published on CD-ROM (The Electronic Beowulf) by the BL.

'Turning The Pages' is an award-winning interactive display system developed by the B L to increase public access and enjoyment of its treasures[6] using touch screen technology and animation.

'International Dunhuang Project' is to promote the study and preservation of pre eleventh century manuscripts and artefacts.

2.2 National Library of Australia (NLA)

A number of digitisation projects have been undertaken by the NLA[7] for long-term preservation of traditional documentary materials.[8] Digitisation provides access to electronic information resources such as image databases, recordings, theses, manuscripts and the full text of periodical articles index by the library[9].

Special collections (including photographs, cartoons, drawings, transparencies, negatives, postcards, maps, printed music and manuscripts) of the NLA has been digitised under a major digitisation programme aiming at providing greater access to its collection. This digital material is made available on the internet through:

❑ Online exhibitions

❑ The Library's catalogue

❑ PictureAustralia

❑ MusicAustralia

❑ Library's multimedia applications.

2.2.1 Pictures

The Pictures' Catalogue provides access to over 65,000 pictorial images available online which are also exposed to internet search engines through a URL list.

2.2.2 Music

Some music albums published before 1930 and all Australian sheet music (composed by an Australian, or on Australia or published in Australia) available at NLA is now available through digitised printed music collection. Initially over 2,000 sheet music items were released which are exposed to internet search engines through a URL list.

2.2.3 Maps

All sheet maps and selected atlases published before 1900 possessed by the National Library of Australia are digitised. Initially, over 1,000 sheet maps from the Nan Kivell rare map collection were released. More maps from the Ferguson rare map collection are expected to be added to this collection thus reaching the collection over 2,000.

2.2.4 Manuscripts

The Library's online exhibition features digitised manuscripts' collections of the NLA (e.g., Diary of William John Wills and the papers of Sir Edmund Barton). All kinds of unpublished records, maps and plans relating to Cook's South Sea voyages have also been digitised. NAL is in the process of digitising oral history and folklore collection and will be made available online very soon. As part of the International Children's Digital Library Project, it is also involved in digitising Australian children's works.

2.3 Library of Congress (LC)

The Library of Congress (LC) is trying to develop a national strategy for the preservation of digital information[10] and has carried out its preservation responsibilities by providing adequate storage conditions,

reformatting materials from their original fragile formats to more stable media and from rare and unique materials with intrinsic value to restoring originals through conservation treatment.

American Memory,[11] the flagship of the library's digital services, offers more than 7.5 million digital items from more than 100 historical collections. Multimedia collections of digitised documents, photographs, recorded sound, motion pictures and text from the American historical collections of the LC and other institutions are now the part of American Memory, the National Digital Library.

2.3.1 Sound Recordings

The majority of sound recordings included in the American memory are taken from cylinder and disk recordings and are offered in three formats requiring special players:

♦ RealAudio delivering sound fast, with a minimum download time

♦ MPEG2, layer 3 (.mp3) offering high quality sound downloading

♦ WaveForm (-wav) higher quality version.

2.3.2 Motion Pictures

The motion pictures, (from 1894-1912) deposited to LC for copyright, were chosen for digitisation. They were all black-and-white and silent whether originating from paper prints or from early film copies in other collections. Current motion pictures are presented in three formats:

❑ .mpg (Motion Pictures Experts Group or MPEG) containing highest quality film files

❑ .mov (QuickTime) offering small downloadable files

❑ .rm (Real Media) allowing viewing of the film as it arrives at the users' computers.

2.3.3 Maps

The digital map collection represents only a small fraction of the Geography and Map Division of the LC holding more than 4.5 million items. Digital images created from maps and atlases are restricted to items not covered by copyright protection. The focus of map collection is American and cartographic treasures of the LC. Maps are available in three images:

❑ Master image in TIFF file format
❑ Reference image in SID file format
❑ Thumbnail image in GIF file format.

2.3.4 Documents (text and page images)

American Memory document collection also includes pictures of pages and illustrations from original documents. They are presented in five formats:

❑ Images of original pages (.tif or .tiff format)
❑ Images of original illustrations (.pcx format)
❑ Page images (.pdf format)
❑ Page images for large-format document material (like newspapers and maps) (.sid format)
❑ SGML viewer option (text).

Online collections of manuscript materials such as Presidential papers, papers of Government officials, organisational records, foreign sources and special collections are available as digital images or as searchable text.

3. CO-OPERATIVE DIGITAL PROJECTS

3.1 Berkeley Digital Library SunSITE

The University of California, Berkeley Library is using digital technology to build a national digital library in collaboration with a number of other organisations to enhance access to information while preserving it for future generations.[12] The Berkeley Digital Library SunSITE[13], sponsored by The Library, UC Berkeley and Sun Microsystems, Inc., builds digital collections and services while providing information and support to digital library developers worldwide.

California Heritage collection[14] is the result of the California Heritage Digital Image Access Project covering photographs, pictures, and manuscripts illustrating heritage of California. Another project Catalan Manuscript contains approximately 900 images of early paper and vellum

manuscripts. 'UC Berkeley/Columbia Digital Scriptorium Project' aims to select Latin dated and datable manuscript (to 1550). An online database, Advanced Papyrological Information Systems (APIS), is of digitised images of papyri. The 'MusiLan Project' has created an online digital listening reserves system for the Department of Music. Historic Topographic Maps of California project provides web access to the historic topographic quadrangles, both 15 and 7.5-minute. 'Aerial Photography Online' is a collection of aerial photographs representing aerial flights of the San Francisco bay area and Yosemite National Park.

'Incunabula', a searchable database of the incunabula collection includes philosophical, theological, scientific, mathematical, historical, legal and poetical works. Siberian Digital Photo Collection, from late 19th century Siberia, is presented by the UC Berkeley Digital Library SunSITE and the UC Berkeley Slavic Collections. The Last Jews of Libya is the site containing photographs, writings, and the spoken words of them. The Jack London Collection contains letters, postcards, telegrams, manuscripts, official documents, photographs, audio and other papers relating to his life and work.

3.2 SCRAN (Scottish Cultural Resources Access Network)

The SCRAN is a Millennium Commission funded project, which aims to build a networked multimedia resource base for the teaching and celebration of human history and material culture in Scotland.[15] SCRAN, a registered charity, began in November 1996.[16] The founding partners are the National Museums of Scotland, the Royal Commission on the Ancient and Historical Monuments of Scotland and the Scottish Museums Council. It was set up as an independent, not for profit, limited company to manage the digital intellectual property rights of its members. Although based on the libraries, archives, museums, and built heritage of Scotland, SCRAN's prime concern is neither with conservation, nor with documentation, but with educational access.[17] SCRAN supports

learning through easy access to images, movies, sound and virtual reality. There are one million records from museums, galleries, archives, media, and contemporary and performing arts. Images are derived from primary source material such as: paintings; drawings; 16th century manuscripts; and trade union banners. It provides easy access to 1.5 million text records of artefacts and historic monuments, and 100,000 related multimedia resources. SCRAN is also commissioning 100 multimedia essays, based on these resources and others, for use in schools and by a wider audience. SCRAN is a rights clearance project, grant aiding the digitisation of assets in exchange for a non-exclusive licence for their educational use. It is also a resource disclosure and delivery project and acts as a metadata repository, pointing to individual digitised assets in its own resource base and to objects in the real world, as well as acting as a gateway to other electronic collections. High quality cultural resources are online including multimedia records, still images, video clips, audio files, virtual reality panoramas and essays.[18]

Each resource has a caption providing valuable background information for reference or research. Thus teachers and learners, through SCRAN, have the best possible opportunity to put ICT into practice in a way that is right for them.

4. ACADEMIC LIBRARIES

4.1 Oxford University Library

The Oxford University Library[19] has extremely rich and diverse resources provided by more than 100 library units. Throughout the last decade, the University of Oxford has been proactive in digitising materials from its own collections and making them available to scholars nationally and internationally. Working concept of the '*hybrid library*' is the base of the forward plan for the Oxford University Library Services (OULS). The Oxford digital Library (ODL) has been established as a key component in the implementation of a comprehensive e-strategy for OULS.[20]

The ODL plans to build up a significant set of digital resources for local and remote online access. It aims to offer a digital library architecture allowing centralised access to digital resources, both created by the Oxford Library and those acquired from outside.[21]

First venture of the University was the Toyota City Imaging Project, instigated by a donation from the Toyota Motor Company and concentrating on automobile ephemera held in the John Johnson Collection. Soon afterwards, through a national grant, the Celtic and Medieval Manuscripts Project began digitising material from the University's extraordinary collections of extant material from the Middle Ages. The momentum has been maintained throughout the 1990s, allowing the University to build expertise in the digitisation of grey literature, photographs and artefacts (e.g. the projects at the Ashmolean Museum and the Beazley Archive), rare print material (e.g. the Bodelian Broadside Ballads Project), and multimedia (notably the Wilfred Owen Multimedia Digital Archive, which digitises audio and video clips).

The Internet Library of Early Journals (ILEJ), a joint venture between the University of Oxford, Birmingham, Leeds, and Manchester is the example of collaborative initiative. Not only this but pioneering work by the Oxford Text Archive and the Digital Library Project have placed oxford in a leading role of providing document delivery and access to structured texts. Potential collections identified for digitisation are rare and unique '*image'* material predominantly in the arts and humanities. The general opinion for increasing access to collections is that material currently in high demand (of both high and low artefactual value) should be targeted initially.

4.2 Cornell University Library

Electronic library components are being created in almost all sectors of Cornell University Library. Cornell's digital library research and development initiatives include projects like CLASS (College Library Access and Storage System), MOA (Making of America), KLIC (Kodak Library Image Consortium), TULIP (The University Licensing Program), NEEDS (National Engineering Education Delivery System), and DAISY (Dissertations Available on Internet Systems).[22]

■ Primary Visual Images-In this digital collection include:

❑ Beautiful Birds: Masterpieces from the Hill Ornithology collection project displays changing techniques during 18th and 19th centuries.

❑ Fantastic in Art and Fiction project is based on graphic material from the Cornell Witchcraft Collections.

❑ Herbert F. Johnson Museum of Art, Museum on-line database search options, over 16,000 images of works of arts of the Johnson Museum of Art.

❑ Icelandic and Faroese Photographs of Frederick W.W. Howell provides an important visual record of Icelandic and Faroese life and landscapes at the close of 19th century along with historical information.

❑ New York School of Human Ecology: Historical Photographs is a collection of over 1400 photographs of students and faculty, buildings and scenes that portray the college from its founding through 1969.

❑ Utopia database, is a collaborative project providing images of European Renaissance Art, primarily from the 15th and 16th centuries.

❑ Cornell University's major multimedia projects are:

❑ Death of the Father examining the closure of political authority following the death of Stalin, Hitler, and Mussolini

❑ Global Performing Arts Database (GloPAD) with the aim to develop a consortium namely GloPAC (Global Performing Arts Consortium).

❑ Agriculture Network Information Center (AgNIC): An internet portal.

❑ French Revolution Collections: Web Exhibits and Guides is based on the collections of French Revolutionary papers of Cornell.

- International Court of Justice Web Mirror Site is the first official web site, for the Court created by the University.

- Making of America (MOA), funded by the Andrew W. Mellon Foundation and the Charles E. Culpepper Foundation, was initially started by the Cornell University Library and the University of Michigan Library. It provides access to 267 monographs and over 100,000 journal articles with 19th century imprints.

The US Department of Agriculture (USDA) through USDA Economics and Statistics System (ESS) provides fast and free electronic access to vital agricultural information via the World Wide Web.

5. DISCUSSION

Thus, it is clear from the above description that many institutions and libraries have undertaken projects to exploit the potential of digital technology for displaying and researching unique and fragile materials. Similarly, many institutions prioritised special collections for digitisation because of their value, rareness or uniqueness, and for the attention such high-profile initiatives can draw to the institution.

The aim of almost all the digital projects discussed above has been to enhance access to rare and unique primary materials that may otherwise have limited scholarly and educational importance at their geographical location, to aid in the preservation of these special materials through the creation of digital surrogates, and to create learning opportunities for remote users, including those in colleges and universities, schools and to the students of distance learning.

REFERENCES

1. Shenton, Helen. From talking to doing: Digital preservation at the British Library. Paper presented at the International Conference on the Preservation and Long-Term Accessibility of Digital Materials, 7-8 December, York, England, 2000.
   http://www.rlg.org/events/pres-2000/shenton.html

2. The British Library. About us: Policies. http://www.bl.uk/about/policies/digital.html

3. The British Library. Collections: Early printed.
   http://www.bl.uk/collections/early/holding bindings.html

4. The British Library. Collections: Early printed.
   http://www.bl.uk/ collections/early/evani on.html

5. The British Library. The Gutenberg bible.
   http://.prodigi.bl.uk/gutenbg/background. asp

6. The British Library. Collections: Treasures.
   http://www.bl.uk/collections/treasures/abo ut.html

7. National Library of Australia. Digital collections.
   http://www.nla.gov.au/digicoll/

8. Woodyard, Deborah. Digital preservation: The Australian experience. Paper presented at the Third Digital Library Conference on Positioning the Fountain of Knowledge, 2-4 October 2000, State Library of Sarawak, Malaysia.
   http://www.nla.gov.au/nla/staffpaper/ dw001004.html#abstract

9. Haddad, Peter. The next years in national libraries: The National Library of Australia. *Alexandria,* 2000, **12**(2), 126-27.

10. Library of Congress, National Digital Information Infrastructure and Preservation Program. Building a national strategy for preservation: Issues in digital media archiving. April 2002.
    http://www.clir.org/pubs/reports/pub106/ contents.html

11. Library of Congress. American Memory: Historical collections for the National Digital Library.
    http://memory.loc.gov/ammem/amhome. html

12. Tennant, Roy. The Berkeley Digital Library SunSITE. *D-Lib Magazine*, February 1996.

http://www.dlib.org/dlib/February96/ucb/0 2tennant.html

13. The Berkeley Digital Library SunSITE. http://sunsite.berkeley.edu/

14. California Heritage Collection. http://sunsite.berkeley.edu/CalHeritage/

15. Royan, B. Cross-domain access to digitised cultural resources: The SCRAN project. Paper presented at the sixty fourth IFLA General Conference, 16-21 August, 1998. http://www.ifla.org/IV/ifla64/039-109e.htm

16. TASI (Technical Advisory Service for Images). SCRAN case study. http://www.tasi.ac.uk/resources/scran. html

17. Royan, B. Content creation and rights management: experiences of SCRAN (the Scottish Cultural Resources Access Network). *Program*, April 2000, **34**(2). http://www.scran.ac.uk/articles/ipr.pdf

18. Royan, B. Scotland in Europe: SCRAN as a marquette for the european cultural heritage network. July 2000 http://www.cultivate-int.org/issue1/scran/

19. Oxford Digital Library. http://www.odl.ox.ac.uk/

20. Oxford Digital Library. Strategic objectives and practical steps on the way to a new core of service of Oxford University Libraries. http://heds.herts.ac.uk/conf2002/heds2002_lossau.pdf

21. Lee, Stuart D. Scoping the future of the University of Oxford's digital library collections: Final report. http://bodley.ox.ac.uk/scoping/report.html

22. Barnes, Susan. Sidebar1: Selected digital library projects in the Cornell University Library System. *Library Hi Tech,* 1999, **47** (**123**), 32-7.

Contributors: **Dr. Rochna Srivastava,** Commonwealth Fellow, Department of Information Science, Loughborough University, Loughborough, Leicestershire LE11 3TU UK. e-mail: rochna1@rediffmail.com