ISSN 0001 253X
Volume **58** Number **1/2** 2006

Aslib Proceedings

New information perspectives

British library and information schools: towards 100 years of educating the information professional at UCL SLAIS

Guest Editors: Andy Dawson and David Brown

Aslib imi
INFORMATION MANAGEMENT

www.emeraldinsight.com

Emerald

ISSN 0001-253X

Volume 58
Number 1/2
2006

Aslib Proceedings: New Information Perspectives

British library and information schools: towards 100 years of educating the information professional at UCL SLAIS

Guest Editors
Andy Dawson and David Brown

CONTENTS

CONTENTS
continued

Access this journal online

www.emeraldinsight.com/ap.htm

As a subscriber to this journal, you can benefit from instant, electronic access to this title via Emerald Fulltext or Emerald Management Xtra. Your access includes a variety of features that increase the value of your journal subscription.

How to access this journal electronically

To benefit from electronic access to this journal, please contact **support@emeraldinsight.com** A set of login details will then be provided to you. Should you wish to access via IP, please provide these details in your e-mail. Once registration is completed, your institution will have instant access to all articles through the journal's Table of Contents page at **www.emeraldinsight.com/0001-253X.htm** More information about the journal is also available at **www.emeraldinsight.com/ap.htm**

Our liberal institution-wide licence allows everyone within your institution to access your journal electronically, making your subscription more cost-effective. Our web site has been designed to provide you with a comprehensive, simple system that needs only minimum administration. Access is available via IP authentication or username and password.

Key features of Emerald electronic journals

Automatic permission to make up to 25 copies of individual articles

This facility can be used for training purposes, course notes, seminars etc. This only applies to articles of which Emerald owns copyright. For further details visit **www.emeraldinsight.com/copyright**

Online publishing and archiving

As well as current volumes of the journal, you can also gain access to past volumes on the internet via Emerald Fulltext or Emerald Management Xtra. You can browse or search these databases for relevant articles.

Key readings

This feature provides abstracts of related articles chosen by the journal editor, selected to provide readers with current awareness of interesting articles from other publications in the field.

Non-article content

Material in our journals such as product information, industry trends, company news, conferences, etc. is available online and can be accessed by users.

Reference linking

Direct links from the journal article references to abstracts of the most influential articles cited. Where possible, this link is to the full text of the article.

E-mail an article

Allows users to e-mail links to relevant and interesting articles to another computer for later use, reference or printing purposes.

Structured abstracts

Emerald structured abstracts provide consistent, clear and informative summaries of the content of the articles, allowing faster evaluation of papers.

Additional complimentary services available

Your access includes a variety of features that add to the functionality and value of your journal subscription:

E-mail alert services

These services allow you to be kept up to date with the latest additions to the journal via e-mail, as soon as new material enters the database. Further information about the services available can be found at **www.emeraldinsight.com/alerts**

Connections

An online meeting place for the research community where researchers present their own work and interests and seek other researchers for future projects. Register yourself or search our database of researchers at **www.emeraldinsight.com/connections**

Emerald online training services

You can also access this journal online. Visit **www.emeraldinsight.com/training** and take an Emerald online tour to help you get the most from your subscription.

Choice of access

Electronic access to this journal is available via a number of channels. Our web site **www.emeraldinsight.com** is the recommended means of electronic access, as it provides fully searchable and value added access to the complete content of the journal. However, you can also access and search the article content of this journal through the following journal delivery services:

EBSCOHost Electronic Journals Service
ejournals.ebsco.com

Informatics J-Gate
www.j-gate.informindia.co.in

Ingenta
www.ingenta.com

Minerva Electronic Online Services
www.minerva.at

OCLC FirstSearch
www.oclc.org/firstsearch

SilverLinker
www.ovid.com

SwetsWise
www.swetswise.com

Emerald Customer Support

For customer support and technical help contact:
E-mail **support@emeraldinsight.com**
Web **www.emeraldinsight.com/customercharter**
Tel +44 (0) 1274 785278
Fax +44 (0) 1274 785204

EDITORIAL ADVISORY BOARD

Aslib Proceedings: New Information Perspectives
Vol. 58 No. 1/2, 2006
p. 4

0001-253X

Note from the editor

British library and information schools

This is the first in a series of special issues of *Aslib Proceedings: New Information Perspectives* devoted to the research output of British library and information schools. With the RAE two years off it gives institutions an opportunity to demonstrate the quality and range of their research and scholarship. It is fitting that we start with The UCL School of Library, Archive and Information Studies (SLAIS), as it was the first library school to be established. The next two schools in the series will be Information Management, Department of Applied Social Sciences, London Metropolitan University and the Department of Information Studies, University of Wales Aberystwyth. Others will follow.

David Nicholas

Aslib Proceedings: New Information Perspectives
Vol. 58 No. 1/2, 2006
p. 5

0001-253X

GUEST EDITORIAL

Towards 100 years of educating the information professions at UCL SLAIS

Andy Dawson
School of Library, Archive and Information Studies, University College London, London, UK, and
David Brown
The British Library, London, UK

Abstract

Purpose – The purpose of this article is to introduce the theme of this special issue which contains a selection of papers written by the staff of UCL SLAIS.

Design/methodology/approach – The history of the School is briefly described and its philosophy of taking a broad and inclusive approach to professional education is reviewed.

Findings – There is considerable benefit to professional education in an institution which incorporates all branches of information work, and retains a sense of the value of traditional skills in tandem with an exploitation of new technologies. The contributors to this issue reflect this both by the range of disciplines which they represent and by the range of topics, from the most traditional to the most modern, and demonstrate the synergy between them.

Originality/value – The article provides an introduction to the special issue and promotes the value of an inclusive approach to professional education in the information disciplines.

Keywords Libraries, Archiving, Records management, Information management, Professional education

Paper type Viewpoint

The University College London (UCL) School of Library, Archive and Information Studies (SLAIS) has a long history of providing professional training and education – indeed as the home of the British School of Librarianship established in the wake of the First World War in 1919 (Shepherd, 2004), it has the longest history of any existing UK higher education establishment providing such a service. As such, its activities have developed proactively through the years to reflect the evolving demands of the information professions, expanding from its original compass of librarianship to include every field of records, archives and information management and science. It established the first UK archival programme in 1947 (with Liverpool also starting such a programme in the same year); it was the first to introduce a professional development MSc in Information Science (as opposed to a first professional qualification) in the 1980s; and most recently it has established the new UCL Centre for Publishing to build on the work of the Centre for Information Behaviour and the Evaluation of Research (CIBER) group at SLAIS, whose Virtual Scholar research programme has already attracted over £1 million for research in electronic publishing and informatics (CIBER, 2005). Today SLAIS is the only UK school to embrace and teach all the information

Aslib Proceedings: New Information Perspectives
Vol. 58 No. 1/2, 2006
pp. 6-9

0001-253X
DOI 10.1108/00012530610649184

skills from the most traditional, such as palaeography and manuscript studies, to the most modern, such as web design and electronic publishing, across all the professional information disciplines, within a single department. It is in part to celebrate that fact, and in part to promulgate why we think such an approach is important, that this special issue of *Aslib Proceedings* has been assembled.

Over the last few decades, the rate of change in both practices and attitudes within the information professions has been very great, particularly with regard to its relationship with information technology (IT). Thirty years ago, Dr Denis Lewis (then head of Technical Intelligence and Library at ICI, and latterly director of Aslib) suggested that by the year 2000 "... the information scientist and the librarian will have gone the way of the brontosaurus ..." and be replaced by the direct utilisation of IT by users (Lewis, 1977). Although his "doomsday scenario" has been proven wrong, the uptake of IT into the everyday work of the professions has been massive, and has led to a great deal of expansion in information science- or computer-oriented programme content in most professional programmes in the UK – sometimes at the expense of the more "traditional" disciplines and skills.

Similarly in that timeframe, there has been much talk of the need for more integration between the information professions. In practice, despite the efforts made in the wake of the Saunders report and the formation of CILIP from the amalgamation of the Library Association and the Institute of Information Scientists, the various branches of the information industry remain remarkably independent despite the obvious commonality of the commodity in which they all deal – the organisation, description and effective use of information in all its varied forms. While the world of work increasingly sees opportunities for hybrid information workers combining different disciplinary skills (e.g. those needing to manage related archival, records and library/information systems in industry and commerce) it is still difficult for new entrants to the information professions to accrue a professional education in more than one branch of the information industry, which arguably leads to further insularity.

While fully embracing technological developments to the professions, and recognising their value, UCL SLAIS believes strongly that the "traditional" skills remain critical to both an understanding of and the successful implementation of IT applications in the information domains. It also believes that by having expertise and research in all the varied information professions gathered together in a single department, the possibilities for learning from each other and the benefits which can stem from the cross-fertilisation of ideas from the different disciplines is greatly increased. These themes, of the value and synergy both between the old and the new, and between the different branches of information work, are the ones which are embodied in the selection of papers presented in this special issue, which span considerations from the paper based world of information to the electronic, from physical to digital environments, and address theory which, even when stemming from one particular branch of the profession, has implications for all.

The issue falls broadly into two convenient "halves", with the first focusing ostensibly on the past and what might be considered some of the more traditional areas of study, and the second ostensibly addressing "the future" and more modern, electronic domains. In reality, however, the papers in both sections have considerable relevance and linkage with one another, and this distinction is largely an artificial one – a point nicely in keeping with the issue themes. The initial papers address

"traditional" issues of information description and management which still need to be addressed in an electronic environment, and the latter ones indicate valuable uses of technology in exploiting both ancient and modern materials more efficiently.

To set the scene in a little more detail, Shepherd reviews the historical development of archives and records management education in UK universities, and gives an overview of the current state of teaching and research in the area in British universities. The research section demonstrates the high degree of involvement with technology and cross-domain relevance of work in the field, and her concluding suggestions of possible future changes, while focused on that branch of the information profession, make interesting reading for comparison with the development and state of other sectors.

Foot's paper considers the development and role of historical bibliography and argues strongly the importance and relevance of its continuation in the face of the increasing transfer of information to electronic formats, and its importance for cultural historical purposes. Although her focus is on manuscripts and books as physical objects of historical study, the points she makes regarding the importance of original physical context should make us consider the similar ramifications for electronic documents of all kinds – how can we establish, and record as evidence, such a thing for such items? Bowman's investigations of description in cataloguing prior to ISBD reveal some interesting historical practices and his conclusions remarking on the primacy of the title page and its value given changes to retrieval are similarly of great relevance to an increasingly electronic environment.

One of the ever-present difficulties of modern-day information retrieval is that of lack of precision, and consequent information overload. Broughton examines the value of facet analysis and classification as an information organisation and retrieval aid, and specifically addresses its value in a "mechanised" environment, identifying its advantages over enumerative and pre-coordinate systems in automated retrieval, as well as for browsing and navigation. Gildersleeves reports on the ongoing research into the Improve Your Library (IYL) initiative and the impact of the Department for Education and Skills' self-evaluation toolkits for school libraries, specifically how the informal pre-research consultations have informed the development of the initial survey instruments.

Moving further into the electronic environment, Hockey looks at the history of text-based humanities computing and discusses the differences between digital and other kinds of information, and new methods of analysis, encoding, manipulation and exploitation of material in the digital medium, particularly in terms of markup languages, and sets the scene for the following papers which focus more on the application of computing technology. Terras reports the development of a computer system which has improved our ability to interpret and understand the Vindolanda scrolls – a perfect example of cutting edge computing techniques being put to good use working in tandem with ancient ink and stylus resources to reveal new information – and Huntington and Nicholas, and Williams and Gunter, both report examples of powerful log analysis methods being used to investigate and model user behaviour, in the first case quite specifically to improve the efficacy of menu design to improve usability, and in the second to establish a formal methodology for the interpretation of user behaviour.

Finally, Miller presents a (thankfully!) "non-technical" overview of some very powerful developments in formal knowledge representation techniques using Event Calculus which, while perhaps quite hard reading for the bulk of the readers of this publication, have some very significant potential for information handling by computer, and the development of knowledge-based information systems.

Despite the breadth of variety in the papers collated within this special edition, all of which have been submitted by the staff within the one institute of UCL SLAIS, it is very much our hope that it is the similarities, overlaps and synergies within them which will be perceived, rather than their apparent differences, and as the department moves towards that centenary of educating the information professions, we hope that the inclusive approach in which it believes, and which this collection is intended to represent and promote, will become an ever-stronger force in the information community.

References

CIBER (2005), "New publishing centre established at UCL SLAIS", press release, available at: www.slais.ucl.ac.uk/press?news=200510240644 (accessed 24 October 2005).

Lewis, D. (1977), "There won't be an information profession in AD 2000", *The Information Worker: Identity, Image and Potential; Proceedings of a One-day Joint Aslib/Institute of Information Scientists Conference, London, 2 November 1976*, Aslib/IIS, London, pp. 38-49.

Shepherd, E. (2004), "Towards professionalism? Archives and archivists in England in the 20th century", PhD thesis in Archive Studies, University College London, London.

About the authors

Andy Dawson is MSc Programme Director, School of Library, Archive and Information Studies, University College London, London, UK. He is the corresponding author and can be contacted at: andrew.dawson@ucl.ac.uk

David Brown is the Head, Publisher Relations, The British Library, London, UK, and can be contacted at: david.brown@bl.uk

Developing a new academic discipline

UCL's contribution to the research and teaching of archives and records management

Elizabeth Shepherd
School of Library, Archive and Information Studies, University College London, London, UK

Received 3 May 2005
Revised 12 September 2005
Accepted 29 September 2005

Abstract

Purpose – The purpose of this article is to examine the historical development of archives and records management education in universities in England and review the state of research and teaching in the discipline in 2005.

Design/methodology/approach – Using a framework that draws on sociological attributes, the main text provides a historical analysis derived from primary and secondary sources, together with a brief overview of current educational provision for the discipline.

Findings – The article finds that graduate education in archives and records management has developed in the UK over a period of 60 years (1947-2005) and is well established, with seven Master's-level programmes offered across England, Wales, Scotland and Ireland, in a variety of learning modes (full- and part-time, open learning, face-to-face) and from different perspectives and contexts (history, information science, digital preservation). The university research community in the discipline has developed more recently and needs to progress quickly and soundly to support the future intellectual life of the profession.

Research limitations/implications – The geographical scope is limited to the UK and focuses in particular on England.

Practical implications – The paper identifies some gaps in educational provision, for instance mid-career cross-domain research opportunities, and an apparent lack of interest in academic research by UK practitioners, which might be further investigated.

Originality/value – No study of the historical development of the academic discipline of archives and records management has previously been published.

Keywords Archives management, Records management, Education

Paper type Case study

Introduction

The archives and records management profession, in common with other professional work groups, requires a complex knowledge base of theory and intellectual technique to underpin its special expertise. Professional work is technical and based on systematic knowledge and training (Wilensky, 1964). Typically, professional education and training is long, specialised and strongly conceptual. Training enables specialised knowledge to be transmitted, and is based on ideas and concepts rather than on

Aslib Proceedings: New Information Perspectives
Vol. 58 No. 1/2, 2006
pp. 10-19

0001-253X
DOI 10.1108/00012530610648644

The article is based on an invited paper given at the International Council on Archives, EURBICA and Section on Archival Education, expert round table at the National Archives of Finland, Helsinki, 11-12 November 2004.

physical objects. In addition to learning skills and knowledge, trainees acquire understanding of the "occupational sub-culture", its values and norms. A training school for new practitioners is usually established, sometimes beginning independently, but always eventually seeking contact with universities, in order to develop standard terms of study, academic degrees and research programmes.

This paper examines the historical development of archives and records management education in English universities, with reference to some significant Welsh and Scottish developments, and provides an overview of the state of archives and records management teaching and research in the universities at the start of the twenty-first century. The conclusion suggests some future changes which might help the discipline to develop to its full potential.

Historical development of archives and records management education in England

In the nineteenth and early twentieth centuries British archivists were recruited to the Public Record Office (PRO) and the few local archives with general historical and classical skills and education and often underwent a period of in-house training (Shepherd, 2004). The need for formal training and examination for archivists had been recognised at least since the publication of the Report on Local Records of 1902 (Bryce, 1902). The Report recommended that custodians of local archives be trained in palaeography and records, which "postulates the existence of some school where the necessary training could be supplied". It recommended that "schools of palaeography should be encouraged at the universities to create the supply of archivists", on the model of the Ecole des Chartes in Paris, whose Director gave evidence to the Committee. Oxford and Cambridge Universities should be encouraged to teach palaeography and medieval history and PRO staff lent to local record offices to disseminate skills (Bryce, 1902).

The disciplines of palaeography, diplomatic, local history and librarianship developed in the universities from around 1900. The teaching of palaeography in the University of London, as an adjunct to historical research, began in 1896, when Dr Hubert Hall of the PRO held classes at the School of Economics. In the 1920s Hilary Jenkinson lectured in sources of English history at King's College, London and became a Reader in Diplomatic and English Archives in the 1930s (University College London Records Office, 1936). Teaching and research in diplomatic and palaeography at Oxford University emerged with the appointment of Reginald Lane Poole to a lectureship in diplomatic in 1897, "the first post of its kind in this country" (Major, 1968, p. 117). In the twentieth century, several very distinguished scholars followed him including V.H. Galbraith, returning from the PRO, C.R. Cheney, Kathleen Major, Neil Ker and Pierre Chaplais, who had studied at the Ecole des Chartes.

At the University of Liverpool, J. Ramsay Bryce Muir and Professor J.M. Mackay set up the School of Local History and Palaeography by 1902. It aimed to stimulate "the study, editing and publication of the history and records ... of the City of Liverpool" (Kelly, 1981, p. 107). In 1908, J.A. Twemlow was appointed Lecturer in Palaeography and Diplomatics at Liverpool. He was a graduate of Oxford and the Ecole des Chartes and spent several months each year in Rome as the PRO's representative at the Vatican Archives (Cantwell, 1991). He helped to re-establish the School of Local History and

Records in 1911, which was described as a "school on the lines of the Ecole des Chartes" (University of Liverpool Archives, 1909).

Formal arrangements were made in 1902 between the Library Association and the London School of Economics for the teaching of librarianship. In 1915 the arrangements lapsed, but immediately after the war an initiative by the Library Association and UCL, supported by the Carnegie Trust, led to the establishment of the first British School of Librarianship at UCL in 1919 (University of London Library, 1918-1919).

These allied disciplines offered skills to archivists and the academic departments acted as a home to new archive qualifications after 1947. Between 1947 and 1955 five universities established first professional qualifications for archivists[1]. Since then it has gradually become usual for archivists to complete a first degree, a period of practical experience and a one-year, university-based Diploma or Masters. Initially, the university Diplomas derived their core content from the examples of the Continental schools (Latin and English palaeography, diplomatic, administrative history, editing, listing and indexing, and archive administration). All of the programmes involved senior professionals as part-time lecturers and offered practical experience as well as academic instruction. Most relied on one or two full-time academic historians to oversee the subject and tutor the students. Between 1965 and 1980 the university programmes matured and changed. The discipline began, albeit in a small way, to develop academic standing through research for higher degrees in archives, available at the University of London from 1966, although they took some time to gain acceptance. In 1968 the archive tutor explained the lack of interest in archival research by saying that the "more restricted nature of the subject" compared with librarianship reduced the pool of candidates and that many archive students were "primarily interested in history and go on to do a thesis in that subject rather than in Archive Studies" (University of London Library, 1968-1974). Archive academics pursued scholarly research, mainly in allied subjects (diplomatic, historical bibliography, history). Michael Cook's appointment at Liverpool University in 1969 signalled a change of direction and his subsequent work in archive administration and description was the first substantive professional research in the UK.

The advent of records management as a distinct part of the archival discipline had two important consequences for education (Shepherd, 2004). First, Northumbria University began new programmes that took an entirely fresh approach to records. The programmes focused on records and information rather than historical archives and showed that this approach could be made to work in complement to the traditional archive qualifications. Second, the traditional programmes all modernised to take account of the new significance of records management. Although archival programmes began to include records management topics from the 1950s (when Peter Walne gave a few classes at Liverpool University), and more formally from 1970 (when a separate module in records management was taught by Felix Hull at UCL), records management only became a significant part of the curriculum after 1980. Some programmes, such as Liverpool University, evolved gradually, expanding the teaching of records management while retaining an essentially archival perspective. Others, such as UCL's SLAIS) after 1990, were altered more radically, embracing records and archives as equal partners.

The British programmes were shaped (and often constrained) by their place within the university. None was a free-standing national school (on the Continental models such as Marburg or the Ecole des Chartes). Some were in departments of history[2], some in library and information studies departments[3] and some attached to the university archive[4]. These administrative arrangements encouraged the academics concerned to focus on different aspects of the discipline: for example, Carr (1995) (in Bangor) (1995) and Oschinsky (1971) (in Liverpool) pursued research interests in history, Hare and McLeod (in Northumbria) (Hare *et al.*, 2001) concentrated on information aspects, while Sayers (1999) (UCL) was a diplomatist. This split the discipline from an academic viewpoint, resulting in a lack of national profile. Research grant awarding bodies failed to recognise the place of the discipline in their areas of interest. Higher education was subject to increasingly rigorous inspection of teaching (Teaching Quality Assessment/Subject Review in 1998-2001 in England (Quality Assurance Agency for Higher Education, 2000)) and research (Research Assessment Exercise periodically from 1986 (RAE, 2005). Archives and records management programmes and research did not fit neatly into the assessment categories, sometimes being included with history, sometimes with information studies and sometimes omitted entirely.

By 2005 archival education had radically altered. Some traditional programmes remained, a few closed, but most were transformed, putting emphasis on intellectual issues and on the wider professional context. New programmes began in new subjects (digital preservation at Glasgow University) or offered in new modes (by distance learning) and increasing the geographical spread (two new archives and records programmes began in Scotland in 2004). External pressures led to further developments such as training and continuing professional development programmes (at Northumbria and Liverpool) and the beginning of a research culture for the discipline (at UCL SLAIS, Northumbria and Glasgow).

Research in archives and records management

An academic discipline of archives and records management gradually developed in England through research. In the 1980s at Liverpool University, Michael Cook undertook funded research projects in archival description (Cook, 1986) and Elizabeth Danbury obtained grants for survey and listing projects for the European Economic Community and the European Coal and Steel Committee. Later, Liverpool and Northumbria Universities were successful in obtaining funds for practice-based research, such as Liverpool University's study of data collection in the archival domain (Williams, 2003). From the 1990s onwards, many archivists were involved in resource enhancement projects, including improvements to archival description standards, new cataloguing, conversion of existing finding aids into structured digital data which could be delivered and searched remotely, the development of the national archive network (NCA, 1998), digitisation of archives and related resources and significant improvements to access. Such projects produced excellent scholarly resources, but are outside the scope of this article.

In 1997-1998 Northumbria University took a lead in records management research when it was a partner in a European-funded REcords Curriculum PROject (RECPRO), which "developed proposals for change in archives and records management teaching that took into account ... the digital environment" (RECPRO, 2003; Valtonen *et al.*,

1998). This was followed by European Training in Electronic Records Management (E-TERM), also European-funded, which included UCL and Northumbria in a group of six partners from five European countries. E-TERM aimed "to design a trans-national vocational training course in the management of electronic records to meet the needs of administrators, information professionals, archivists and records managers" (E-TERM, 2003; Shepherd, 2000). Such collaborative research projects enabled the discipline to develop new skills and knowledge.

Northumbria University continued its research programme in the 2000s through the Records Management Research Group. A project assessing the impact of the international standard for records management ISO 15489 was funded by the Arts and Humanities Research Board (AHRB) (McLeod, 2003). The Joint Information Systems Committee (JISC) programme to develop records management practices in higher and further education institutions awarded Northumbria four projects: Managing student assessment records, Managing primary research data and records for research in higher education institutions, Developing records management programmes in further education institutions and The electronic records management training package (Northumbria University, 2003).

Glasgow University developed a track record in funded research, making use of its expertise in digital preservation and its skilled archival staff in the university archive and business records centre. Glasgow undertook a range of interdisciplinary research in humanities computing, focusing on four main areas: the relationship between digital and analogue objects, digital creation and storage, user evaluation and information retrieval (Glasgow University, 2005). Projects included the Electronic Resource Preservation and Access Network which disseminated information, best practice and skills development in the area of digital preservation of cultural heritage and scientific objects (erpanet, 2004), and the Digital Curation Centre (DCC, 2005).

UCL's SLAIS supported research projects including a major developing country records management study under Dr Anne Thurston in the 1980s, funded by Leverhulme. In 2001-2004 Professor Susan Hockey directed the Linking Encoded Archival Description to Electronically Retrievable Sources (LEADERS) project, funded by the AHRB (LEADERS, 2003). It developed a web-based demonstrator system to bring together encoded archival finding aids, transcriptions of records, contextual information on the persons and organisations involved and digital images for the specialist user. LEADERS utilised three XML encoding systems as a basis for developing a generic toolset: encoded archival description (EAD) for finding aids, encoded archival context (EAC) for authority records and the text encoding initiative (TEI) for transcripts. In addition, LEADERS developed a segmentation model for profiling types of users of archival sources (Yeo *et al.*, 2004).

Information policy legislation and the e-government target of 2004 (The National Archives, 2001) placed archives and records management firmly on the government's agenda, but also increased the need for research to enable the domain to respond effectively. There was a research opportunity to evaluate and assess the impact of policy and practice within the widening arena of information held globally, which would help the profession to meet major challenges including the long-term preservation and management of digital records and information policy requirements. In response, UCL SLAIS established in 2005 an international centre for research and evaluation of key issues in the discipline of archives and records management

(ICARUS) to develop knowledge and enhance understanding of the creation, management and use of records and their role in society, and map, monitor and evaluate significant changes in the archives and records domain using robust, evidence-based methods. Initial projects under way include studies of the implementation of the Freedom of Information Act (Screene, 2005; Shepherd, 2005) and impact assessment.

The university research community in archives and records management in the UK is in the early stages of development, but it needs to progress quickly and soundly to ensure that it is capable of competing in the increasingly global academy. Universities also need to inculcate a sense of the value of research for the proper future expansion of the discipline among employers, individual professionals, academics in other disciplines and with policy bodies so that they accept the validity of academic research in archives and records management and support its development.

Present provision of archives and records management education

The UK archival domain work force needs educational opportunities at many levels (MLA, 2003), from volunteers to high-flyer, strategic and leadership candidates. Historically most university programmes were taught as first professional qualifications at graduate level. In 2005 full and part-time one-year Masters programmes in archives and records management were available at seven universities[5], two of which began in 2004. They range from those that focus on records management (as at Northumbria and Aberystwyth), to those which integrate records and archives (at UCL SLAIS) to traditional archival programmes (in Dublin). Some are taught face-to-face on campus and three are available by distance learning. All fulfil (or aspire to fulfil) the requirements of the Society of Archivists accreditation criteria (Rider, 1996).

Until recently no universities offered undergraduate programmes in archives or records management, even as a minor subject, although course units were offered to students studying other disciplines. For example, the BSc in Information Management at UCL had a compulsory first year unit on records management from 1995. Since 1999, Liverpool University has offered an undergraduate Diploma or Certificate in Professional Studies: Records and Information Management, by distance learning that takes between one and four years to complete. The programme was developed in conjunction with The National Archives, UCL and the University of Northumbria, and was delivered (with Northumbria until 2005) as the rm^3 partnership. The programme is aimed at government records staff, but is suitable for people working within a public sector records management environment (LUCAS, 2004).

No university programmes are targeted at unqualified clerical workers and volunteers, although thematic study days and summer schools are sometimes offered. Skills development tends to be offered in-house or by short courses by professional bodies or regional groups. Most volunteers work in archives and cultural services: not many volunteers work in records management, except as part of the work experience requirements for entry to a professional programme at university.

Archivists and records managers seeking mid-career qualifications usually focus on management skills such as those offered in a Diploma in management studies or Master of Business Administration (MBA). Increasingly records professionals need to understand principles and practices in other domains, such as libraries, museums, risk

management, or corporate governance, but there are few educational opportunities to study across the discipline boundaries. UCL's SLAIS offers an MRes in Library, Archive and Information Studies for librarians, archivists, records managers, museum curators and other information professionals who already have a degree and a first professional qualification. It gives students the opportunity both to develop their research skills before registering for a higher degree and to follow a flexible but guided programme of study to develop leadership, management, information technology and professional skills, including modules from the Centre for the Advancement of Learning and Teaching at UCL.

One route for development of strategic and leadership candidates is through a research degree. A PhD in Archive Studies has been available in the University of London since 1966. The Universities of Liverpool, Northumbria, Glasgow and Wales at Aberystwyth each also offer PhD and research programmes. In the 1980s and 1990s a steady number of research students came from overseas, in particular to study at UCL with Dr Anne Thurston (from Ghana, Canada, Swaziland, Turkey, Sri Lanka, Kenya and Botswana). In 2005 half a dozen students from four different countries were registered for archives and records management research at UCL: few of these students were from the UK (UCL SLAIS, 2005). In the UK a research degree is not of clear value to records practitioners for career progression, so there is little incentive for mid-career professionals to obtain one. The prospective development of "professional" doctorates with some taught components and a research thesis may make this a more attractive route (AHRC, 2005).

An alternative route to educational development for leadership is the Clore Leadership Programme, an initiative designed to help develop potential leaders in the cultural sector (Clore Leadership Programme, 2004). The MLA supported three fellowships in the first (2004) group, one each for libraries, museums and archives. The programme recruits between 20 and 30 fellows each year who undertake an individual programme of activities and experience, including a secondment to a cultural service different from the one in which they work, a two-week leadership course, tuition in business and management skills and a research project. The first archives fellow undertook a research project into the management of digital archives, supported by the UK Data Archive at Essex University; future fellows might be attached to active research groups at other universities.

Conclusion

Archives and records management as an academic discipline is still in a state of formation and is not yet stabilised (Forsyth and Danisiewicz, 1985). Archival graduate education sets parameters for professional work, defines the range of the profession, provides a gateway (and barrier) for entry and lays the foundations of career development. Archives and records management has benefited from quality structured university programmes for 60 years. However, the close association between archival education taught at universities and practitioners, especially those at the former Public Record Office, in some ways inhibited the development of an academic discipline. Until at least the 1970s, research into archives and records management was not considered appropriate or necessary: research was undertaken in allied disciplines such as history or diplomatics. Very few academic posts in archives and records management existed until recently; the handful of universities offering the subject in the UK in the 1980s

each had a single academic post. Even in 2005, there were fewer than 20 UK academics in archives and records management. However, universities have now established strategies to ensure that they develop research programmes and promote the intellectual development of the discipline, while maintaining high quality, professionally-focused, teaching programmes, especially at Masters level. Gradually a cadre of UK archival educators is emerging who are qualified both academically (PhD) and professionally (MA) to undertake teaching, research and publication. Groups such as the Forum for Archives and Records Management Education and Research (FARMER, 2005) and the Network of North Western European Archival Educators provide essential academic networks. The first UK conference for doctoral research in archives and records management was held in June 2005, sponsored by the AHRC, a mark of the need to foster research in the discipline but also of its potential and vitality.

Notes

1. University College London, University of Liverpool, Bodleian Library Oxford, University of Wales at Aberystwyth and at Bangor.
2. University of Liverpool, University of Wales at Bangor and, for most of its life, at Aberystwyth.
3. University College London, Loughborough University, Northumbria University.
4. Bodleian Library Oxford, University College Dublin, and, initially, University of Glasgow.
5. University of Liverpool, Master/Diploma/Certificate of Archives and Records Management (MARM); University College London, MA in Archives and Records Management; University of Wales, Aberystwyth, Archive Administration (MSc Econ / Diploma) and Records Management (MSc Econ / Diploma) both also available by distance learning; Northumbria University, Records Management MSc by distance learning; University College Dublin, Higher Diploma in Archival Studies; Glasgow University, MSc in Information Management and Preservation (Archives and Records Management); University of Dundee, Archive and Records Management (ARM) Distance Learning Programme.

References

Arts and Humanities Research Council (AHRC) (2005), "Working group on the UK doctorate in the arts and humanities", available at: www.ahrc.ac.uk/university_staff/postgrad/working_group_on_the_uk_doctorate_in_the_arts_humanities.asp# (accessed 29 April 2005).

Bryce, J. Chair (1902), *Report of the Committee Appointed to Enquire as to the Existing Arrangements for the Collection and Custody of Local Records and as to Further Measures which It May Be Advisable to Take for the Purpose*, Cd 1335, HMSO, London.

Cantwell, J. (1991), *The Public Record Office 1838-1958*, HMSO, London.

Carr, A.D. (1995), *Medieval Wales*, Macmillan, Basingstoke.

Clore Leadership Programme (2004), "Programme overview: a general overview of the inititative", available at: www.cloreleadership.org/programme.htm (accessed 22 March 2005).

Cook, M. (1986), *Manual of Archival Description*, Gower, Aldershot.

Digital Curation Centre (DCC) (2005), available at: www.dcc.ac.uk/ (accessed 22 September 2005).

Electronic Resource Preservation and Access Network (ERPANET) (2004), available at: www.erpanet.org/ (accessed 24 August 2005).

European Training in Electronic Records Management (E-TERM) (2003), available at: www.ucl.ac.uk/e-term/ (accessed 29 April 2005).

Forsyth, P.B. and Danisiewicz, T.J. (1985), "Toward a theory of professionalization", *Work and Occupations*, Vol. 12 No. 1, pp. 59-76.

Forum for Archives and Records Management Education and Research (FARMER) (2005), available at: www.liv.ac.uk/lucas/FARMER/ (accessed 29 April 2005).

Glasgow University (2005), "HATII (Humanities and Advanced Technology and Information Institute)", available at: www.hatii.arts.gla.ac.uk/ (accessed 25 August 2005).

Hare, C., McLeod, J. and Whitman, J. (2001), "BIAP: balancing information access and privacy", *Journal of the Society of Archivists*, Vol. 22 No. 2, pp. 253-74.

Kelly, T. (1981), *For Advancement of Learning: The University of Liverpool 1881-1981*, University Press, Liverpool.

Linking Encoded Archival Description to Electronically Retrievable Sources (LEADERS) (2003), available at: www.ucl.ac.uk/leaders-project/ (accessed 29 April 2005).

Liverpool University Centre for Archive Studies (LUCAS) (2004), "Courses and programmes", available at: www.liv.ac.uk/lucas/courses.htm (accessed 3 November 2004).

McLeod, J. (2003), "Assessing the impact of ISO 15489 – a preliminary investigation", *Records Management Journal*, Vol. 13 No. 2, pp. 70-82.

Major, K. (1968), "The teaching and study of diplomatic in England", *Archives*, Vol. 8 No. 39, pp. 114-18.

Museums, Libraries and Archives Council (MLA) (2003), "Discussion paper 6a: investing in the future: training and development: a further paper", available at: www.mla.gov.uk/action/archives/atf_meetings.asp#4 (accessed 2 November 2004).

(The) National Archives (2001), "E-government policy framework for electronic records management", available at: www.nationalarchives.gov.uk/electronicrecords/pdf/egov_framework.pdf (accessed 10 February 2005).

National Council on Archives (NCA) (1998), *Archives On-line*, NCA, London.

Northumbria University (2003), "Records management research", available at: http://online.northumbria.ac.uk/faculties/art/information_studies/Imri/rarea/rm/rmresearch.htm (accessed 29 April 2005).

Oschinsky, D. (Ed.) (1971), *Walter of Henley and Other Treatises on Estate Management and Accountancy*, OUP, Oxford.

Quality Assurance Agency for Higher Education (2000), *Subject Review Handbook 2000-2001*, QAA, Gloucester.

RAE (2005), "Research assessment exercise 2008", available at: http://rae.ac.uk (accessed 22 August 2005).

RECPRO (2003), available at: http://online.northumbria.ac.uk/faculties/art/information_studies/Imri/rarea/rm/rmproj/rmcomp/recpro/recpro.htm (accessed 29 April 2005).

Rider, C. (1996), "Developing standards for professional education: the Society of Archivists' accreditation criteria", *Journal of the Society of Archivists*, Vol. 17 No. 1, pp. 85-95.

Sayers, J.E. (Ed.) (1999), *Original Papal Documents in England and Wales from the Accession of Pope Innocent III to the Death of Pope Benedict XI*, OUP, Oxford.

Screene, L. (2005), "How prepared are public bodies for the implementation of the UK Freedom of Information Act in January 2005?", *Records Management Journal*, Vol. 15 No. 1, pp. 34-42.

Shepherd, E. (2000), "Report on the proposal for a European training project for administrators, archivists and information managers: E-TERM", *Proceedings of the DLM Forum on Electronic Records. European Citizens and Electronic Information: The Memory of the Information Society, Brussels, 18-19 October 1999*, Office for Official Publications of the European Communities, Luxembourg, pp. 244-9.

Shepherd, E. (2004), "Towards professionalism? Archives and archivists in England in the twentieth century", PhD in Archive Studies, University of London, London.

Shepherd, E. (2005), "Freedom of information and records management in the UK: what has been the impact so far?", paper presented at the DLM-Forum, Budapest, 5-7 October.

University College London Records Office (1936), *Printed Report Prepared for the University of London Committee on Palaeography, June 1936, School of Librarianship, Miscellaneous 1933/34*, file 18/3, University College London Records Office, London.

University College London, School of Library, Archive and Information Studies (UCL SLAIS) (2005), "Research at SLAIS", available at: www.slais.ucl.ac.uk/research (accessed 29 April 2005).

University of Liverpool Archives (1909), *Printed Report of Public Meeting January 1909*, file D.399/1/3, University of Liverpool Archives, Liverpool.

University of London Library (1918-1919), *Correspondence between Carnegie Trust, LA and Vice-Chancellor of University of London about Funding: University Central File, General Correspondence, School of Librarianship and Library Association 1918-1919*, file CF/1/19/208, University of London Library, London.

University of London Library (1968-1974), *Board of Studies in Librarianship and Archive Administration Minutes 1968-74*, file AC 8/34/1/2, University of London Library, London.

Valtonen, M., Hare, C., Horsman, P. and Schokenhoff, V. (1998), "RECPRO – developing a European records management programme", *Records Management Journal*, Vol. 8 No. 3, pp. 55-61.

Wilensky, H. (1964), "The professionalization of everyone?", *American Journal of Sociology*, Vol. 70 No. 2, pp. 137-58.

Williams, C. (2003), "Data collection and management in the archival domain", *Journal of the Society of Archivists*, Vol. 24 No. 1, pp. 65-81.

Yeo, G., Sexton, A., Turner, C. and Hockey, S. (2004), "Understanding users: a prerequisite for developing new technologies", *Journal of the Society of Archivists*, Vol. 25 No. 1, pp. 33-49.

Corresponding author
Elizabeth Shepherd can be contacted at: e.shepherd@ucl.ac.uk

The study of books

Mirjam M. Foot
School of Library, Archive and Information Studies, University College London, London, UK

Received 18 April 2005
Revised 26 July 2005
Accepted 21 September 2005

Abstract

Purpose – This paper aims to show how the concept of "Bibliography" has changed since the late nineteenth century. It proposes discussing what "Bibliography" did and did not include in the various stages of its development; how the study of "Bibliography" moved from the UK to the USA; how it narrowed down from an originally much wider concept and how, under the influence of French historians over the past three decades, it has widened out again, reaching a better synthesis of the study of books as material objects with the study of the history of the book.

Design/methodology/approach – A discussion and critical assessment of the writings of the major main stream bibliographers and book historians is presented.

Findings – From an original (nineteenth century) emphasis on enumerative bibliography, the concept of "Bibliography" widened out (from the end of the nineteenth century) to include historical bibliography and the study of books as material objects; in the mid-twentieth century this wider approach narrowed down, as a consequence of much emphasis being placed on descriptive, analytical, critical and textual bibliography. Under influence of French book historians the emphasis has changed again and the value of a wider historical approach and greater inclusivity in subjects has brought the study of historical bibliography and that of the history of the book much closer together.

Research limitations/implications – This research looks only at Western Europe and the USA.

Practical implications – Practical implications of this study are: the widening-out of the subject to include all physical manifestations of the book; the dimension of creative reading; and the emphasis on the importance of artifactual evidence for correct establishment and interpretation of texts has had implications for preservation.

Originality/value – This paper is a critical assessment of the literature, drawing the logical consequences of its findings. It presents an argument for the inclusion of all aspects of the book as a physical object, as well as for the importance of using all available evidence.

Keywords Bibliographies, History

Paper type Literature review

On 21 October 1920, A.W. Pollard wrote to George Watson Cole:

> I gave an inaugural lecture this very week as Professor of English Bibliography in the University of London. Greg followed this afternoon with one on the differences & similarities between MSS & Printed Books. McKerrow carries on for 3 more lectures on Elizabethan printing habits, and Dover Wilson winds up with one on the printing of plays. So in London at any rate Bibliography is recognised as a University subject (Cole Papers, quoted in Stoddard, 2000).

Here, in one short letter from one pillar of bibliographical studies and a world-class librarian to another across the Atlantic, are five giants who have shaped the development of bibliography into what Greg (1933, p. 271) came to call "the study of books as tangible objects".

Now 86 years later "Historical Bibliography" is still taught at UCL, but in the intervening period the subject has changed its focus, and expanded from what was

Aslib Proceedings: New Information Perspectives
Vol. 58 No. 1/2, 2006
pp. 20-33

0001-253X
DOI 10.1108/00012530610648653

mainly descriptive, analytical and textual bibliography to a much wider "History of the Book". How the concept of "Bibliography" changed, and what it did and did not include in the various stages of its development will be the main subject of this paper.

Pollard and his colleagues were not the first to extend the concept of bibliography beyond the sometimes mundane, but essential exercise of compiling lists of works by specific authors, on specific subjects, or from specific presses or publishers, nor did they ignore this enumerative aspect. Even Greg (1930, p. 259), who in 1930 called the compilation of bibliographies by bibliographers "mere prostitution", later justified enumerative bibliography on the grounds that "Before books can be studied they must be known to exist and survive" (Greg, 1945, p. 25). The very basis for bibliographical studies, for the study of books and for the study of the history of books, is formed by enumerative bibliography – not least because it permits, in D.F. McKenzie's words, "the resurrection of the most marginal texts and their makers . . . and thereby the study of all who were kept from the centres of power . . . [which] opens up the possibility of a far more comprehensive reconstruction of cultural history" (McKenzie, 1992, p. 296).

Pollard (n.d., quoted in Dover Wilson, 1948) saw "the business of the bibliographer [as] primarily and essentially . . . the enumeration of books . . . the lowly task of finding out what books exist, and thereby helping to secure their preservation". But Pollard's own concept of bibliography was much wider. Already in 1903 he realised that:

> [. . .] so long as literature in order to be communicated has to take material form . . . every point which concerns this material form should be carefully and thoroughly investigated.

And ten years later:

> Between us and the author of any old book stand scribes or printers, publishers, and even binders, and until we have eliminated the errors due to these we cannot reach the true text (Brown and Pollard, 1903, pp. 158, 161; Pollard, quoted in Sadleir, 1945, p. 157).

Henry Bradshaw, Librarian of Cambridge University Library from 1867 to 1886, whom Pollard called "a master of method", was one of the first English scholars to stress the importance of the structure of books and the relation this bears to their texts; he also used the evidence of printing types to reach bibliographical conclusions and drew attention to what he called "the habits of printers". In order to describe the book's structure, Bradshaw developed the collational formula (in 1861), essentially as it is used to-day (Needham, 1988, p. 9 and *passim*).

Bradshaw, in reaction to the amateur approach to the study of literature in particular, and influenced by the new interest in science of his own time, stressed the "scientific" nature of what was in essence analytical and descriptive bibliography, a notion which has caused far too much ink to flow and one on which minds and pens far sharper than mine have been blunted. This categorisation proved a red herring in the end, and was finally dismissed by Thomas Tanselle in 1974:

> Whether bibliography can be defined as a "science" . . . is of less importance than understanding . . . what in fact it does, what its methods of procedure are, what its strengths and weaknesses may be (Tanselle, 1974, p. 73).

Shades of A.W. Pollard, who, after a wrangle between Gaselee and Greg in *The Library* (1932-1933) about defining bibliography, concluded that, although definitions differ:

> We all want the same things: the question at issue is as to whether we want them as *bibliographers*, or in some other capacity. My own preference is for a big umbrella (Pollard, 1932a, p. 258).

Pollard's own work, and the subjects it covered – from early illustrated books and copyright, to the transmission of Shakespeare's texts, Chaucer and the technical description of incunabula – demonstrated his concept of bibliography as a big umbrella and his students gained by his catholic views. Notwithstanding a bad stammer, which made public speaking an ordeal, he taught at King's College London, where he became in 1919 "honorary Professor in Bibliography", a post he held until 1932. His self-perception showed in a letter to Falconer Madan:

> I never called myself Professor while I was one – as it was a very side show – and now I've given it up (Pollard, 1932b).

George Watson Cole, the recipient of Pollard's letter of 1920, with which I began, was taught by Melvil Dewey (of the DD system) and encouraged by him to become a librarian. He was the first professional librarian of the library of Henry E. Huntington and was one of the first, at least in the USA, to apply scholarly standards to bibliographic work. He moved between bibliographical study and historical research, occasionally combining the two, thereby perhaps pointing the way towards what we now call the history of the book.

Cole's notion of what bibliography encompasses widened out from his somewhat restricted view in 1916, when he stated that for the bibliographer "the only perfect book is the one caught on its way from the printer's office to the binder's" (Cole, 1916, p. 124), to that four years later, when he realised that the bibliographer needs to understand not only "the art of printing and be able to detect in every instance the points in which the printer has mistaken the ideas of the author", but that he must also "be able to detect the places where the binder has mistaken the order in which the printer intended the various parts of the book to be bound", showing that he was clearly prepared to look beyond the printed sheets (Cole, 1920, p. 15). More striking, when we consider the focus of the work of Greg, McKerrow and their successors, was Cole's insistence on treating books not only "as products of the printing-press but as emanations from the minds of their authors and of their standing in the history of literature", and – one step further still:

> Interest is added if some details can be given concerning the author … and incidents connected with the writing of the book or with its publication and public reception (Cole, 1990, p. 171, 1916, p. 141).

The real importance of the latter had to wait for about 70 years before being fully acknowledged and expanded.

Cole's vision was expressed in an article in the *New York Herald Tribune* of 1925:

> The realm of books is boundless and … resembles the universe … The work of the bibliographer … may … be compared with that of the astronomer … there are constantly more and still more worlds being created for him to conquer – those already in existence and those that have yet to see the light.

How Cole would have relished the challenge of developing bibliographical methods for the study of electronic transmission of ideas.

Three more English bibliographers are referred to in Pollard's letter: Greg, McKerrow and Dover Wilson. The latter has from time to time had a bad press as being too fanciful, too imaginative, not "scientific" enough. Bowers accused him of speculative textual insight and historical imagination, but he also received acclaim for his innovative work on the complete edition of Shakespeare. He comes through as a most engaging character and certainly one with wide appeal. His *What Happens in Hamlet* (Dover Wilson, 1935) reads almost like a novel. Dover Wilson was one of the advocates of Greg's "new bibliography" and its relevance to editorial theory. He paid attention to palaeography and to the role of copytext, and also saw the importance of stage directions, which he was not averse to re-writing.

McKerrow's "lectures on Elizabethan printing", mentioned by Pollard, may well have led to his Sandars lectures on the printing of Elizabethan plays (1928). He lectured in bibliography at King's during the First World War until 1919, and when Pollard became professor, he resurrected a short technical course on the subject. He was a great educator, and indeed, his *An Introduction to Bibliography for Literary Students* (McKerrow, 1994), first published in 1927, still educates the library students at UCL (and elsewhere) today. Although this indispensable work focuses mainly on English printing in the sexteenth and seventeenth centuries, McKerrow's interests were somewhat wider. In 1921 he had suggested that bibliography should be extended to European literature and culture, as he wanted to get away from the notion that England is the centre of the Western world. Moreover, he argued for the importance of the study of social history. Six years later he stated that a book should be considered "from the points of view of those who composed, corrected, printed, folded, and bound it; in short, so that [it is seen] not only as a unit but as an assemblage of parts, each of which is the result of a clearly apprehended series of processes" (McKerrow, 1994, p. 4). Nevertheless, elsewhere in the same book, he dismissed binding, marketing, reading patterns and other aspects of the relationship between authors, publishers and readers, which he considered not to bear directly on the mechanics of textual transmission.

McKerrow's strict contemporary at Trinity College, Cambridge, was W.W. Greg, one of the most famous – and most discussed – figures in English bibliography. Although Greg called Pollard "the real pioneer" of the "new bibliography", it is Greg's own name that has been most firmly associated with this movement (Greg, 1945, p. 28; Wilson, 1945, p. 76). The theme of the lecture referred to in Pollard's letter, the differences and similarities between MSS and printed books, is one that pervades much of Greg's work. In a paper to the Bibliographical Society of 1912, he regarded "the distinction between written and printed books as irrelevant", calling bibliography "a system of investigation and a method of description . . . [which] can be made to apply to clay cylinders and rolls of papyrus as well as to codices", applicable "to the transmission of all symbolic representation of speech or other ordered sound or even of logical thought" (Greg, 1912, p. 42). This important point has been expanded by later bibliographers, notably by McKenzie (1986, p. 33):

> [. . .] think not of books as the only form of textual artefact, but of texts of many different kinds in many different material forms, only some of which are books or documents.

Any signs that express meanings form a "text". Greg considered bibliography to include "the study of book-making and of the manufacture of the materials of which books are made . . . a knowledge of the conditions of transcription and reproduction, of

the methods of printing and binding, of the practices of publication and bookselling – … the whole of typography and the whole of palaeography … [all] processes that leave their mark on the character of the finished book", but, he added, "to the bibliographer the literary contents of a book is irrelevant" (Greg, 1912, pp. 44-6). Greg's concept of bibliography and what it does and does not include seems to have moved through various stages, from his liberal concept of 1912 and 1920 to a narrower view in the 1930s, emphasising the importance of books "only as the vehicle by which [their] contents reach us" (Greg, 1930, p. 250), while two years later he suspected the study of bookbinding to be "mere antiquarianism", calling it a "purely bookish art", not a factor in the transmission of texts. Having been dismissive of one branch of bibliographical studies, admittedly one that at the time was still in its infancy, Greg (1932, pp. 118, 121) showed his open-mindedness by declaring himself open to "light and guiding from every branch of bookish lore".

Throughout his work he stressed the importance of the history of the text: "for an author … has been and has meant something different and individual to each ensuing age". He also gave what is perhaps still the best definition of bibliography, concluding that:

> Bibliography … deals with books as more or less organic assemblages of sheets of paper, or vellum, or whatever material they consist of, covered with certain conventional but not arbitrary signs, and with the relation of the signs in one book to those in another (Greg, 1933, pp. 270-1).

In 1945, in the Bibliographical Society's *Studies in Retrospect*, Greg both expanded his earlier views and contracted them:

> […] the object of bibliographical study is … to reconstruct for each particular book the history of its life, to make it reveal in its most intimate detail the story of its birth and adventures as the material vehicle of the living word.

However, an earlier-expressed limitation is still apparent: "The ornamentation of a manuscript or a printed book may notably enhance its beauty …; it cannot alter in any fundamental way the character of the book itself or the lessons it is able to teach" (Greg, 1945, pp. 26-7), something we would now consider a regrettable blind spot. The character of a book is changed noticeably by the addition of illustration, ornamentation and even by its binding. It reaches out to a different public and fulfils a different role. The lessons a book is able to teach can be enhanced enormously by ornamentation, and also by the structural particularities and decoration of its binding. These lessons are not only related to the book's provenance – as Greg did acknowledge – but they explain for us the reason for the book's production, the function it was meant to fulfil and the public it was aimed at.

Greg was not only a prolific writer and influential speaker, he also aspired to be a teacher. As early as 1912 he described "a dream of my own. It is of a course of lectures on English bibliography which may one day be delivered at one of our so-called seats of learning."(Greg, 1912, p. 50). Eighteen years later he came back to this, although the course he had dreamt of had not yet been delivered: "there has now been a bibliographical chair in the University of London for over ten years … [filled by] Alfred Pollard". He criticised the fact that it was a chair of "English bibliography" … "to suppose that bibliography can be parcelled out into linguistic, literary, or regional departments is to harbour a dangerous fallacy … [Nevertheless,] London not only

boasts a professorship, but has also, at University College London, a lectureship, which for several years has been held with distinction by Mr. Esdaile. Its subject, however, is what I should describe as librarianship" (Greg, 1930, pp. 242-3).

Indeed, Arundell Esdaile, one of the founding fathers of the first British School of Librarianship, established at UCL after the Great War, was in 1919 invited to become a part-time lecturer in bibliography there, a post he held for 20 years. His lectures, published in 1931 as *A Student's Manual of Bibliography*, did not cover historical bibliography, even though Esdaile recognised the importance of the subject.

Greg himself did get the opportunity to fulfil his dream of 1912. In his *Biographical Notes, 1877-1947* (Oxford, 1960), he wrote: "I did ... accept in 1932, and with some misgivings, an honorary lectureship in bibliography at University College, London. (The College wanted to make it an honorary professorship – like that held by Pollard at King's – but found they had no power to do so under the new statutes.) ... I found", said Greg, "that my difficulty was not to know what to say or how to say it, but to stop".

Theodore Besterman, known especially for his *World Bibliography of Bibliographies* (Oxford, 1939-1940), a work that went through many revisions and editions, also lectured at the School of Librarianship at UCL (1931-1938). He was followed by Frank, later Sir Frank, Francis, one-time director and principal librarian of the British Museum, for many years a leading figure in the Bibliographical Society, and a champion of the scholarly pursuit of knowledge. His bibliographical interests ranged widely and those who knew him will remember him as infinitely approachable. He taught bibliography from 1945-1959 and was succeeded by Howard Nixon, also a prominent figure in the Bibliographical Society and at the British Museum; a librarian with an extensive knowledge of, and a profound love for, rare books; an inspiring and enthusiastic teacher and a prolific writer on the subject he had made especially his own: the history of English and European bookbinding.

Others were influential too and their work as teachers and bibliographers has left its mark. Nicolas Barker, of catholic tastes in bibliographical studies and achievements; David McKitterick, the most recent historian of the Cambridge University Press; Paul Quarrie who certainly would have pleased Greg by not only focusing on English bibliography; and Robin Alston who, in several – at the time ground-breaking – lectures, applied the study of bibliography to electronic formats.

Over the decades, the course has been expanded into a more rounded history of the book, giving equal emphasis to the history of printing, paper making, and bookbinding, with some excursions into the history of publishing, the booktrade, and collecting. Descriptive and analytical bibliography, however, are still important components.

In 1990 Robin Alston proved conclusively that, for the study of post-1964 books, the book itself no longer reveals its own production methods, so that archival evidence is essential (Alston, 1990), a point of view that can also be applied to the hand-produced book. This was shown already in the 1930s, when several bibliographers, grouped round John Johnson and Strickland Gibson in Oxford, concerned themselves with wider historical evidence, including booktrade documents – Graham Pollard was a particular advocate for the importance of the history of the trade. They extended bibliography from the history of the making of books to that of their publishing and

distribution and they laid the foundation for the study of the history of the book in Britain.

Nevertheless, the struggle between the old school of bibliographers for whom the book itself was the beginning and the end, epitomised by Fredson Bowers, and that of a younger generation who, influenced by the work of French historians, such as Lucien Febvre and Henri-Jean Martin, believed that archival evidence could be equally important, is still continuing, although a better synthesis of the study of books as material objects with the study of the history of the book is now being reached.

Fredson Bowers, although accused of being both narrow and dogmatic, was a great scholar who, in direct succession to Pollard, McKerrow and Greg, moved the centre of bibliographical studies from England to the USA. His *Principles of Bibliographical Description*, first published by the Princeton University Press in 1949, has influenced generations of bibliographers – and still does. Thomas G. Tanselle called it "a manual for the writing of... the history of the published forms of books as physical objects", with the aim of reconstructing past events and thoughts through focusing on the objects that have survived to "give us a tangible link with a previous time" (Tanselle, 1994, p. xiii). Like Greg, Bowers held that the general principles of bibliography have a wide application: "No matter what the field of study, the basis lies in the analysis of the records in printed or manuscript form". To Bowers, one of the main purposes of descriptive bibliography was to create "a detailed, analytical record of the physical characteristics of a book ... as a trustworthy source of identification". He emphasised the importance of examining "every available copy of an edition of a book in order to describe in bibliographical terms the characteristics of an ideal copy of this edition", and he made a strict separation between those bibliographical facts that are the province of analytical bibliography, which can be derived solely from inspection and analysis of the book as a physical object, and what he called "collateral facts", such as printers' and publishers' records, archival evidence, inscriptions in books and the "evidence ... of a presentation copy" (Bowers, 1994, pp. xv, 1-34). In his later work Bowers continued to emphasise the separation between "the physical characteristics of the sheets of a book" and what he classed as "ancillary" material (Bowers, 1953, p. 5).

Thomas Tanselle, Bowers's closest disciple, came to accept, more than Bowers did, the value of evidence external to the book. At the same time he held that if this external evidence, which he called "secondary evidence", is not supported by the evidence found in the book itself, which is the "primary evidence", it should be dismissed. Here he differed from D.F. McKenzie who pointed out that bibliographers have analysed books in isolation and taken them out of the context of production in the printer's shop. Both Bowers and Tanselle observed the evidence presented by the book itself and postulated this into a pattern or a "normal" way of behaviour: "expectations of regularity based on past experience", as Tanselle (1974, p. 75) called it. They both agreed that the amount of evidence is vital, hence the need to inspect every copy of every edition, while McKenzie, looking at external evidence and realising that there simply is not enough evidence to postulate "normality", concluded that the "unpredictable complexity" is such that the norm does not exist. "The idea that a 'pattern' must be significant because it appears to indicate a regular method of work", warned McKenzie, "is one of the most perniciously seductive presuppositions of current bibliographical analysis" (McKenzie, 1969, pp. 7, 38). A view also held by McGann (1983).

With hind sight, the issue is not so much which kind of evidence is "primary" evidence, but how extensive it is, how complex and varied, and whether evidence of all kinds has been correctly understood and interpreted. The acceptance of the importance of all external, archival evidence, as well as all evidence that can be derived from analysis of the book itself, is where the bibliographer and the book historian meet. But included in the evidence as presented by the book or document as a physical object is much of what Bowers, and indeed Greg before him, considered to be "collateral" or "ancillary" evidence. Meanwhile several bibliographers took different views and especially those who chose the history of the booktrade or that of bookbinding as a field of study, included in their investigations not only anything they could derive from minute observation of the book itself, but also made a first attempt at placing the book in a wider historical and cultural context. E.P. Goldschmidt considered the history of bookbinding as part of European cultural history, while Graham Pollard (1960) felt himself concerned with "the Book Efficient and the Book Economic", using evidence provided by binding structure as indicators of how books were sold. Historians of the book, such as Robert Darnton (1983), extended the function of the book as a physical object beyond its text, drawing attention to the significance of books for the societies in which they were produced and read.

Nevertheless, for main stream bibliographers, the text remained the focus of attention. Bowers (1994, p. 9) saw "true bibliography" as "the bridge to textual, which is to say literary, criticism", a position which scholars such as John Carter, have criticised:

> Scholarship in this context [i.e. bibliography] surely does not end with the mere text: it embraces book-structure, publishing practice, copyright regulations and the remuneration of authorship, distribution methods and reading habits (Carter, 1952, p. 80; quoted by Bowers, 1953, p. 3).

Earlier, Sadleir (1945, p. 147) had exclaimed:

> [. . .] we want to know . . . what were the changing processes, trade customs, methods of distribution and book-buyers' tastes which made . . . [the books] what they are; . . . and [we] want to weave them into the lives of their authors, producing . . . a garment of biography with a texture of bibliography.

Bowers himself gradually got away from his restricted approach when in 1953 he declared that although he held "in highest estimation that form of bibliography which is intended to illuminate textual problems", he was also "unwilling to see analytical bibliography limited only to a direct application to textual problems". But he never abandoned the importance of the text: "If the history of our culture is ultimately the history of the writings of our great men", we must ensure "that we are reading them aright and not in corrupt texts" (Bowers, 1953, p. 4; 1966, p. 25), a point that was developed by Tanselle in 1981:

> It is hard to see how historians or philosophers could claim that their purposes are served by any less meticulous recording of textual variants and editorial alterations than that employed in . . . literary works . . . [The] widespread feeling that historians, philosophers, sociologists – all scholars, in fact . . . – read for content and need not be concerned with the niceties of punctuation or other formal features, . . . [is of a] *naïveté* [that] . . . is astounding.

Form and content are inseparable and "a reader cannot be satisfied with any text that conceals or obscures … evidence" (Tanselle, 1981, pp. 7-8).

McKenzie also maintained the importance of the text: "our knowledge of the past derives mainly from texts", but he demanded more than the books themselves: the "history of the book should help to explain how these particular texts were created, why they took the form they did, their relations with other media, … and what influence they had on the minds and actions of those who heard, read or viewed them" (McKenzie *et al.*, 1999). Earlier, in his 1986 Panizzi lectures, McKenzie had stressed the importance of dissemination and readership.

Since G.W. Cole's interest in public reception of texts in 1916, bibliographers had not paid much attention to this. There were, of course, exceptions. Wytze Hellinga, the Dutch bibliographer, concerned himself with the history of books "as bearers of *content*, which traces how the product came into being on the press and how it finally went about fulfilling its function in its readers' hands" (Hellinga, 1962, p. xix).

In his Panizzi lectures McKenzie redefined what bibliography is and does and extended it to include the "meaning [of texts] for, and … their creative regeneration by, readers". "Readers", said McKenzie (1986, pp. 10, 20) "inevitably make their own meanings"; he went further: "new readers … make new texts, and … their new meanings are a function of their new forms". Jerome McGann (1983) also took up this point. Book-historians, especially in France, have emphasised the role of the reader in re-creating the identity of the text and so have others. In the early 1960s, Professor Toyama argued that the text transforms itself through the audience's added action (see Davison, 1992, p. 7). This goes perhaps a little far, but if we consider the reception of electronic texts, the reader, in formatting the transmitted files and thereby creating the format that will affect the content, is assuming a role of crucial importance.

Whatever the role of the recipient, we must not forget that every text has an author and it would be regrettable if the baby of authorial intention were to be thrown out with the bathwater of the "social dynamics of … reception". What writers thought they were doing in writing texts, or printers and booksellers in designing, producing, publishing and disseminating them, or readers in making sense out of them, are all issues of concern to the bibliographer and to the historian in their attempt to understand the book as a cultural force (McKenzie, 1986, pp. 8, 10).

How important are the various forms in which the author's thoughts are transmitted, recreated and received? In a series of articles that appeared between 1981 and 1998, Tanselle, laying the connection between what he called "verbal works" and their "tangible representations in physical objects", came back again and again to the inseparability of form and content and to the way one affects the other. "Verbal works … come to us tied to objects … the study of such works … cannot be separated from the study of artefacts" (Tanselle, 1981, 4, pp. ix, xiii), a point McKenzie (1986, p. 4) made in 1986: "bibliographers should be concerned to show that forms effect meaning".

Tanselle (1998b, pp. 39-40) has been one of the most outspoken advocates for books as objects and already in 1998, in an analogy with museum objects, he posed:

> When an object does not display any words or writing, it is easier to see that form and content are one and that the form cannot be altered without changing what the object communicates … Books … contain words … and because the same words can be printed on different physical backgrounds, people … think that the message conveyed is independent of the

> vehicle carrying the words . . . the history of the transmission of texts shows conclusively that the content of texts is affected by the mechanical processes of transmitting them.

This is equally true for the transmission of thought in electronic format. Here too are intangible works at the mercy of tangible processes of transmission.

Taking on board this further dimension of creative reading, strengthens the argument that many bibliographers have adduced for the indispensable evidence provided by every copy of every edition of every book, irrespective of its age or of its means of production.

In 1953 Bowers wrote:

> From a strictly bibliographical point of view, which considers all books only as they are tangible objects, the twentieth edition is just as interesting and even as important as the first, since it is a part of the total publishing and textual history of the author (Bowers, 1953, pp. 7-8).

Tanselle (1998c, p. 95) went further:

> Book-reading is the activity of reconstructing the texts of intangible works from the physical evidences for them found in books. [Therefore] Every physical characteristic of every copy of every edition is potentially relevant for the activity of reading.

In 1995 the Modern Languages Association stated that:

> Not only do editions differ from one another, but also copies within an edition (of any period) often vary among themselves; as a result, every copy is a potential source for new physical evidence, and no copy is superfluous for studying an edition's production history . . . The loss of any copy of any edition . . . diminishes the body of evidence on which historical understanding depends (Tanselle, 1998d, p. 336).

The logical conclusion to all this is that:

> The importance of artifactual evidence to the reading of every genre of verbal text means that the objects containing verbal texts in all fields must be preserved as assiduously as other artefacts (Tanselle, 1998a, Preface, p. x).

But there is an additional argument for this. Ruth Perry stated in 1993:

> [. . .] the object itself is part of the material culture of its own time, and carries with it something of the social context that produced it (Perry, 1993, p. 61).

Already in the 1960s Professor Hellinga emphasised to his students the importance of reading texts in the context in which they had been produced. McKenzie (1993, p. 25) stressed that:

> [. . .] original artefacts [are] products of distinctive contexts", and it is only in their own context that we can interpret them.

The inseparability of the text from its physical form as a product of, and functioning in, its social context, thus forming firm historical evidence, confirms the imperative need to preserve it as an object. As a principle this is sound, but as the responsibility for such preservation rests with public institutions, archives and libraries, this has implications.

Most archivists and librarians, and many, although by no means all, bibliographers, realise that it is unrealistic to expect that all currently surviving books, manuscripts

and other documents can be saved and kept indefinitely in their original physical format, as their survival depends on the materials of which they are made and "on the nature of events that befall them". In a purely practical sense, it also depends on money and resources, but, as the MLA pointed out "the attitudes that people hold about them can be instrumental in either mitigating or exacerbating the destructive effects of these factors" (Tanselle, 1998c, p. 336).

Decisions about priorities for preservation still have to be made, but they should be considered in the recognition of the artifactual value of each object. This does not lighten the burden placed on the archivist and librarian and it increases the burden on those who try to educate them to work in a twenty-first century environment. Librarians have not had a good press lately. They have been accused – by novelists, journalists, by scholars and by fellow librarians – of vandalism, ignorance, obsession with space saving, in charge even of deliberate destruction. It is fashionable among library administrators to envisage a future with smaller book stacks and much information transferred to electronic format. Robert Zich's "Idols in the Library" described the stacks in the Library of Congress as "emptied of most routine books and magazines and other routine paper" (Zich, 1990). Book stacks may be gradually abandoned as texts are incorporated in machine-readable databases. Book-centred librarianship is seen as outmoded and special-collections librarians as an insignificant element in the structure of information retrieval; librarians as managers, information providers, hell-bent on "access", ignorant of the fact that access to physical evidence is an essential kind of access, and, to quote Tanselle (1998e, p. 9) once more, convinced that "the cultural heritage transmitted in books can survive independently of the physical objects".

This is not new. In 1937 Randolph G. Adams devoted an article in the *Library Quarterly* to "Librarians as enemies of books" (Adams, 1937), while Fredson Bowers on more than one occasion wrote sharply and amusingly of his own unfortunate encounters with academic librarians who wanted him "to know that we are not interested in rare books here" or wished that "you would give me one good reason why a library should keep a duplicate!" Bowers, – quite rightly – blamed library schools, who "do not know enough to teach analytical bibliography … and this deficiency sometimes also means, unfortunately, that books cannot be recorded accurately in libraries" (quoted in Tanselle, 1985, p. 7; Bowers, 1966). In 1966 he exclaimed: "of all forms of librarianship, the training of the rare-book librarian has been most neglected" and he continued to develop the argument for training librarians as scholars rather than as technicians, referring – with some sadness – to "the great days of English librarianship [when] the librarians were often better scholars than many of the persons whom they served". These days have gone and Bowers (1966, pp. 1-3, 9, 26) called the "tendency [of librarians] to become managers and technicians … dissociated from the active work of scholarship … most unfortunate". Paul Raabe, the great German librarian who built up the book-history-centred studies at the Herzog August Bibliothek in Wolfenbüttel, confirmed: "the librarian who is not a scholar cannot effectively promote scholarship" (Raabe, 1983, p. 252).

Bowers lamented the lack of attention paid to the physical book in library schools and Tanselle saw as the reason for the closure of so many library schools in the USA, that they had switched their attention from "objects" to "information", thereby competing, unsuccessfully, with strong computing and management departments in

the universities. If "the traditional emphasis of library schools had been different, they would not now be regarded as dispensable. Their neglect of rare-book training, their failure to insist on it for all librarians", according to Tanselle, "is the key; ... Library schools and their graduates over the past century have not established a climate in which people can see why the study of books as physical objects is central to cultural history". And he regretted the "failure to recognise the relationship between the form and the content of books" (Tanselle, 1998e, p. 19, b, p. 24).

Tanselle wrote about libraries and library schools in the USA, but we, in this country, are rather inclined to follow in the footsteps of our trans-Atlantic colleagues, and we would do well to heed their example as a warning.

We are fortunate at UCL that manuscripts and rare-books librarianship are taught, and not as second-rate or outmoded subjects, but we must not forget that – as the twenty-first century moves on and as the focus is ever more on the role of libraries to facilitate the exchange of information – that information cannot be exchanged without sufficient attention being paid to the artefacts that provide the physical means for its transmission. If we understand the uses of the artefact and its vital role in the shaping and transmission of ideas, we will also realise that the role of the "book" librarians (by which I mean all librarians concerned with physical objects, be they clay tablets, scrolls, codices, written or printed) in the future will be ever more important. The existing books and manuscripts in library stacks must never be abandoned, because they will remain crucial as the original pieces of evidence of their own production and as evidence of the creation, transmission and reception of ideas and their impact on society. Recognising this takes nothing from the enthusiastic embracing of the technology for electronic dissemination of texts. The wide accessibility, the sophisticated search options, and the great convenience of reception are very welcome, but they come at the price of removing such texts from their original physical context with the consequent loss of evidence. Originals can never be rendered irrelevant, however inconvenient this may be. Bibliography and book history combined in "the study of books" has taught us that.

References

Adams, R.G. (1937), "Librarians as enemies of books", *Library Quarterly*, Vol. 7, pp. 317-31.

Alston, R.C. (1990), "Bibliography, computers and research", lecture delivered to the London Bibliographical Society, available at: www.r-alston.co.uk/essay2.htm (accessed 3 May 2005).

Bowers, F. (1953), "Purposes of descriptive bibliography", *The Library*, Vol. VIII No. 1, 5th series, pp. 1-22.

Bowers, F. (1966), *Bibliography and Modern Librarianship*, University of California, Berkeley, CA.

Bowers, F. (1994), *Principles of Bibliographical Description*, Princeton University Press, Princeton, NJ (originally published in 1949).

Brown, J.D. and Pollard, A.W. (1903), "Practical bibliography", *The Library*, Vol. 4, new series, pp. 144-62.

Carter, J. (1952), "Some bibliographical agenda", in Ray, G.N., Weber, C.J. and Carter, J. (Eds), *Nineteenth Century English Books*, University of Illinois Press, Urbana, IL.

Cole, G.W. (1916), "Bibliographical problems with a few solutions", *Papers of the Bibliographical Society of America*, Vol. 10, July, pp. 119-42.

Cole, G.W. (1920), "Bibliography – a forecast", *Papers of the Bibliographical Society of America*, Vol. 14, pp. 1-19.

Cole, G.W. (1990), "Bibliographical method", in Dickinson, D.C. and Cole, G.W. (Eds), *1850-1935*, Scarecrow Press, Metuchen, NJ, pp. 165-72 (reprint, originally published in 1929).

Darnton, R. (1983), "What is the history of books?", in Carpenter, K.E. (Ed.), *Books and Society in History*, Bowker, New York, NY/London, pp. 3-26.

Davison, P. (1992), "Introduction", in Davison, P. (Ed.), *The Book Encompassed*, Cambridge University Press, Cambridge.

Dover Wilson, J. (1935), *What Happens in Hamlet?*, Cambridge University Press, Cambridge.

Dover Wilson, J. (1948), "Alfred William Pollard 1859-1944", *Proceedings of the British Academy*, Vol. XXX.

Greg, W.W. (1912), "What is bibliography?", *Transactions of the Bibliographical Society*, Vol. XII, pp. 39-53.

Greg, W.W. (1930), "The present position of bibliography", *The Library*, Vol. XI No. 3, 4th series, pp. 241-62.

Greg, W.W. (1932), "Bibliography – an apologia", *The Library*, Vol. XIII No. 2, 4th series, pp. 113-43.

Greg, W.W. (1933), "The function of bibliography in literary criticism", *Neophilologus*, Vol. XVIII.

Greg, W.W. (1945), "Bibliography – a retrospect", *The Bibliographical Society 1982-1942*, Studies in Retrospect, Bibliographical Society, London, pp. 23-31.

Hellinga, W.G. (1962), *Copy and Print in The Netherlands*, North-Holland Publishing, Amsterdam.

McGann, J.M. (1983), *A Critique of Modern Textual Criticism*, University of Chicago Press, Chicago, IL/London.

McKenzie, D.F. (1969), "Printers of the mind", *Studies in Bibliography*, Vol. 22, pp. 1-76.

McKenzie, D.F. (1986), *Bibliography and the Sociology of Texts*, British Library, London.

McKenzie, D.F. (1992), "History of the book", in Davison, P. (Ed.), *The Book Encompassed*, Cambridge University Press, Cambridge, pp. 290-301.

McKenzie, D.F. (1993), "What's past is prologue", Bibliographical Society Centenary Lecture, 14 July 1992, Hearthstone Publications, London.

McKenzie, D.F., McKitterick, D.J. and Willison, I.R. (1999), "The Cambridge history of the book in Britain", foreword, in Hellinga, L. and Trapp, J.B. (Eds), *The History of the Book in Britain*, Vol. III, Cambridge University Press, Cambridge.

McKerrow, R.B. (1994), *An Introduction to Bibliography for Literary Students*, new ed., St Paul's Bibliographies/Oak Knoll Press, Winchester/New Castle, DE (originally published in 1927).

Needham, P. (1988), *The Bradshaw Method*, University of North Carolina, Chapel Hill, NC.

Perry, R. (1993), "Embodied knowledge", *Harvard Library Bulletin*, Vol. 4 No. 1, Spring, pp. 57-62.

Pollard, A.W. (1932a), "'Note' at the end of Gaselee, S., 'The aims of bibliography'", *The Library*, Vol. XIII No. 3, December, 4th series, pp. 255-8.

Pollard, A.W. (1932b), letter to F. Madan, dd. 23 August 1932, Bibliographical Society Archives, London.

Pollard, G. (1960), "Commentaries on the physical form of books" three unpublished lectures, May, typescript, MS Pollard 284, fols. 68-127, Bodleian Library, Oxford, pp. 68-127.

Raabe, P. (1983), "Library history and the history of books", in Carpenter, K.E. (Ed.), *Books and Society in History*, Bowker, New York, NY/London, pp. 251-4.

Sadleir, M. (1945), "The development during the last fifty years of bibliographical study of books of the XIXth century", *The Bibliographical Society 1982-1942, Studies in Retrospect*, Bibliographical Society, London, pp. 146-58.

Stoddard, R.E. (2000), "BSA presidential address 'Dear Lawrence', 'Dear Bill': William A. Jackson, Lawrence C. Wroth, and the practice of bibliography in America", *Papers of the Bibliographical Society of America*, Vol. 94 No. 4, December.

Tanselle, G.T. (1974), "Bibliography and science", *Studies in Bibliography*, Vol. 27, pp. 55-90.

Tanselle, G.T. (1981), *The History of Books as a Field of Study*, University of North Carolina, Chapel Hill, NC.

Tanselle, G.T. (1985), "The achievement of Fredson Bowers", in Tanselle, G.T. (Ed.), *Fredson Bowers at Eighty*, Bibliographical Society of America, New York, NY, pp. 3-18.

Tanselle, G.T. (1994), "Introduction to: Bowers", in Bowers, F. (Ed.), *Principles of Bibliographical Description*, St Paul's Bibliographies/Oak Knoll Press, Winchester/New Castle, DE.

Tanselle, G.T. (1998a), *Literature and Artifacts*, Bibliographical Society of Virginia, Charlottesville, VA (collection of essays, 1977-1996).

Tanselle, G.T. (1998b), "Bibliographers and the library", in Tanselle, G.T. (Ed.), *Literature and Artifacts* (collection of essays, 1977-1996), Bibliographical Society of Virginia, Charlottesville, VA, pp. 1-56.

Tanselle, G.T. (1998c), "The latest forms of book-burning", in Tanselle, G.T. (Ed.), *Literature and Artifacts* (collection of essays, 1977-1996, originally published in 1993), Bibliographical Society of Virginia, Charlottesville, VA, pp. 89-95.

Tanselle, G.T. (1998d), "Statement on the significance of primary records adopted by the executive council of the MLA of America, 19th May 1995", in Tanselle, G.T. (Ed.), *Literature and Artifacts* (collection of essays, 1977-1996, originally published in 1995), pp. 335-7.

Tanselle, G.T. (1998e), "Libraries, museums, and reading", in Tanselle, G.T. (Ed.), *Literature and Artifacts* (collection of essays, 1977-1996, originally published in 1991), Bibliographical Society of Virginia, Charlottesville, VA, pp. 3-23.

Wilson, F.P. (1945), "Shakespeare and the 'New Bibliography'", *The Bibliographical Society 1982-1942*, Studies in Retrospect, Bibliographical Society, London, pp. 76-135.

Zich, R. (1990), "Idols in the library", in Cole, J.Y. (Ed.), *Research Collections in the Information Age*, Library of Congress, Washington, DC, pp. 10-14.

Corresponding author
Mirjam M. Foot can be contacted at: m.foot@ucl.ac.uk

The development of description in cataloguing prior to ISBD

J.H. Bowman
School of Library, Archive and Information Studies, University College London, London, UK

Received 9 August 2005
Revised 22 September 2005
Accepted 29 September 2005

Abstract

Purpose – The purpose of this article is to compare Anglo-American cataloguing codes and practices for description over the past 150 years and assess the contribution that they made to International Standard Bibliographic Description (ISBD).

Design/methodology/approach – The major international codes, and those of major British libraries, are examined point by point, using as a basis the main areas of description as outlined in AACR2. Cataloguing textbooks are also referred to when appropriate.

Findings – The article finds that general order of elements has been remarkably constant throughout the period, most variation being seen in the physical description area. Primacy of the title page as a source of information is long established. Publisher's name was often of minor importance, and in public libraries physical description was greatly reduced. It is clear from wording that codes often adopted rules from one another, and evidently some libraries made a great attempt to adopt the latest thinking. Solutions to some problems evolved gradually over a long period. Even some apparently minor aspects of punctuation have a long pedigree. ISBD clearly drew on a long-established consensus as far as possible.

Practical implications – In revising AACR2 it is important to be able to see how we have reached where we are now, and in particular to avoid repetition of past mistakes. In an increasingly international publishing environment it is vital to solve the problem of multiple places of publication.

Originality/value – This subject has never been tackled in this way before, and the findings are timely for the ongoing revision of AACR2.

Keywords Cataloguing, Information management

Paper type Research paper

Introduction

Although there have been several histories and comparisons of cataloguing rules, these have tended to focus on entry points and headings, rather than on how the body of the description was done. Hanson (1939), for example, explicitly excludes description, and Norris (1939) has merely a few mentions of the topic. Strout (1956, p. 272) makes only brief reference to transcribing the title. Henderson (1976) provides an excellent account of descriptive cataloguing in general, but includes hardly any detail. Gorman (1989) deals only with specific aspects of entry points, headings and card catalogues. Hitchler (1903) is the only comparative study to include both access points and description and to go into detail; her work is very useful, but it is brief, it includes several codes which are hardly remembered now, and of course it is over 100 years old. Hagler (1977) gives an excellent overview of description but, with the exception of one set of comparative examples, does not go into specific detail. More recently, Blake (2002) has included a few passing references to description, but there has been no comprehensive treatment. This is hardly surprising given that, until the arrival of the online catalogue – and still

Aslib Proceedings: New Information Perspectives
Vol. 58 No. 1/2, 2006
pp. 34-48

0001-253X
DOI 10.1108/00012530610648662

in AACR2 –, it was the entry points that allowed readers to retrieve material. Description was probably regarded by most librarians as very much secondary, perhaps even as "a rather disagreeable and tedious necessity" (Gorman (1971, p. 3), although obviously not his own opinion). An exception was Charles Coffin Jewett, librarian of the Smithsonian Institution in Washington from 1848 to 1858. His desire to construct a general catalogue from separate units necessitated some form of standardisation, but the plan was short-lived (see London, 1980, pp. 256-62). At a time when work is proceeding on the revision of AACR3 to become Resource Description and Access it seems appropriate to review the history of the subject, and the purpose of this paper is to fill a gap by looking at how description has changed and to see which features were incorporated into ISBD. The emphasis is on British practice over the last 150 years or so, but much of this will inevitably reflect the American situation also. The discussion is restricted to books, which for most of the period under consideration have constituted the bulk of libraries' stock.

Our current AACR2 owes a great deal to International Standard Bibliographic Description (ISBD), which was the first serious international attempt to standardise the "description" part of cataloguing. It was an outcome of the International Meeting of Cataloguing Experts held in 1969 under the auspices of the IFLA Committee on Cataloguing, and its background is well summarised by Hagler (1977). The first ISBD to be published was that for monographic publications (ISBD(M)) in 1971 (IFLA, 2004, p. iii); the latest version of this was published in 2002 (IFLA, 2002).

All ISBDs, and therefore AACR2 also, break bibliographic description down into a number of "areas", as follows:

- title and statement of responsibility;
- edition;
- publication, distribution, etc.;
- physical description;
- series;
- note; and
- standard number and terms of availability.

We may ignore the last of these because for most of the period it is irrelevant, but otherwise it is appropriate to use this list as the basis of comparison. As will be shown, some elements of description have moved between areas over the years, but even in the earliest days most of these main areas can be identified in some form. There is also, as Gorman (1971, p. 4) found, a surprising degree of agreement regarding the order of the elements, and this extends to many other countries' codes. It is apparent that some of the differences from modern practice arise from the fact that in the period before automation the descriptive part of the catalogue record could never (with the exception of Jewett's proposed system) be seen in isolation from its heading, and this will be noticed several times.

The codes that I shall be using for the purpose of the comparison are listed in chronological order as follows, and they will be referred to using the specified abbreviations, followed by the relevant rule numbers:

- British Museum (1841) *Alphabetical Catalogue of Printed Books: Rules to Be Observed in Preparing and Entering Titles* (BM). (It is often more convenient to consult *Rules for Compiling the Catalogues of Printed Books, Maps and Music in the British Museum* (British Museum, 1936), but it must be remembered that this is a brief guide issued to help readers, and not a code for practical cataloguing.)
- Library Association. Cataloguing Rules of the Library Association of the United Kingdom (as revised at the Liverpool Meeting, 1883). Included in the *Library Association Year Book 1884*, (Library Association, 1884, pp. 65-8) (LA).
- Cutter (1904) *Rules for a Dictionary Catalog* (Cutter).
- *Cataloguing Rules: Author and Title Entries* (Committees of the Library Association and of the American Library Association, 1908) (AA).
- *Anglo-American Cataloguing Rules* (American Library Association *et al.*, 1967) (AACR67).
- *Anglo-American Cataloguing Rules*, 2nd ed. (American Library Association *et al.*, 1978) (AACR78).
- *Anglo-American Cataloguing Rules*, 2nd ed. 1888 revision (Gorman and Winkler, 1988) (AACR88).
- *Anglo-American Cataloguing Rules*, 2nd ed., 1998 revision Joint Steering Committee for the Revision of AACR, 1998) (AACR98).
- *Anglo-American Cataloguing Rules*, 2nd ed., 2002 revision (Joint Steering Committee for the Revision of AACR, 2002). (AACR2002) + updates 2003, 2004, 2005.

Note: when no differentiation is needed between the various revisions of the 2nd ed., this is referred to simply as AACR2.

Sources of information

An aspect of description, which has assumed great importance since the coming of ISBD, is the question of which parts of the publication the cataloguer is allowed to use as sources of information for the various areas of the description. Earlier rules are somewhat ambiguous about this, but we can nevertheless see a consensus from the outset that the title page had some kind of primacy. Moreover, some codes agreed on a distinction that does not exist today, that information taken from elsewhere in the work should go in round brackets, while additions from external sources went in square (see Cam 51 and 58). Bod for example gives no specific instruction on location of information for the title, although it states (36) that additions are to be enclosed in square brackets; Cutter (242) says the same. AA (136) refers to "an exact transcript of the title-page", but is unspecific about how far this applies to different parts of the description. LC (3:4) states that "The data given in the body of the entry are mainly those on the title-page of the work . . . " and it is clear from the examples that this is intended to apply to the whole of the description. One of the points of difference between the British and American editions of AACR67 was that the British edition (132A.1) allowed the edition statement, imprint and series statement to be taken from places other than the title-page without the use of square brackets, while the American edition did not. However, no specific list of alternative sources is provided. It was not

until the publication of ISBD that it became an important part of the cataloguing rules to specify exactly what sources could be used, and for these to vary in a clearly defined way according to the area of the description.

Title and statement of responsibility

Despite the undoubted importance of the title page, in the mid-nineteenth century cataloguers were much less scrupulous than their modern counterparts are expected to be in the extent to which they transcribed exactly the information that appeared on it. Not surprisingly, it was expected that the language of the original would be maintained and that the spelling would be copied (BM XVIII: "the original orthography to be preserved"), and this is something that has been carried through all the codes and into AACR2.

What about capitalisation? A quick glance at the British Museum *General Catalogue of Printed Books* reveals that the titles are delightfully full of capital letters in a way that would be quite unacceptable now; this catalogue is perhaps unique in this respect. There is no reference to capitals in BM 1841, and indeed the examples included there are almost entirely devoid of them. Chaplin (1987) says nothing about capitalisation, and it was presumably a practice which developed naturally, simply because it is natural to put capital letters in titles. Later rules were usually specific in its prohibition: Cutter (288), for example, says "Capitals are to be avoided, because in the short sentences of a catalog they confuse rather than help the eye", which seems as good a reason as any. He does, however, permit, but not prescribe, ("may" in rule 289) a capital letter for the first word after an article, to make filing clearer. AA also includes an appendix on capitals and other matters. Here only the first word is to have a capital letter, except that in title main entries the first word after an article should also have one (AA 172). This exception surprisingly enjoyed a brief revival in AACR78 (A.4D) and AACR88, applying also to series, and was cancelled only in the 1993 amendments.

Moving outside the rules and looking at other contemporary accounts, we find Wheatley (1889, pp. 155-9) giving several examples with capitals, although he makes no specific mention of them, and it is unclear whether he is simply following the capitalisation of the title pages in question. Quinn (1899, p. 21) states that the old custom of using capitals for most words has "fallen into disuse", the implication being that this was fairly recent. Jast (1899/1990, p. 120) specifically recommends capitals for the principal words:

> I am well aware that this is in direct antagonism to Cutter and other authorities, [. . .] but I must confess that uncapitalised titles always seem to me bizarre and ugly. And so I believe they strike everyone, save the librarian who has cultured a liking for them, as one cultures a taste for olives and tomatoes.

Edmond (1903, p. 78) partially anticipates AA by advocating a capital letter for the first word after an article; this would not have been restricted to title main entries, and its intention, as with Cutter, was to help filing. (This practice, in defiance of AACR67, was still adopted in the library where the author of this paper first worked as a graduate trainee in the mid-1970s.)

There was never an assumption that the whole of the title should be transcribed. Cutter (223-225) was very clear about abridgment of the title, saying "Many a title a yard long does not convey as much meaning as two well-chosen words." Much

depended on the nature of the library and its users. Initial articles should be omitted "when it can be done without altering the sense or too much offending the ear" (224), and in "short cataloguing" much more than this could be omitted, so that the title "Observations upon an alteration of the charter of the Bank of England" might be entered as "Alteration of charter of Bank of England". Cam (51), using words so similar to BM XVIII that they must be based on it, states that the title should be "expressed in as few words (and those only of the title-page) as may be needed to show what was meant to be conveyed, but especially so much as is needed to distinguish the work from another bearing the same title". (They might have added "and by the same author", as there was no way of retrieving a work by its title in this catalogue.)

Sometimes it was possible to cut down the title considerably. Jast (1900/1901) recommended omitting it altogether if the work was a biography, but this was only possible because in a dictionary catalogue the subject would be given as a heading. More generally, Cam (51 fn.) suggested that the title *The Letter of Charles Lamb* would be entered just as Letters, and as late as 1949 LC (3:5) prescribed: "The first words of the title are always included in the transcription except that an initial article is omitted if this will make it possible to omit the author statement that would otherwise be necessary only because of the grammatical structure of the statement." The result of this is that *The Works of Shakespeare* is transcribed simply as Works, and it was clearly not thought necessary to show that anything on the title-page had been omitted. All this was possible, of course, only because the description could never be seen in isolation from its main heading.

In some cases omissions were marked as such: three dots were to be used to mark the omission of "mottoes and non-essential matter" (AA 136; cf. LA 2); words "omitted from the *chief* part of the title" were to be marked with three dots (Bod 48) and similarly Cam (51), but these would not be used at the beginning or end of the title. BM clearly used "etc." for omissions from the end of a title. AACR67 (132B) allowed the abridgment, with dots, of long titles, and this is still the case in AACR2 (1.1B4).

At the other extreme we can consider cases where the wording of the title was expanded. An example of this arose (LC 3:5C) where a title in a foreign language started with a numeral, in which case the spelt-out version would be added in square brackets, for example:

> 100 [i. e. Honderd] jaar spoorwegen in Nederland.

The reason for this was of course filing, because the numerals would be expected to be filed as words in the appropriate language. This rule is not to be found in the other rules under consideration, but a relic of it survived as an option in AACR67 (Appendix IV, D), and judging by the number of records containing these expansions it must have been widespread (see Bowman, 2001, pp. 48-9).

Punctuation of the title was generally to be copied (AA 136) –or not copied (LA 1). The rule in AACR2 (in 1.1B1) about changing square brackets appearing in a title into round, and dots into a long dash, was anticipated by Cam 58, but otherwise the only recognition of this problem seems to have been by the Library of Congress, which used angle brackets < > for square in this situation (Cutter 242 fn.). LC (Appendix II, A.1.c) says that dots on the title-page should be represented by "bold face periods", which must have been quite hard to distinguish from ordinary dots.

Errors were to be marked with "[sic]" or "[!]" (Cutter 242, AA 138), by "[sic]" only (Bod 43, Cam 51), or by "[sic]" or "[i. e.]" followed by a correction (LC 3:4). The exclamation mark has gone out of use, but otherwise there has been no change in AACR2 (1.0F1) except that since AACR67 (132A.2) it has been possible to interpolate missing letters in square brackets, and that "i.e." now has no space between the letters.

Since AACR78 we have had *Eileen Ford's A* [or, from 1993 a] *More Beautiful You in Twenty-one Days* entered in full with no omissions, but one cannot imagine any of the early cataloguers doing anything so obtuse as to include the author's name as part of the title in this way. Most such titles would probably have been treated like *The Works of Charles Lamb* mentioned above and shortened as much as possible. On the other hand, the ambiguity that arises here is shown by the fact that in inflected languages these words might be retained; there are many examples of the type *Homeri opera* in the British Museum catalogue, for example.

Other title information

Moving on to what we now call "other title information", Cam (62) makes a nice distinction between "an explanatory and alternative title", which is introduced by a colon, and "words that are explanatory only", the latter being introduced by a semi-colon. Examples given are:

ROBERTS (Sydney Castle). Doctor Johnson in Cambridge: essays in Boswellian imitation.

CHATTERLEIGH (Harriet). Hearts of Oak; being a history of the British Navy simply told.

Such a distinction must have been very hard to make in practice. Quinn (1899, p. 21) uses the semi-colon for an alternative title and the colon for a sub-title. Jast (1899/1990, p. 119) advocates the colon "between main and second titles" (i.e. sub-titles, or "other title information"), and the semi-colon to separate different treatises bound together. In the latter respect he was to some extent unconsciously anticipating AACR2 (1.1G3). Bell (1908/1909) likewise recommended the colon for "explanatory sub-titles" but the semi-colon for alternative titles and to precede any other information that followed the title.

Because most British public libraries produced printed catalogues it was desirable to keep entries short, and some writers therefore advocated shortening the sub-title as much as possible by eliminating unnecessary words. Jast (1899/1990, p. 119) gives the following examples, where the words in parentheses should be omitted:

: (being the) Report of the Education Section.

for (the Use of) Craftsmen.

: (Account of a) Journey in Tibet.

: (a Series of) Critical Biographies with (Illustrative) Extracts.

Mad Tour: (or, a Journey Undertaken in an Insane Moment) Through Central Europe on Foot.

Such omissions would probably not have been acceptable in other contexts such as academic libraries, and with the exception of Cutter they are not advocated by the other codes.

AACR2 (1.1A2) allows the transposition of elements if they do not appear in the required order on the title-page; a sub-title appearing at the top of the page can therefore be transcribed after the title proper. This question seems not to be mentioned in the earlier rules, and it appears that it was not until LC in 1949 that it became part of the philosophy of description (Hagler, 1977, p. 606). As specific evidence that prior to this such transposition was not practised we may look at Taylor (1948, pp. 24 and 29). She gives the example of a title-page that starts "THIRD EDITION. FIFTEENTH THOUSAND" at the top, rendering this as by "...", with "Third edition" then appearing in square brackets in its normal place. Fortunately such a transcription would be quite wrong nowadays, and one wonders how widespread it ever was, as it seems to be taking the primacy of the title page too far.

Statements of responsibility

AACR2 prescribes that statements of responsibility should be included, with the exception that in Level 1 cataloguing a statement may be omitted if it simply restates the main entry (1.0D1). In older codes, because the description could never be seen without the heading, if a name had already appeared in the heading it was not felt necessary to include it in full, or at all, in the description. Cutter makes no mention of it and would probably have been astonished at any suggestion that such a name could be included. AA (136), unusually, states that the author's name is to be included in the transcription of the title, and it is clear from the examples that it is to be given in full. Despite this, if there was only one author, it became normal to omit this from the description altogether unless its form differed markedly from that used in the heading (Bod 40, Cam 51). This rule continued in AACR67 (134: "Author statement"), though it was implied that any author statement that was transcribed would be transcribed fully; it would be preceded by a comma.

Elsewhere it was acceptable, and recommended, to shorten the form of authors' names, for example, by replacing full forenames on the title page with initials (explicit in Bod 40, Cam 51, and assumed in BM). Similarly Wheatley (1889, pp. 159-60) allows that there are circumstances in which the author's name needs to be represented, if only by initials.

AACR2 allows two kinds of omission from statements of responsibility: the first is when the number of names in a statement of responsibility exceeds three, and this requires the mark of omission ("...") (1.1F5); the other omission is of certain prefixed titles, post-nominal letters and suchlike associated with a person's name, and in these cases the omission is not marked (1.1F7).

The first of these (omitting names when there are more than three) seems to be found first in LC (3:6C), the omitted names being replaced by "and others" if the title-page was in English, or "*et al.*" if in any other language. It seems that neither brackets nor the mark of omission were used, which means that this was a somewhat inaccurate representation of the title page. It is not clear what would have happened prior to this rule, but it would have been very rare until comparatively recent years to find more than two authors anyway. By the time of AACR67 (134D) brackets had been introduced, though there was still no mark of omission. Degrees and suchlike, following authors' names, were omitted without indication (Cam 51, unstated elsewhere), while other omissions were shown by dots (Cam 51). LC (3:6D) provided for the omission of "titles and abbreviations of titles of address, honor and distinction (but

not nobility), initials of societies, etc.", again without indication. A similar rule was included in AACR67 (134E), and this has been expanded (although one case has been dropped) into AACR2 1.1F7, which still retains some of the same examples from LC. The whole question illustrates the tension that has always existed between the primacy of the title page and the desire to eliminate needless verbiage, and the resulting ambiguity is perhaps unsatisfactory.

As for punctuation, if a statement of responsibility was included, it was preceded by a comma if it was the author, or a semi-colon for an editor, translator, etc. (Cam 62). "Full stops are to be avoided in the body of the Description" (Cam 62). Cutter (233-240) gives various examples depending on the level of cataloguing being adopted; here most statements of responsibility are introduced by a semi-colon. Other rules say nothing about it, but the examples (e.g., at AA 1) indicate that a comma was the norm.

Abbreviations in titles were permitted by Bod (45), although these are naturally restricted to those of a "bibliographical" nature, such as "ed." for edition, "illustr." for illustrated, etc., and it does not seem likely that they would have been used in the title proper. A similar list was provided by Cam (57). "Words in the title proper are not to be abbreviated" (AA App. 1; cf. LC App. III).

Finally in this area we should mention the parallel title, now introduced by the equals sign. The term was not used until AACR2, and treatment prior to this varied. There is a great deal of flexibility, for example: give priority to the roman alphabet, or use the title in the original language (AA 145); prefer the one in roman characters, or the one in the language of the book (Cam 52); give the first, and any other that is in English (LC 3:5B, AACR67 133C). Exactly how it was separated from the first title is not stated.

Edition

BM (XVIII) includes edition as part of the title: "The number of the edition to be stated when appearing in the title." The accompanying example shows this in full:

> The fifth edition.

It seems to have continued to be BM practice to transcribe edition statements in the exact form in which they appeared in the book. Cutter (254-6) and AA (148) likewise consider the edition to be part of the title. Both say that it is given in the language of the book "except that customary abbreviations may be used". Cutter clearly sees the utmost abbreviation as being essential, and states that "2d, 3d" should be used, rather than "2nd, 3rd". (It is hardly likely that these would have found favour in Britain.) AA 174 states that "arabic figures may be used", and they seem to be encouraged, for the example gives:

> 8th ed.; *not* Eighth ed.

Bod (42(1)) gives little instruction as to how the edition statement should be transcribed except to say that it is enclosed in brackets when not taken from the title-page. Cam (61) states that the number of the edition should always be given in arabic figures, so that "The fifth edition" becomes "5th ed.". Examples from French, German and Italian are also included, and in the case of the first two, as with English, their practice is exactly as in present-day AACR2 (1.2B1). The edition statement is to be preceded by a full stop. LC (3:7) gives no specific instruction about abbreviating edition statements,

but it becomes apparent from Appendix IV that the number is to be converted to an arabic numeral, as in current AACR2; one assumes that as regards the word "edition" Appendix III was used, making "ed.". AACR67 (135) specifies that the form in the book should be followed, but reference to the appendices on abbreviations and numerals leads to the same shortened forms as just mentioned. AACR2 (1.2B1) now says: "Transcribe the edition statement as found on the item. Use abbreviations as instructed in Appendix B and numerals as instructed in Appendix C." These two sentences are on the face of it contradictory, and it would be helpful if they were connected with "but".

Publication, distribution, etc.

The three elements in ISBD (with required punctuation) are place: publisher, year. We can look at these in turn.

Place

"Then the place where the book was printed; and in particular cases, as in the instance of early, or very eminent typographers, the Printer's name to be specified" (BM XXVII). Notice here that the reference is to printing, not publishing, but this must have long ceased to be useful in dealing with modern books. It seems, therefore, to have come to be interpreted as meaning publisher.

For Cutter (257-275) the Imprint "consists of place of publication, publisher's name, date, number of volumes, number of pages, number of maps, engravings, and the like, and typographic form", and yet he also has a separate section entitled Collation (276-280) which includes the elements listed after the date, i.e. what we would now call Physical description. He recognises the differing needs of different kinds of library, saying that in the majority of municipal libraries "neither the character of the readers nor of the books" justifies the insertion of imprints. This is certainly borne out by an examination of British public library catalogues of the late nineteenth and early twentieth centuries, where it is rare to find more than the date of publication. Quinn (1899, p. 19) says that normal practice is to omit the place if it is London, the implication being that otherwise it should be included. In line with his recommendations elsewhere Cutter says that abbreviations, sometimes even down to one letter, should be used for the place, e.g. Balt., Berl., Bost., Camb., L. (London), P. (Paris), Phila. and several others. In the case of more than one place, "take that which proves on examination to be the real place of publication", unless they are joined with "and", in which case take the first. Exactly how examination should prove which place is real is not stated.

AA (150) gives very little detail: "After the title give the place or places of publication in the language of the title". The only additional rule refers to anglicised versions of Latin names; there is no instruction on how to deal with multiple places of publication, and this and other problems were dealt with only by some "Library of Congress supplementary rules" following AA 157.

Cambridge allows abbreviations for places too: "exactly as it appears in the book, except that abbreviations may be used. If more than one place of publication is mentioned, choose that which appears on examination to be the real one; where there is any doubt, choose the first" (Cam 55(2)). The wording here is so close to Cutter's rule that it must have been based on it, which again shows that, as one might expect, Cambridge was thoroughly familiar with other cataloguing rules and was happy to adopt what it took to be best practice. At Oxford the form used in the book was to be

followed "but customary abbreviations may in all cases be used, for example, Oxf., Lond., Edinb., Leipz., Lugd. Bat., Phila." Bod 42(4)). When more than one place of publication, "&c." to be added (Bod 42(4)).

The problem of multiple places was addressed further by LC (3:10), with priority being given to the first, but with exceptions. We also here begin to see the practice of including a second place if it was in the country of the cataloguer, in this case the United States. This was continued in AACR67 (British ed. 139B, American ed. 138B). Multiple places are separated by a comma. The problem of multiple places is, if anything, of even more importance now because of the international interchange of records and the increasing globalisation of the publishing industry. If priority is given to the home country of the cataloguing agency (AACR2 1.4C5) then records obtained from a different country are likely to be "wrong". This is a problem that needs to be tackled, but what the solution might be is difficult to say.

Names of counties, states, etc., could be added in parentheses (AACR67 140A); there is no differentiation according to whether or not they are being copied from the book, and abbreviations are to be used for "most such additions". It is interesting that although one of the examples shows "Boston (Lincs.)" there is no abbreviation for any British county in the appendix, just as there is still none in AACR2. AACR2 (1.4C3) introduced the distinction of using square brackets for additions that were not present in the source, and simply a comma if they were, and this has been maintained.

Absence of place of publication is marked by leaving a space (Cam 55(2)) but otherwise not mentioned until LC (3:10C), which prescribed the use of "n. p.". AACR67 (139C) prescribed "[n.p.]", and this seems to have been the first use of square brackets in this position. Presumably the assumption prior to this had been that from the meaning of the letters it could be inferred that the information was not being copied from any source and that therefore square brackets were superfluous.

Publisher

The name of the publisher seems always to have assumed much less importance than the place, and is omitted in several codes, including Bod and Cam. It was not until 1880 that the British Museum began to include the publisher systematically; initially they did this for all English books and for foreign books printed before 1700 (Chaplin, 1987, p. 74). The placing of this element was unusual and unique, in that it preceded the place from which it was separated by a colon. Perhaps it was felt more convenient to place it here because the place and date were usually separated from the rest of the description by a sizeable gap. Cutter (262) includes the publisher, though with the implication that it is not always necessary, and in LA (7) it is likewise optional. AA (152) also includes it, with no guidance as to how fully to copy the name, though there is a note that "In a written card catalogue the place and publisher's name may generally be abbreviated". AA includes additional rules from the Library of Congress giving more guidance on the amount of detail for both place of publication and name of publisher. These rules are divided according to whether or not the book is published in the USA. Priority is given to the "relative importance" of the places within the USA, and it is not surprising that this rule is not found again later as it must have been difficult to interpret. In cases where two places were connected by "and" or its equivalent, this word was to be retained in the description, with both places. LC (3:12A) gives useful guidance on the

shortening of publishers' names. AACR67 (141) separates the publisher from the place by a colon, and repeats most of this guidance.

Absence of publisher was not noted (LC 3:10C); later, indicated by "[n.pub.]" (AACR67 139C). Elsewhere the question is not mentioned.

Date

Of the three elements, this is the only one that Jast (1899/1990, p. 285) includes, and, judging from their surviving printed catalogues, his recommendation seems to have been followed by the vast majority of public libraries. The intention was to save space by providing only the minimum of information. For novels not even this detail should be included, although Jast (1899/1990, p. 287) does suggest that for reprints it will be helpful to add a note of the original edition.

All rules that mention the matter say that the date is to be expressed in arabic figures even if it appears in a different form in the source (Cutter 272, LA 7(g), AA 155, Cam 55(3)). If no date is given, it has always been essential to make an attempt, and to put estimates like "[17–?]", "[189-?]" (Cutter 271, Cam 55(3)), though Cutter and AA (156) did allow "n. d." (again without brackets) if it was unavoidable. The system of estimating dates has been carried right through into AACR2 (1.4F7). Bod (42(5)) allowed dates from prefaces, etc., to be included, in round brackets, and dates from elsewhere in the text in square brackets. The abbreviation "n.d." was permitted, but an approximate date was preferable. In AACR67 (142K) "n.d." was still allowed, though it is implied that an approximate date should be given if possible.

A copyright date was allowed by Cutter (265) if the date was not on the title-page: for example, "[c. 1892]". AA (157) also permitted it if there was no date of publication; in this case the abbreviation was a superscript "c" attached to the year, the whole being in square brackets, for example, [^{c}1894]. One wonders how many library users would have understood the significance of this, which must be the origin of the present-day AACR2 usage of "c" to mean "copyright". A similar instruction is given by LC (3:13F), though without necessarily using the brackets. Instructions about copyright dates are rather more informative in AACR67 (142G) than in the current edition. We are told that if the copyright symbol © appears, the date accompanying it can be used without brackets or any other note; this seems eminently sensible, and it is a pity that the current rule (1.4F6) does not explicitly say this. If, however, the date of the book is earlier than the Universal Copyright Convention of 1952 then any copyright date should be preceded by the word "[copyright]". This is very much clearer than the present use of the single letter "c" (AACR2 1.4F6), and it is puzzling to know why the change was made, especially as ISBD gives "cop.".

Physical description

The elements which now make up this ISBD area tend to be scattered in different places in the various codes under consideration. BM (XXVII) says no more than "The form to follow, whether fol., 4to, 8vo, &c.", meaning following the date of publication. This was as close as one could get to the size of the book. There was no indication of the presence of illustrations, or of pagination, though it is clear from looking at the catalogue that often a note such as "[With a portrait]" would be added directly to the title. It was not until 1880 that the number of pages was added, and this was inserted before the name of the publisher and the place of publication (Chaplin, 1987, p. 74).

Useful examples of the variety in content and positioning of the different elements are shown by Hagler (1977, p. 610).

For the LA (7) the elements to be listed were: number of volumes if more than one; number of pages; number of separate illustrations, maps or portraits; size; but only the first and last of these were compulsory. All preceded the place, publisher and date, and in this respect the rule was similar to BM. Jast (1899/1990, p. 285) follows a similar order in stating that the "imprint" consists of the following:

(1) Volumes.
(2) Illustrations:
    - portraits;
    - diagrams; and
    - facsimiles.
(3) Tables.
(4) Maps.
(5) Plans.
(6) Date.
(7) Series.

Clearly most of these constitute what we should now call the physical description, only Date and Series belonging elsewhere. Jast's abbreviation for illustrations was "il."; their number should be given if not exceeding six, or if the number was stated on the title page. He also recommends using numbered footnotes in the catalogue to amplify information given about certain illustrations, but obviously this was possible only in a printed catalogue. Writing a little later Jast (1902) criticises the LA rules for not making the mention of illustrations, maps and portraits compulsory, saying that "the best modern practice" always gives these. Conversely hardly any public library catalogues provided place of publication.

Cutter (257) is the first code to adopt the same order as that of the present AACR2, i.e. number of pages (or volumes), illustrations, size, and this was followed by AA (158-164). Cutter tells us to give the last number of each paging, which is more or less what AACR2 says today. As with the previous area, little instruction is given as to how much detail of pagination should be included, the only unusual thing being the possible addition of a plus sign (+) to indicate any final unnumbered pages. Different types of illustration should be included (Cutter 278), and later there developed a set order for these (AA 161), unlike the AACR2 order, which is alphabetical. Plates here count as a kind of illustration, rather than, as now, being part of the pagination element. Finally (Cutter 279) comes the size, which is to be given either in centimetres or using a special system of notation drawn up by the American Library Association using capital letters. This scheme had presumably been abandoned by 1908, for it does not appear in AA 164, which simply states that the height of the book is given in centimetres, this time rounded up or down to the nearest half-centimetre, so that 172 mm would be shown as 17^{cm}, while 173 mm would be $17\frac{1}{2}^{\text{cm}}$. Note that the abbreviation is shown superscript. This must have been very time-consuming compared to the present-day AACR2. Just as now, books of unusual format were to have both dimensions given. As with the publication area, AA includes some Library of Congress supplementary rules,

giving further detail about the whole of this area. This detail goes far beyond what is provided by other codes until LC, and there is no need to discuss it here as it is not comparable with anything else.

AACR67 (143B) begins to give more detail, of the kind with which we are familiar from AACR2. Plates are mentioned as part of the number of pages element rather than as illustrations and most of the rules that are in force now are to be found here, with the exception of the treatment of unnumbered sequences. The different types of illustration (in addition to "illus." –"ill." from 1975) are listed in alphabetical order (143D1b), as they are now.

Bod (42), like BM, places the number of volumes and illustrations *before* the place and date of publication, followed then by the size. Illustrations are to be mentioned only "when they are of more than ordinary importance" (Bod 42(3)). Size consists of height rounded *down* to the last complete centimetre, in the form "cm. 15" and including breadth when equal to or greater than the height (Bod 42(6)). This in some respects anticipates AACR2 (2.5D4). Cam (55) places the height of book (rounded *up* to the next whole centimetre, e.g. "18 cm.") before the place of publication, and then the date. Again both dimensions are required for unusual shapes.

LC appears to be the first code to give a reason for including the size of a book (3:14D): "The size of the work is included in the catalog entry as an aid in finding the work on the shelves and as an aid to the user of the catalog in selecting a desirable edition. It also serves the reader who wishes to borrow the work through interlibrary loan or who wishes to order a photocopy of the work or a part of it." By this time too we find that the height is to be rounded up to the next whole centimetre, as it is in AACR67 (143E) and in AACR2 (2.5D1) today.

Series

The use of parentheses to enclose series statements has a long pedigree. It is found in Cutter (237, "after the imprint"), AA (166 "after the collation") and Cam (40, where it appears after the main title and before and size and imprint), and has continued in AACR67 and AACR2. BM dealt with series in two different ways, according to whether an independent entry was made for the series title. No detail is given in 1841, and presumably the system had not then been worked out, but it appears in the 1936 edition (rule 30). If the whole series was shelved together there would be a series entry in addition to the normal entries for individual volumes. The latter included the series statement in italics and square brackets at the end. In other cases, where no series entry was made, the series statement appeared as a note in italics at the foot of the entry, the actual series name being in double quotation marks.

Notes

Because the Note area in AACR2 covers such a wide range of kinds of information, it is difficult to generalise about it, and it is perhaps best to restrict discussion to its positioning. Cutter is not explicit about this, but LA says that both Contents and Notes should be "subjoined" and "in smaller type". Likewise all the examples in AA (167-168) are in smaller type and in separate paragraphs. Bod (36, 47), however, seems to assume that most notes would be added immediately after the title itself. Round brackets were to be used for notes taken from elsewhere in the work itself, and square for notes from elsewhere. Similarly Cam 54(5) states that notes about limited editions, privately

printed works, etc., were put in italics and round brackets following the series (which followed the edition (see above)).

Conclusion

We can see that there has been great variation in both the position and the content of what we now know as AACR2's "areas". The general order of title – edition – publication details has been fairly constant, but the physical description has been broken up in various ways and put in different places. Certain rules in AACR2 have surprisingly long histories and have been anticipated in different libraries' codes. The ISBD arrangement was therefore clearly born of a consensus developed gradually over a long period. The primacy of the title page has always existed in tension with the desire to avoid unnecessary detail. Now that retrieval can in theory be based on any part of the catalogue record, not simply on "access points", it is perhaps time for this primacy to be questioned again. Some other problems too, such as treatment of copyright dates, and multiple places of publication, are still with us and provide fresh challenges for the global bibliographic environment of the twenty-first century.

References

American Library Association, the British Library, the Canadian Committee on Cataloguing, the Library Association and the Library of Congress (1978), in Gorman, M. and Winkler, P.W. (Eds), *Anglo-American Cataloguing Rules*, 2nd ed., Library Association, London.

American Library Association, the Library of Congress, the Library Association and the Canadian Library Association (1967), *Anglo-American Cataloguing Rules*, Library Association, London (British text).

Bell, E.J. (1908/1909), "Punctuation of catalogues", *Library World*, Vol. 11, pp. 38-9.

Blake, V.L.P. (2002), "Forging the Anglo-American cataloging alliance: descriptive cataloging, 1830-1908", *Cataloging & Classification Quarterly*, Vol. 35 Nos 1/2, pp. 3-22.

Bowman, J.H. (2001), "Sic catalog sydrome: title page transcription as a barrier to retrieval", *Cataloging & Classification Quarterly*, Vol. 32 No. 1, pp. 39-54.

British Museum (1841), *Alphabetical Catalogue of Printed Books: Rules to Be Observed in Preparing and Entering Titles*, British Museum, London.

British Museum (1936), *Rules for Compiling the Catalogues of Printed Books, Maps and Music in the British Museum*, rev. ed., British Museum, London.

Chaplin, A.H. (1987), *GK: 150 years of the General Catalogue of Printed Books in the British Museum*, Scolar, Aldershot.

Committees of the Library Association and of the American Library Association (1908), *Cataloguing Rules: Author and Title Entries*, Library Association, London.

Cutter, C.A. (1904), *Rules for a Dictionary Catalog*, 4th ed., GPO, Washington, DC.

Edmond, J.P. (1903), "On the use of capital letters in cataloguing", *The Library*, Vol. 5, pp. 77-80.

Gorman, M. (1971), "The standard bibliographic description", *Catalogue & Index*, No. 22, pp. 3-5.

Gorman, M. (1989), "Yesterday's heresy – today's orthodoxy: an essay on the changing face of descriptive cataloguing", *College and Research Libraries*, Vol. 50 No. 4, pp. 626-34.

Gorman, M. and Winkler, P.W. (Eds) (1988), *Anglo-American Cataloguing Rules*, 2nd ed., 1988 revision, Library Association, London (prepared under the direction of the Joint Steering Committee on Cataloguing, the Library Association and the Library of Congress).

Hagler, R. (1977), "Changes in cataloging rules: rules for description", *Library Trends*, Vol. 25 No. 3, pp. 603-23.

Hanson, J.C.M. (1939), *A Comparative Study of Cataloging Rules Based on the Anglo-American Code of 1908*, University of Chicago Press, Chicago, IL.

Henderson, K.L. (1976), "'Treated with a degree of uniformity and common sense': descriptive cataloging in the United States 1876-1975", *Library Trends*, Vol. 25 No. 1, pp. 227-71.

Hitchler, T. (1903), *Comparative Cataloguing Rules: 20 Points in 10 Codes Briefly Compared*, George G. Peck, New York, NY.

International Federation of Library Associations and Institutions (IFLA) (2002), "ISBD(M): international standard bibliographic description for monographic publications", available at: www.ifla.org/VII/s13/pubs/isbd_m0602.pdf (accessed 21 January 2005).

International Federation of Library Associations and Institutions (IFLA) (2004), "ISBD(G): general international standard bibliographic description", available at: www.ifla.org/VII/s13/pubs/isbdg2004.pdf (accessed 21 January 2005).

Jast, L.S. (1899/1900), "Classified and annotated cataloguing: suggestions and rules", *Library World*, Vol. 2, pp. 118-21, 229-31, 285-8.

Jast, L.S. (1900/1901), "Classified and annotated cataloguing: the individual biographical entry", *Library World*, Vol. 3, pp. 29-31.

Jast, L.S. (1902), "The Library Association cataloguing rules", *Library Association Record*, Vol. 4, pp. 579-82.

Joint Steering Committee for the Revision of AACR (1998), *Anglo-American Cataloguing Rules*, 2nd ed., 1998 revision, Library Association, London.

Joint Steering Committee for the Revision of AACR (2002), *Anglo-American Cataloguing Rules*, 2nd ed., 2002 revision, CILIP, London.

Library Association (1884), *Library Association Year Book*, Library Association, London, pp. 65-8.

London, G. (1980), "The place and role of bibliographic description in general and individual catalogues: a historical analysis", *Libri*, Vol. 30 No. 4, pp. 253-83.

Norris, D.M. (1939), *A History of Cataloguing and Cataloguing Methods 1100-1850: With an Introductory Survey of Ancient Times*, Grafton, London.

Quinn, J.H. (1899), *Manual of Library Cataloging*, Library Supply Co., London.

Strout, R.F. (1956), "The development of the catalog and cataloging rules", *Library Quarterly*, Vol. 26 No. 4, pp. 254-75.

Taylor, M.S. (1948), *Fundamentals of Practical Cataloguing*, Allen & Unwin, London.

Wheatley, H.B. (1889), *How to Catalogue a Library*, Elliot Stock, London.

Corresponding author
J.H. Bowman can be contacted at: j.bowman@ucl.ac.uk

The need for a faceted classification as the basis of all methods of information retrieval

Vanda Broughton
School of Library, Archive and Information Studies, University College London, London, UK

Received 21 July 2005
Revised 14 September 2005
Accepted 26 September 2005

Abstract

Purpose – The aim of this article is to estimate the impact of faceted classification and the faceted analytical method on the development of various information retrieval tools over the latter part of the twentieth and early twenty-first centuries.

Design/methodology/approach – The article presents an examination of various subject access tools intended for retrieval of both print and digital materials to determine whether they exhibit features of faceted systems. Some attention is paid to use of the faceted approach as a means of structuring information on commercial web sites. The secondary and research literature is also surveyed for commentary on and evaluation of facet analysis as a basis for the building of vocabulary and conceptual tools.

Findings – The study finds that faceted systems are now very common, with a major increase in their use over the last 15 years. Most LIS subject indexing tools (classifications, subject heading lists and thesauri) now demonstrate features of facet analysis to a greater or lesser degree. A faceted approach is frequently taken to the presentation of product information on commercial web sites, and there is an independent strand of theory and documentation related to this application. There is some significant research on semi-automatic indexing and retrieval (query expansion and query formulation) using facet analytical techniques.

Originality/value – This article provides an overview of an important conceptual approach to information retrieval, and compares different understandings and applications of this methodology.

Keywords Classification, Information retrieval

Paper type Conceptual paper

Introduction

In 1955, the recently formed Classification Research Group (CRG) of the UK issued a statement which was published in the *Library Association Record* (Classification Research Group, 1955) and which proclaimed their desire to see faceted classification as the basis of all information retrieval. The group that signed this paper consisted of the leading exponents of classification theory of the period: academics, teachers and researchers, and practising librarians. The purpose of the current paper is to see, 50 years on, how far has that objective been achieved.

At the time, faceted classification was a relatively new phenomenon. Ranganathan (1960) is generally credited with introducing the concept of facet analysis in his Colon Classification and in his theoretical writings (Ranganathan, 1967); although it is only fair to say that a number of earlier writers had advanced similar notions, albeit in a more limited manner[1]. The majority of Ranganathan's writings had been published in the 1930s, within 20 years of the formation of the CRG in 1952. Members of the CRG at that time began to use facet analysis in a somewhat experimental way as the basis of

Aslib Proceedings: New Information Perspectives
Vol. 58 No. 1/2, 2006
pp. 49-72

0001-253X
DOI 10.1108/00012530610648671

a number of new classifications, primarily for special libraries, and confined to very specific subject areas. Soon they would embark on a project to construct a new general scheme of classification based on faceted principles (Classification Research Group, 1964), and some major part of their activities was devoted to this; ultimately the work never resulted in a classification scheme, but much of it contributed to the PRECIS indexing system developed by Austin (1984) for the British National Bibliography.

Thus, even from the beginning, the methods of facet analysis were not confined to the creation of classification schemes per se; they were seen to be relevant to alphabetical subject indexing, and later to the development of thesauri (Aitchison *et al.*, 1969), in the 1950s still an emerging tool, but today very much the preferred indexing language for many environments.

Why did the CRG regard the faceted classification as so superior to its predecessors? What were the features of faceted classification that ensured effective information retrieval? Although not explicitly stated, various comments in the paper suggest the following:

- the display of useful generic relationships;
- full and accurate cross-referencing;
- accurate application of principles of division;
- a clear citation order;
- established rules for compounding; and
- an appropriate notation.

The CRG at that time were concerned with the application of classification only to the organisation of print media, whether this was the physical organisation of a collection of books or documents, or the arrangement of document surrogates such as card catalogues or printed indexes and bibliographies. As a consequence some of the perceived advantages of the faceted classification are related to the difficulty of producing a linear order for non-linear (i.e. compound) subjects. The faceted classification, with its deconstructed vocabulary and clear rules for combination through the medium of the citation order, has a wonderfully unambiguous syntax leaving little room for doubt about the placing of compounds. When retrieving from a linear arrangement, whether of items or their surrogates, this property of predictability is a matter of central importance.

We know that the situation in a digital environment is really rather different. Linearity does not concern us overmuch (although it may still be relevant to the way in which information is displayed on the screen if not its physical order in the information store or repository). The concerns in managing the digital information store are not those of arranging the material, but rather of adequate object description (labelling the items to support subject retrieval), providing search tools that support browsing, navigation and retrieval, and, to a more limited extent, the presentation of results. Viewed in this context, faceted classification offers the following benefits:

- the capacity to express through synthesis the complexity of subject content that is typical of digital documents;
- a system syntax that ensures this is managed in a regular and consistent manner;

- a rigorously logical structure that is compatible with machine manipulation at whatever level;
- a structure that is compatible with a graphical interface for end-user navigation and query formulation;
- the facility through variation or rotation of the citation order to allow approaches from a number of angles (i.e. cross domain searching);
- a structure and methodology that permits conversion to other index language formats (i.e. subject heading lists and thesauri); and
- features of these integrated tools that allow modifiable keyword searching through mapping vocabularies and vocabulary control via the thesaurus, and provide tools for browsing and display via the subject heading list.

We shall investigate the way in which these characteristics of faceted classification are increasingly seen, in whole or in part, in a range of systems and tools for information retrieval and subject access. The impact of facet analysis on the "conventional" schemes of bibliographic classification, and their terminological counterparts, the thesaurus, and the subject heading list, will be reviewed. The role of facets in the development of new subject access tools, the concept map and the ontology, and the part played by faceted structures in the application and implementation of retrieval systems on the web will also be considered.

Basic principles of facet analysis and faceted classification structures

Before looking at the impact of facet analysis on the general schemes of classification, it is important to have a clear understanding of what a faceted classification is. This is necessary because the nature of faceted classification is often misunderstood.

A common misperception of a faceted scheme is one in which there is analytico-synthesis; that is to say, there is some element of "deconstruction" of a compound subject (analysis) and reconstruction of it (synthesis) using the terminology and combination rules of the controlled language to create a classmark or a subject heading. Most modern classification schemes are to some degree analytico-synthetic: both DDC and UDC use auxiliary schedules to build classmarks containing commonly occurring concepts such as form and place. The same phenomenon can be observed in LCSH where compound headings can be constructed by adding topical, geographical and free floating subdivisions to main headings. Even in the Library of Congress classification (the least analytico-synthetic of all the major systems) subject extension is often possible by the use of tables. The UDC takes this process further by allowing the combination of any selection of concepts in the classification, whether these are commonly occurring or not.

However, this analytico-synthetic function does not make these systems faceted. Facet analysis implies a fundamental structural organisation of the vocabulary of an indexing tool starting from a "bottom-up" position that cannot be discerned in these existing tools.

Fundamental features of the faceted indexing language

Most of the examples of structure are taken from the schedules of the second edition of the *Bliss Bibliographic Classification* (BC2) (Mills and Broughton, 1977-) which in many

ways represents the full flowering of CRG theory. Since the late 1960s work on the revision of BC2 had been brought to the CRG table for comment and discussion, and many members of the CRG were active in the creation of schedules. Consequently BC2 has been the testing ground, and is the major vehicle for the expression of CRG work in a classification. BC2 has become the CRG's general classification scheme, which was not realised in the 1950s project, and it will be used here to demonstrate that theory.

Faceted classification simplified

Many introductory books on classification present faceted classification in a simplified but rather limited manner by using an example based on the attributes of entities. This simplification is important because it is frequently the view of faceted structures that is adopted by many newcomers to the concept (particularly those using it in web applications).

The example given in Table I is trivial but typical, and involves the classification of socks.

We can select terms from Table I to define a collection of socks, combining them to represent all the characteristics present in a particular sock. So a given sock could be a black wool work sock, or a blue striped polyester football sock. If the structure has a specified order of combination, or citation order, it can be populated with combinations of attributes to generate a more complex structure very similar to an enumerative classification, but with a more rigorous and logical pattern to it:

Grey socks
Grey wool socks
  Grey wool work socks
  Grey wool hiking socks
    Grey wool ankle socks for hiking
    Grey wool knee socks for hiking

However, the sort of selection of terms we see above is a faceted classification with only one facet: that of entity, or personality in Ranganathan's understanding. The structure is a taxonomy rather than a classification, since there are no concepts outside that of the primary facet, and the organisation is into arrays within a facet rather than into facets.

Such an arrangement is often presented as an example of a faceted classification, and it does give quite a good sense of how a faceted classification is structured. A faceted bibliographic classification has to do a great deal more than this, and a proper faceted classification will have many more facets, covering a much wider range of terminology.

Colour	Pattern	Material	Function	Length
Black	Plain	Wool	Work	Ankle
Grey	Striped	Polyester	Evening	Calf
Brown	Spotted	Cotton	Football	Knee
Green	Hooped	Silk	Hiking	
Blue	Checkered	Nylon	Protective	
Red	Novelty	Latex		

Table I.
Classification of object attributes

Source: Adapted from Broughton (2004, p. 262)

Fundamental categories

From the early days of classification and indexing, a number of writers and compilers of systems had noticed the regular occurrence of common attributes such as place, time and form. The first published schedules of UDC (1905-1907) made provision for these to be achieved by synthesis from generally applicable tables, as did Bliss (1910) in the first drafts of the Bibliographic Classification.

Ranganathan's contribution, and the major breakthrough in facet analysis, was to see that not only were there recurrent concepts common to virtually all subjects, but also that there were common types of concepts within the subjects themselves. Some were activities or actions, which he labelled the Energy facet; others related to substances constituted the Matter facet; the core concepts representing the primary object of study within a discipline Ranganathan called the Personality facet since it represented the essence of the discipline. The members of the personality facet are very often (although not exclusively) entities of some sort or another: plants, animals, chemical compounds, astronomical bodies, geographical formations, religions, manufactured objects, and so on. To these were added the commonly occurring Space and Time to give the famous Ranganathan facet formula, PMEST.

Within a discipline or subject domain, all the concepts or terms could be organised into these five categories: personality, matter, energy, space and time. These are rather too few categories for some disciplines, and the Colon Classification which Ranganathan created using them, often has to employ more than one P, or E category, which are then labelled as different rounds and levels, P1, P2, E1, E2 and so on. The CRG expanded these fundamental categories to 13: thing, kind, part, property, material, process, operation, agent, patient, product, by-product, space and time. Such categories can accommodate the vocabulary of most existing disciplines, although arts and humanities often require some additional ones (form, style, genre) and there is nothing to say that new fundamental categories cannot be discovered, or perhaps invented, for the essence of facet analysis is that it is a practical art.

These fundamental categories form the basis of facet analysis, as defined by the CRG statement. The application of each *category* to the containing discipline as a broad principle of division generates a specific and discrete set of concepts, or *facet*. In the example below we can see the allocation of terms to categories in the discipline of medicine to create the facets of the subject:

(Thing) Human beings

(Kind) Women, children, old people, etc.

(Part) Head, legs, muscles, bones, heart, brain, lungs, etc.

(Process) Respiration, digestion, reproduction, disease, etc.

(Operation) Surgery, drug therapy, physiotherapy, etc.

(Agent) Doctors, nurses, equipment, buildings, etc.

Within each facet it will be necessary to further organise terms into arrays (sometimes called sub-facets) on the basis of their attributes as discussed in the sock example. Order within an array is usually decided on a pragmatic basis; chronological and developmental orders are common, as are orders based on geographical proximity, size

and other physical attributes. This list of prisons from BC2 Class Q uses the degree of security as the basis of arrangement:

QQS K (Types of prisons)
QQS M Maximum security
N Medium security
P Minimum security, open prisons
R With semi-liberty (living in institution, working outside)
S With restricted liberty (living at home, working at institution)

An important thing to notice about the members of an array is that they are all mutually exclusive classes. If the analysis is accurate there should be no difficulty about this. Enumerative systems on the other hand often produce groupings of classes that are not mutually exclusive, and that is a sure sign of a "non-faceted" structure.

It is perhaps worth stating here that in both the Ranganathan and CRG models, the universe of discourse is the discipline. Neither method attempts a unitary application of the fundamental categories to the whole of knowledge; although there has been some considerable discussion within the group regarding the feasibility of this, in practice it is seen to be unworkable (or at least very difficult). The faceted general classification is a series of subject classifications, each with its own facet structure and facet formula or citation order. The primary facet is not therefore "discipline" as stated in some sources (Priss, 2000); although there is an initial division into disciplines, this is external to the application of the facet analysis proper.

Relationships between concepts

A second major feature of the faceted scheme is clarity in the expression of the relationships between concepts, both the intra-facet relationships (semantic relationships) and the inter-facet relationships (syntactic relationships).

As far as intra-facet relationships are concerned, because all the terms within a facet come into the same category (they are all things or parts or processes) the relationships between them will be those of hierarchy, or broader, narrower, and co-ordinate terms. In a faceted scheme these are equally likely to be found in the operations and processes facets as they are in those facets dealing with entities or objects, as this example from BC2 Class Q shows:

QP Police work, law enforcement
QPD Police work narrowly
QPD O Communications
QPD P Patrol and surveillance
Q Patrolling, beat
R Stopping and questioning
S Search and seizure
T Surveillance
U Pursuit and apprehension
V Pursuit
W Apprehension, arrest and charge
QPE Criminal investigation, detection

Where a faceted classification differs most significantly from an enumerative classification is in its potential to combine terms from different facets, and it is here that another major feature of the faceted scheme comes into play: the relationships

between facets, and between terms from different facets – the inter-facet relationships. The number and variety of these relationships seem unique to the faceted classification, and although they are seldom the object of discussion they represent a degree of sophistication that does not seem to exist elsewhere. In a structure using, for example, eight of the categories of standard citation order to create facets, the number of potential inter-facet relationships of the "agent-operation" or "process-product" type, will be 56, a significant difference from the three standard thesaural relations, or the common "is-a" or "is-a-part-of" relationships of ontologies.

Citation order

Citation order controls several aspects of the classification and the classified order:

- it gives rules for the order of combination of terms when classifying;
- it determines which aspects of a subject are brought together and which are distributed; and
- it affects the logical structure of the system, the predictability of locating compound subjects, and hence the effectiveness of retrieval.

Any system that allows combination will have a citation order. Citation order need mean no more than the order of combination of terms and in this sense even pre-coordinated systems like LCC and LCSH have citation order, although this is more usually implicit in an enumerated list of compounds, than explicit as a means of synthesising them.

Citation order in the faceted scheme is more complicated because there are a great many more facets to enter into the equation. Rather than make a decision about citation order at every potential place of compounding, the faceted scheme has a general rule for a default citation order, known as standard citation order. This states that in the majority of cases the order of combination will be that of the categories as they are normally listed, i.e.:

> Thing – kind – part – property – material – process – operation – patient – product – by-product – agent – space – time

There are three major theoretical arguments for this order:

(1) the order progresses from concrete to abstract;

(2) each facet is dependent on preceding facets; and

(3) it gives the most useful grouping of compounds.

Schedule order

Although it is not a feature of the Colon Classification, most faceted schemes following the British tradition employ what is known as schedule inversion. This means that the order of facets in the schedule reverses that of the citation order, beginning with the more "abstract" facets, such as time and space. Classmarks for compound subjects are built retroactively, working backwards through the schedule to add notation from earlier filing (but later citing) facets. The notion of schedule inversion, and how and why it functions, is a complex one and because it is mainly relevant to linear ordering will not be pursued here; other sources provide a full explanation of the principle (Broughton, 2004).

The forgoing discussion of faceted classification follows the line of Ranganathan as it was subsequently developed by the CRG. Although this is widely comprehended as a "standard" of faceted classification theory, it is not the only model operating in different information sectors. As we shall see in the following sections, what constitutes a "facet" and the precise nature of "facet analysis" is subject to different interpretations; some of these use facet analysis in a much more limited way, others understand facets to embrace rather more than the semantic aspects of documents. Each section will examine how facet analysis is defined, and how this affects the way in which the method is applied.

The influence of facet analysis on conventional classifications

Facet analysis is important because it provides a rational methodology for building a classification (Hjorland, 2002) in contrast to the entirely pragmatic groupings of classes, which characterise classifications built prior to its invention. It also provides a coherent body of theory and formalises a great deal of the good practice of earlier classificationists. Features that were previously regarded as of practical value, such as the rule of general-before-special, now have a proper philosophical basis.

It is possible to see in all the general schemes of classification a tighter and more rigorous approach to the structure and organisation of classes than previously. Elements of the faceted scheme, such as schedule inversion and organisation of concepts into facets and arrays are now much more common, as is the consistent application of citation order, and an increasing level of analytico-synthesis.

Initially the methods of faceted classification did not impact greatly on the wider world; Ranganathan's Colon Classification had hardly been used outside the Indian subcontinent[2], and the efforts of the CRG were largely confined to special schemes. By the early 1990s this had begun to change, and the influence of facet analysis could be seen in the general classification schemes. The Dewey decimal classification which from the 1950s onward had introduced an increasing number of analytico-synthetic features, now began to speak of "facets" and "citation order":

> Since the 1950s the impact of library classification theory and technique development on the DDC has been very direct. The most obvious results ... have included ... the increasing use of subject faceting and notational synthesis in the system (Miksa, 1998, p.80).

In recent editions of the scheme this has been made absolutely explicit:

> The Decimal Classification Editorial Policy Committee ... has endorsed the general trend towards more faceting in the Classification. Why are facet indicators and notional synthesis important? The use of facet indicators to identify meaningful components in a number and the use of uniform notation to express recurring aspects of topics within a schedule expand retrieval possibilities by providing access to information represented by parts of a number (Mitchell, 1996, p. xx).

While this suggests that the faceted element of DDC is restricted to extended use of synthesis and facet indicators, in reality the effect of facet analysis can be seen in better structured, more consistent and logical schedules. In 1982 the revision of Class 78 for music displayed a clearly faceted structure, and incorporated many features of faceted music schemes such as that of the *British Catalogue of Music*.

The Universal Decimal Classification (UDC) is also committed to a programme of radical revision designed to base the classification on facet analytical principles. In a

key paper McIlwaine and Williamson (1994) proposed the conversion of UDC from an analytico-synthetic scheme to a fully faceted one. Some efforts had already been made by McIlwaine as Editor of the UDC to begin this process. A rolling programme of revision to remove pre-coordinated classes from the main tables and to extend and improve the systematic auxiliary tables had been instituted, and there was a much more faceted feel to the classification overall. The core of the new project was to be the incorporation of the BC2 structure and vocabularies into the existing UDC tables as part of the programme of ongoing revision of the scheme (McIlwaine and Williamson, 1993). There were difficulties associated with this as a methodology; while it was highly flexible and in theory hospitable to conversion to a completely faceted structure, some aspects of the UDC were at odds with the "culture" of BC2. These can be listed as follows:

- Although it is highly analytico-synthetic the UDC still has some considerable element of pre-coordination in the schedules; over the years combinations had not always been placed consistently, or following a standard citation order.
- The correspondence between notation and schedule structure in UDC is extremely close, and difficulties were encountered in absorbing a structure which had been developed quite deliberately independently of the notation.
- A by-product of the notational problem was the number of un-notated "classes" in BC2 in the form of "principles of division" and other conceptual labels in the schedules. The relatively flat structure of the faceted schedule contrasted with the highly detailed UDC with its many levels of hierarchy, all closely linked to the expressive notation. The conversion of these structural markers to containing classes has occupied much time.
- Although a default order of combination is implicit in UDC, the citation order for compounds is very much at the discretion of the classifier, and the imposition of schedule inversion with retroactive building posed some problems for schedule construction that had not needed to be faced before.
- Similarly, the method of number-building in UDC does not impose a specific citation order on the classifier in the way that most faceted schemes do.
- The general method of linking classes, the colon, resulted in very long class marks for most compounds, and a more elegant alternative was sought.

Most of the exploratory work in this area has been done by Williamson in connection with the medicine tables, and to date a number of proposed new tables have been published (McIlwaine and Williamson, 1995). A new and fully-faceted revision of Class 2, Religion, was published in 2000 (Broughton, 2000), which tackled some of the problems of compound number building described above.

The work on UDC has been particularly interesting in the problems that have arisen in connection with the holding of the classification in its CDI-ISIS database. In order to meet the needs of the database (which controls much of the display of the classification including the management of built classmarks and examples of combination) the structure of the classification data has been minutely examined; a number of conceptual problems have emerged related to the handling of instantive relationships which have caused us to think hard about the conceptual structure. This is a particular problem in the humanities disciplines with their proliferation of unique and named

entities. Work on the data structure of UDC is described in several papers by Slavic and Inês Cordeiro (2004), and has been central to the work of the FATKS project described below.

Faceted classification and subject headings

Even the Library of Congress, the most conservative of institutions where its subject access tools are concerned, shows signs of the need to accommodate the new thinking. Although the Classification itself is unlikely to incorporate any elements of faceted structure, the Library of Congress Subject Headings (LCSH) do lend themselves to its application. The LCSH context is perfectly hospitable to the use of standard categories of terms and the application of consistent orders of combination, since there is already some regularity of practice in this area. Facet principles were used for subject headings early on, both in the development of systems such as PRECIS, and in the development of a theory for the alphabetical approach alongside the systematic (Coates, 1960). The Faceted Access to Subject Terminology (FAST) (FAST, 2005a) project makes some progress along the road to consistent analytico-synthesis, although it is not faceted in the sense that most UK professionals would recognise.

The project aims to simplify the complex and sometimes inconsistent structure of LCSH by rationalising the way in which headings can be constructed; this is with a view to making them more accessible to the untrained end-user, or those not very conversant with LCSH practice:

> The first phase of the FAST development includes the development of facets based on the vocabulary found in LCSH topical and geographic headings and is limited to six facets: topical, geographic, form, period, with the most recent work focused on faceting personal and corporate names. This will leave headings for conference/meetings, uniform titles and name-title entries for future phases. With the exception of the period facet, all FAST headings will be fully established in a FAST authority file (FAST, 2005b).

This is clearly not a model of facet analysis that is familiar to a British audience. Facet analysis here is not at all concerned with managing the complexity of subject description but rather with the consistent application of rules for the construction of headings; the emphasis is on the order of combination, together with vocabulary control, both of subject terms and of names. While certainly not within the mainstream of thinking, we should welcome the improvement in consistency and predictability of LCSH structured headings, if only because they are the most widely-used headings in the world, and at present they can be very puzzling indeed to the novice user.

The faceted thesaurus

At the time of the CRG statement, the thesaurus was only just emerging as a tool for retrieval. In the 1950s most post-coordinate indexing was done using keyword lists, these often consisting of terms extracted from the documentation itself, and lacking any system of cross referencing, vocabulary control or underlying structural principles.

Most of the features of the thesaurus which address the first two of these aspects were established, along with the standard format, by the *Thesaurus of Engineering and Scientific Terms* (*TEST*) in 1967 (Engineers Joint Council and US Department of Defense, 1967). The appendices to *TEST* explain the procedures used in constructing the thesaurus, the decisions made as to the choice of preferred terms, and many aspects

of the form of entry; the hierarchical and associative relationships are also discussed. In addition to the alphabetical display of terms *TEST* had a rudimentary systematic display, albeit with a very limited number of categories at a broad subject level. This supplementary "subject" display became a feature of thesauri from this period onwards, but at first the two were not necessarily developed in conjunction with each other.

The first example of an "integrated' thesaurus and classification was compiled at the library of the English Electric Company in 1969. A faceted classification scheme had been developed for the library some years earlier, and it was decided to create a thesaurus of the terms in the classification, and to publish the two as a single entity. This dual tool was named *Thesaurofacet* (Aitchison *et al.*, 1969) and it seems to be the first example of a thesaurus systematically derived from a classification.

It was followed by several more faceted thesauri which were mainly the work of Jean Aitchison (1977, 1996), and soon the technique was well established as a methodology for thesaurus construction. Whilst working on a thesaurus for the (then) Department of Health and Social Security (DHSS) using the Department's classification scheme, the Health sciences class of the newly published second edition of the BC2, Aitchison published a formal statement of the methodology of converting BC2 to a thesaurus (Aitchison, 1986). She subsequently elaborated on this, and a more developed and sophisticated form of thesaurus construction using a faceted classification as a starting point is contained in the standard UK manual (Aitchison *et al.*, 2000).

The largest and best known online thesaurus, the *Art and Architecture Thesaurus* (*AAT*) hosted by the Getty Institute (1994) is also built on faceted principles. The default display looks very much more like a standard classification than the conventional thesaurus format, although the system can be searched for individual terms and produces a typical thesaurus entry record:

> Facets constitute the major subdivisions of the AAT hierarchical structure. A facet contains a homogeneous class of concepts, the members of which share characteristics that distinguish them from members of other classes. For example, *marble* refers to a substance used in the creation of art and architecture, and it is found in the Materials facet. *Impressionist* denotes a visually distinctive style of art, and it is found in the Styles and Periods facet (Getty Institute, 1994).

The facet structure of *AAT* is very clearly explained on the site, and the correspondence between *AAT*'s facets (physical attributes, styles, periods, agents, activities, materials, objects) and those of "standard" facet analysis is very evident, although the activities facet merges the standard categories of operation and process, and style and period are also collocated where they might be more clearly separated in a classification. The "hierarchies" of *AAT* are equivalent to sub-facets or arrays in the UK tradition.

It is striking that the Getty Institute no longer devotes any part of the vocabularies pages to an explanation of facet analysis; certainly five years ago they did feel the need to do this, and one can only assume that the principles of facet analysis are now sufficiently well known not to require explanation.

It is now firmly established that a classification of some sort is the best starting point for the construction of a thesaurus. Not only does the classification provide a source of vocabulary for the thesaurus, the very structure of the classification helps the identification of the relationships between terms that is

essential to the thesaurus. Any classification might in theory do the job, but faced with disentangling the pre-coordinated classes in an enumerative classification, the thesaurus compiler may well find the "top-down" structure less than ideal for clearly identifying relationships. The faceted classification on the other hand is ideal for the purpose. The assignment of terms to functional facets in the faceted scheme means that all the terms in a facet should be of the same facet "value" (i.e. they are all processes or all agents). Hence, by definition, the only relationships to be displayed intra-facet will be hierarchical ones. These will be predominantly of the BT/NT type, although terms in array, and other hierarchically coordinate terms will of course generate RTs or associative terms. These relationships are evident from the schedule display. Relationships between terms from different facets are de facto associative relationships.

In practice the generation of a thesaurus from its equivalent faceted classification is almost as automatic a process as thesaurus construction can ever hope to be. All decisions about the hierarchical relationships and most of the associative ones have been made in the process of assigning the concepts in the initial analysis. The faceted structure removes much of the ambiguity concerning associative relationships between terms in different facets, and does much of the preliminary work of synonym control. The classification can nevertheless create difficulties in the form of entry if the classification has been created without the possibility of conversion to a thesaurus, since the more liberal (and sometimes cumbersome) use of words in the class heading or caption is not suitable for thesaurus entry.

Nonetheless the relationship between the thesaurus and an underlying classification is now so well established that the working party for the revision of the British Standard for monolingual thesauri proposes to re-label this as the British Standard for structured vocabularies, and will include a model for classification construction as part of the new standard. They are also investigating the wider range of RTs that are exposed by the use of facet analysis in the systematic structure; whereas in the standard thesaurus format no subdivision of the associative terms is possible, in a faceted thesaurus it is perfectly easy to identify many more specific examples of this category, such as the "operation-agent" relation, "entity-process" relation, or "operation-product" relation. In fact any combination of facets can generate a relationship of this type, except for the thing-kind-part facets which are already accommodated by the BT/NT relationship and its variants narrower term instance (NTI) and narrower term part (NTP).

The new draft NISO standard for thesaurus construction *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies ANSI/NISO Z39.19-200x* has an interesting and slightly unusual perspective on the role of facet analysis. It acknowledges the part played by facet analysis, stating that:

> Controlled vocabularies – especially large ones consisting of thousands of terms – may be easier to use if they are organised in some way other than hierarchically (NISO, n.d.).

A brief description of the origins of facet analysis follows, and then a summary of its potential applications:

> Facet analysis is particularly useful for:
> - new and emerging fields where there is incomplete domain knowledge, or where relationships between the content objects are unknown or poorly defined;

- interdisciplinary areas where there is more than one perspective on how to look at a content object or where combinations of concepts are needed;
- vocabularies where multiple hierarchies are required but can be inadequate due to difficulty in defining their clear boundaries; or
- classifying electronic documents and content objects where location and collocation of materials is not an important issue (NISO, n.d.).

The understanding here seems to be that the faceted structure is less specific and rigorous than its comparable enumerative counterpart, whereas one would normally consider that it is more so. Certainly it would seem difficult to construct a faceted scheme where the relationships between document content cannot be established. It seems likely that the comparison is really between the pre- and post-coordinate forms of controlled languages, the faceted language being perceived as of the latter type, with the relationships implicit, rather than actualised as in the pre-coordinate system.

The draft of the standard brings out another interesting (and rather typically US) concept of the facet, in that it extends the meaning of facet to encompass various non-semantic aspects of a document. The draft says that "facet analysis is sometimes used to indicate the attributes of content objects" (NISO, n.d.), and lists as potential facets: topic, author, location, format, language, and place of publication. This list, which would hardly be entertained by most UK facet analysts, has a great deal in common with the FAST project (see above), where "topic" (or subject) is regarded as one facet among a list of non-subject elements of bibliographic description. The notion of facet here seems to be more or less equivalent to a database field, and in fact these "fundamental facets" can easily be mapped onto the MARC fields or those of Dublin Core. This application of facet analysis seems to be unique to the USA.

Faceted classification and the web

The web has now become a major vehicle for the dissemination of information about faceted classification, and, in addition to a number of "faceted classification" web sites such as that of the Knowledge Management Connection (www.kmconnection.com/DOC100100.htm), one can even find bibliographies of faceted classification there (Denton, 2003; Fast *et al.*, 2003). It is also the case that faceted classification has itself become an important method of information organisation and display on the web. This seems to be a fairly recent development; Foskett writing in 2000 on the future of faceted classification makes no reference to it at all, nor indeed to other electronic applications of facet analysis, asking instead "will the use of computers in information retrieval make the work of classification redundant?" (Foskett, 2000, p. 78).

The logical and predictable structure of the faceted system undoubtedly makes it compatible with the requirements of mechanisation in a way that enumerative and pre-coordinated systems are not. Even where linear order is not a major consideration and the aspects of the classification related to combination and display of compounds are very much secondary, the simplicity and logic of the faceted approach is appealing.

The usefulness of a faceted approach in automated information retrieval had been appreciated early on. Its appropriateness to a managed environment is fairly evident, and we have seen above (*Draft NISO Thesaurus Standard)* that a faceted structure bears some similarity to the field based structure of databases and document description templates associated with them. More than ten years ago Godert (1987, 1991) and Ingwersen and Wormell (1992) had tested faceted structures in database

searching, and concluded that they greatly facilitated efficient retrieval. Ingwersen and Wormell (1992, p. 199) were able to state with confidence that:

> ... the discussion demonstrates the suitability of the faceted categorization, not only for textual documents, but also with other forms of carriers of information. Faceted categorization may provide multi-dimensional and hence structured entry points to document contents, and thus give intellectual access to generated and stored knowledge.

By the 1990s some had begun to consider the benefits of the faceted approach to unmanaged digital resources, and how a "culture" of facet structure could inform searching and browsing as well as indexing and digital object description. Several papers by Ellis and Vasconcelos (1999, 2000) address the idea of facet analysis as applied to the web, and conclude that "it can alleviate some problems in searching the WWW by being applied to using subject directories or search engines" (Ellis and Vasconcelos, 2000, p. 111).

The following section of the discussion considers how facet analysis is, or might be, exploited as a web tool, and where this requires some human intervention in the process. This intervention could be of a conventional nature (i.e. using indexer assigned metadata), or the broader categorisation of resources, or using visual displays structured in a faceted manner; here the emphasis is on guided navigation, browsing and, to some extent, query formulation. The other approach has focused on a faceted knowledge structure underlying the interface and not necessarily visible to the end-user; here the interest has been in query modification (usually by means of mapping to the control language) via semantic expansion techniques and a more sophisticated system syntax.

The first approach is seen largely in the application and actual use of faceted structures both in the academic world (where it has been used to manage digital libraries and portals of various kinds) and in the commercial sector where it is increasingly frequently encountered as a navigation tool particularly in internet selling. The second approach is principally to be met with in research projects where the emphasis is on conceptual work and the development of the facet supported structure.

Faceted structures as navigational tools

The advent of Windows technology in the 1990s brought with it a variant on search techniques of an earlier age. The user was now able to employ a hybrid "browse-and-select" search technique, with available options displayed via drop-down menus. This was the perfect vehicle for a faceted structure, and what began as a means of on-screen help, and a navigational aid of a fairly basic kind, soon was adopted as the basis of a more sophisticated kind of searching. This employed multiple drop-down windows based on a faceted structure that allowed the searcher to browse the conceptual structure of the information store, and, more importantly, to combine concepts from different facets (windows or menus). The technique was therefore not only a browsing and navigational aid, but also a tool for query formulation.

The architect of this approach was Stephen Pollitt (Pollitt *et al.*, 1996, 1998), and he called it "view-based searching". View-based searching based on Windows is now a very well established means of information structuring on the web. The technique can operate at several different levels; in its simplest manifestation there is a linear progression through the material with successive filtering of retrieved items by the

addition of new search criteria; in a more sophisticated model the starting point can be varied and the "extra facets' added in any order.

In the very typical example given in Figure 1 terms from three facets from the field of medicine (disease by part of the body, therapy, and patient by age) can be viewed simultaneously by the searcher and combined as wished; further hierarchies can also be opened by clicking on the folders.

The view-based system has taken the facility of facet analysis to reduce the multi-dimensionality of subjects to a linear order, and moved it up a gear to do the same in a two-dimensional graphical interface. It also incidentally provides a visual model of the data structure that is more immediately grasped by the end-user, and thus brings him closer to an understanding of the structure of the subject.

Faceted portals and web sites

This view-based approach underlies most of the "faceted" classifications to be seen on the world wide web. The wine retailing site shown in Figure 2 is fairly typical of the genre, with its four "facets" of type, region, winery, and price, and it is commonly referred to within the sector as a model of a faceted classification. The method is described more fully in a number of web published articles, mainly written by computer specialists, knowledge managers and others from outside the traditional documentation world.

Recently a number of writers on classification and knowledge organisation have studied this type of application. LaBarre (2004) has made an in-depth scholarly study of the phenomenon of faceted classification of all types on the Web. Adkisson (2005), in

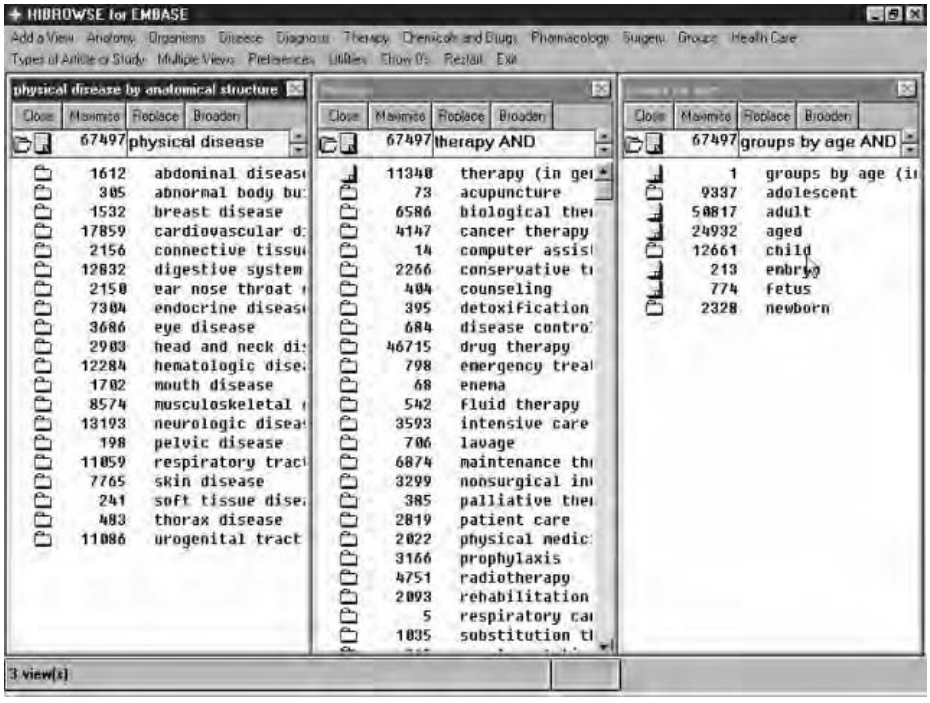


Figure 1.
View-based searching in the EMBASE database

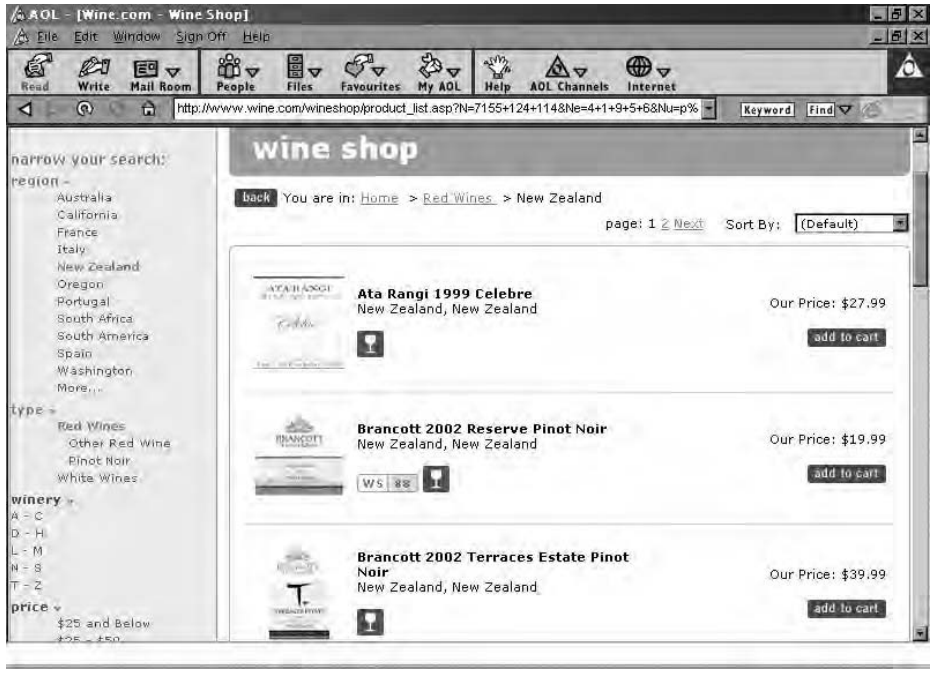


Figure 2.
Faceted navigation on the Wine Shop site (http://www.wine.com/)

a briefer survey of 75 e-commerce sites, discovered that 69 per cent of them made use of some form of faceted classification of products:

> Unlike a simple hierarchical scheme, faceted classification gives the users the ability to find items based on more than one dimension. For example, some users shopping for jewelry may be most interested in #browsing by particular type of jewelry (earrings, necklaces), while others are more interested in browsing by a particular material (gold, silver). "Material" and "type" are examples of *facets*; earrings, necklaces, gold, silver are examples of facet *values*.

The type of interface used is very similar to Pollitt's original model, employing drop down menus. The complexity of this can be quite variable, and Adkisson identifies the existence of several modes of use of faceted systems:

> Of sites providing faceted navigation 67 percent did so at a single point in the browse path. For example, on a top-level page, the user may have the option to browse by brand or category, but no additional faceted navigation options are presented along the browse path. 28% of sites providing faceted navigation at more than one point in the browse path, creating a progressive filtering experience based on multiple criteria. 4% of sites providing faceted navigation presented multiple filtering options on the page in a search-like interface. The user selects one or more values from a pull-down, clicks submit, and the page displays a filtered list of links based on the selected values.

Of course this would not be regarded as faceted classification by purists; what is classified here are the objects themselves, and not subjects or documentation. As a result the classification is relatively simple, and a matter of organisation of different

attributes of the objects. This statement from a web site about faceted classification typifies this view:

> Now, faceted classification isn't inherently innovative. In fact, objects tend to have a fixed set of facets by which they are organized. Where innovation comes is through user research that listens to how the users/customers/audience think about and approach a task, and providing tools to allow them to approach it meaningfully (Merholz, 2001).

We should regard this as a classification with only one facet (that of the object to be organised) and the classification is restricted to the identification and labelling of arrays within that facet. Adkisson's facet values here are equivalent to the terms in array. Although in this example she gives two facets: type and material, which might be regarded as equivalent to two fundamental facets, this is not in fact the case. The only thing under consideration is the type of jewellery, either type (by function) or type (by material). Nowhere is there any information about the material itself (gold, silver, etc.).

Ontological applications of facet analysis

A more rigorous approach to information retrieval tools is displayed in the creation and application of ontologies. In describing the use of the ontology McGuinness (2002) specifies a number of potential ways in which it may be used:

- as a controlled vocabulary;
- for site navigation and support;
- to set expectations (i.e. provide a quick overview of the site);
- as umbrella structures from which to extend content;
- for browsing support;
- for search support; and
- to sense (i.e. semantic) disambiguation support.

Thus it appears that the overall purpose of the ontology, and what may be expected of it, is broadly similar to that of a faceted knowledge structure.

Prieto-Diaz (2002) says that "building ontologies is more difficult than it seems", and outlines three stages in the building of the ontology:

(1) ontology capture;

(2) ontology coding; and

(3) ontology integration.

Facet analysis certainly addresses items one and three, which involve the collection of concepts, specification of relationships, insertion of terminology from external sources, and mapping onto those sources. Although faceted classifications use formal coding systems to express both content and relationships (through the use of notations and facet indicators), the coding here referred to is mathematically based, allowing for some manipulation of the ontology. Prieto-Diaz's (2002) definition of an ontology as "a taxonomy where the meaning of each concept is defined by specifying properties, relations to other concepts, and axioms narrowing down the interpretation", puts the ontology firmly into the same category as the faceted structure, and interestingly he makes the same comparison with the more limited structure of the taxonomy. His

methodology for building a simple ontology would be recognised immediately by any LIS student required to construct a simple faceted classification, and provides us with a practical demonstration of the role that may be played by facet analysis in organising the domain of the ontology.

McGuinness (2002, p. 4) is particularly concerned with the exact expression of relationships, and is critical of the Yahoo hierarchy:

> ... the general category apparel includes a subcategory women (which should be more accurately titled women's apparel) which then includes subcategories accessories and dresses. While it is the case that every instance of a dress is an instance of apparel ... it is not the case that a dress is a woman, and it is also not the case that a fragrance (an instance of a woman's accessory) is an instance of apparel ... Without true subclass (or true 'isa') relationships, we will see that certain kinds of deductive uses of ontologies become problematic.

This is a restatement of the criticism of the enumerative classification with its unclear hierarchies and mixed principles of division by facet analysts from the mid-twentieth century onwards. In a more recent paper (McGuiness and Noy, n.d.), McGuiness is more specific about the relationships in the ontology, although the terminology used is very different to that of the LIS world. "Facet" is used without being defined, but seems to be equivalent to "array". McGuiness refers to "object properties that can become slots" (principles or characteristics of division, and facets or arrays) and identifies the part relationship. "Common facets" are stated to include:

- slot cardinality (the number of "values" a slot can have);
- slot value type (string, number, Boolean, enumerated, instance-type); and
- domain and range of a slot.

The correspondence with the concepts of facet analysis is evident if the terminology used is specific to this discipline. The discussion also covers other relationships, principally those of the hierarchy, and is concerned with the accuracy of the broader and narrower classes, the siblings in a "class hierarchy", multiple inheritance, and the distinctions between classes and instances; naming problems are also included.

It would appear that the ontology is conceptually very close to the faceted structure, and that facet analysis may have much to offer as a method for building ontologies. Surprisingly, the number of relationships covered in the ontology does appear more limited than those inherent in the faceted scheme.

The two examples given here give only a taste of the work on facet analysis and ontologies; other notable projects include that of Damiani *et al.* (1999), the Ontosaurus project (Swartout *et al.*, 2005), and an interesting combination of ontology and view-based searching carried out by the Finns in connection with the Finnish Museums Portal (Hyvonen *et al.*, 2004).

Closely related to the ontology is the topic map, although this type of structure had been around for rather longer – long enough to have generated an ISO standard for the format (ISO/IEC, 1999). Topic maps also acknowledge the role of facet analysis as a basis for organising concepts, and the understanding of facet is similar to that employed by the ontology; the standard defines facet as the:

- subset of information objects that share an externally applied property; and
- values given to a particular property externally applied to a set of information objects.

This property may be referred to as the facet type, and a facet value is a specific member of the set of values attached to the facet type (i.e. a member of the set defined by the facet type).

The topic map has given rise to a mark-up language for the dissemination and exchange of faceted metadata, exchangeable faceted metadata language or XFML. The specification for XFML core (Van Dijck, 2002) covers the conceptual model, an XML format, and instructions for its application. The XFML concept of a facet sits somewhere between the topic map model and the LIS one. The XFML specification notes that:

> ... the term "facets" is used with somewhat different meanings by different people ... The term "facet" in XFML is used very much like the library-science definition of the term.

More far-reaching applications of facet analysis: the search for solutions to the problem of the semantic web

While the above examples of web applications show the usefulness of faceted structures as organisational and navigational tools, researchers from a number of fields have begun to examine the technique for its potential in more sophisticated searching and retrieval, particularly in the search for tools to underpin automated systems.

Tzitzakas *et al.* (2004) create a fairly conventional looking faceted structure dealing with aspects of tourism; facets include locations and sports. Within the context of this taxonomy they investigate the possibilities for the compounding of terms, noting that compounds may be created between terms from different facets, but also between terms in the same facet, and that some combinations are valid while others are not. They construct an algebra consisting of four operations which can be used to control the combination of terms; in other words it formalises the system syntax of the faceted scheme. They are quite clear that this is distinct from other algebras developed for ontology engineering. The use of the algebra can generate dynamically navigation trees which are suitable for browsing and can also be exploited during the indexing process (to aid the indexer and prevent indexing errors).

In the second example Ali and Du (2004, p. 501) describe a faceted scheme that is also fairly close in conception to the LIS model, using six facets related to object-oriented software to create a classification. They note that the "major advantages of faceted classification are flexibility, expandability, adaptability and consistency", which is familiar enough, but also suggest that "another important concept of faceted systems is the use of a conceptual distance graph. This graph can be used to calculate the degree of closeness between the descriptor of the target component and that of components in the repository".

The phenomenon of semantic closeness has also been investigated in the FACET project (Binding and Tudhope, 2004) at the University of Glamorgan. Working with the faceted *Art and Architecture Thesaurus*, Binding and Tudhope have shown how a thesaurus can be incorporated into a search interface to support query formulation. User input can be mapped to the controlled vocabulary, the systematic structure can be browsed, and a view-based interface aids query building. The FACET demonstrator has also exploited the relationships of terms in array to allow a degree of "best-match" searching based on computed degree of similarity between concepts. Their

demonstrator allows the searcher to frame the search, and ranks results based on the degree of correspondence of automatically coordinated searches with the original input. Thus there is a strong element of "conceptual" searching, since the system will return results even where input and index terms do not match, and even where there is some difference in the exact meaning of terms.

The FATKS project[3] at SLAIS looked at the potential of faceted controlled languages as tools for digital humanities resource management within a subject portal (Broughton, 2002). One of the services involved, the Arts and Humanities Data Service, had a digital repository of millions of objects, and although it had tried to use conventional subject cataloguing tools, these had turned out to be inadequate, both for complete subject description and for retrieval purposes. Although the original project involving the merger of the two UK humanities-related portals did not ultimately come to fruition, the project has provided us with an opportunity to look at some aspects of controlled vocabulary management in a digital context, and has thrown up some interesting aspects of problems associated with faceted structures in the humanities:

- existing faceted vocabularies require further rationalisation before they are suitable for machine use;
- in particular, the structure of both BC2 and UDC contain many examples of enumerated instances within built classes which cannot be handled in this manner;
- since this is common in the humanities alternative means of coping must be found;
- it is possible to build a database that accommodates the complexity of both the relationships and the syntax in a faceted vocabulary; and
- encoding of facet status can allow a considerable measure of automatic manipulation of the system syntax, and hence of query formulation.

Conclusion

It is clear that faceted classification in some form or another now plays an integral part in most methods of information retrieval. It is very well established as a method of construction in classification schemes and thesauri, and has affected the development of even the most conservative of systems in the area of traditional document description and organisation. It is popular as a navigational tool for web sites of all sorts, helping to structure all manner of objects and information about them, from children's shoes to the artistic output of the High Renaissance. It is beginning to be taken up by researchers in the fields of automatic indexing and the semantic web as a conceptual tool to assist in the understanding of the most complex relationships between objects.

What constitutes a facet may be very variously interpreted: it can be no more than a name for the fields in a basic bibliographic format or metadata template, where subject is a single facet among others concerned with the structure, provenance, and identification of the object; it may be limited to the listing and display of physical objects based on their various properties, albeit in a highly structured and regular manner; it can engage with the whole range of fundamental categories of classical facet theory to create highly complex but ordered models of the information universe that begin to support some degree of automated object description and retrieval.

Facet analysis is significant for the clarity of the light it shines upon the relationships between objects and entities, and abstract concepts and their associated labels. It gives a rational, scientific, methodology for the construction of systems; it enables the full and precise description of objects of considerable structural complexity and of multi-dimensional semantic composition; it provides a flexible syntactical apparatus for the combination and ordering of concepts where this is required.

The faceted system can function as a tool for browsing, for navigation and for retrieval; it can act as a means of spatial organisation of information in both physical and digital stores; it has a role as a standard in vocabulary control, for mapping onto and between terminologies, and in query formulation and modification; it is also a powerful means of information retrieval.

In 50 years it has advanced from utilising these characteristics to fulfil the basic need to reduce complexity to linear order: what one might call the faceted structure in one dimension of the 1930s to the 1970s. In the 1980s and onwards it provided a visual model of the information store in two dimensions through the medium of Windows-based graphical interfaces. In the twenty-first century it has begun to be exploited in three or more dimensions through vehicles such as the ontology and other mathematical structures.

Notes

1. Kaiser (1911) identifies concretes, processes and place as categories of indexing terms. Also Otlet and La Fontaine (1905) in the first edition of UDC introduced a whole range of generally applicable tables for such categories as place, time, form, language, and point of view.
2. The library of Christ's College Cambridge is a notable exception in the UK, but in 2005 it is now almost completely replaced by the Library of Congress classification.
3. Full details of the Facet Analytical Theory in Knowledge Structures (FATKS) project, structured vocabularies and the prototype classification database and search interface are available on the project web site at www.ucl.ac.uk/fatks

References

Adkisson, H.P. (2005), "Web design practices: use of faceted classification", available at: www.webdesignpractices.com/navigation/facets.html (accessed April 2005).

Aitchison, J. (1977), *Unesco Thesaurus: A Structured List of Descriptors for Indexing and Retrieving Literature in the Fields of Education, Science, Social Science, Culture and Communication*, Unesco, Paris.

Aitchison, J. (1986), "A classification as a source for a thesaurus: the bibliographic classification of H.E. Bliss as a source of thesaurus terms and structure", *Journal of Documentation*, Vol. 42 No. 3, pp. 160-81.

Aitchison, J. (1996), *International Thesaurus of Refugee Terminology*, 2nd ed., United Nations High Commissioner for Refugees, Geneva (compiled under the auspices of the International Refugee Documentation Network, New York, NY).

Aitchison, J., Gilchrist, A. and Bawden, D. (2000), *Thesaurus Construction: A Practical Manual*, 4th ed., Aslib, London.

Aitchison, J., Gomershall, A. and Ireland, R. (1969), *Thesaurofacet: A Thesaurus and Faceted Classification for Engineering and Related Subjects*, English Electric, Whetstone.

Ali, F.M. and Du, W. (2004), "Toward reuse of object-oriented software design models", *Information and Software Technology*, Vol. 46 No. 8, pp. 499-517.

Austin, D. (1984), *PRECIS: A Manual of Concept Analysis and Subject Indexing*, 2nd ed., The British Library Bibliographic Services Division, London.

Binding, C. and Tudhope, D. (2004), "KOS at your service: programmatic access to knowledge organization systems", *Journal of Digital Information*, Vol. 4 No. 4, available at: http://jodi.tamu.edu/Articles/v04/i04/Binding/ (accessed April 2005).

Bliss, H.E. (1910), "A modern classification for libraries, with simple notation, mnemonics and alternatives", *Library Journal*, Vol. 35, pp. 351-8.

Broughton, V. (2000), "A new classification for religion", *International Cataloging and Bibliographic Control*, No. 4, pp. 2-4.

Broughton, V. (2002), "Organizing a national humanities portal; a model for the classification and subject management of digital resources", *Information Research Watch International*, June, pp. 2-4.

Broughton, V. (2004), "Faceted classification", *Essential Classification*, Facet, London, pp. 257-83.

Classification Research Group (1955), "The need for a faceted classification as the basis of all methods of information retrieval", *Library Association Record*, Vol. 57 No. 7, pp. 262-8.

Classification Research Group (1964), *Some Problems of a General Classification Scheme*, report of a conference held in London, June 1963, Classification Research Group, London.

Coates, E.J. (1960), *Subject Catalogs: Headings and Structure*, Library Association, London.

Damiani, E., Fugini, M.G. and Bellettini, C. (1999), "A hierarchy-aware approach to faceted classification of object-oriented components", *ACM Transactions on Software Engineering and Methodology*, Vol. 8 No. 3, pp. 215-62.

Denton, W. (2003), "Putting facets on the web: an annotated bibliography", October, available at: www.miskatonic.org/library/facet-biblio.html (accessed April 2005).

Ellis, D. and Vasconcelos, A. (1999), "Ranganathan and the net: using facet analysis to search and organise the world wide web", *Aslib Proceedings*, Vol. 51 No. 1, pp. 3-10.

Ellis, D. and Vasconcelos, A. (2000), "The relevance of facet analysis for world wide web subject organization and searching", *Journal of Internet Cataloging*, Vol. 2 Nos 3/4, pp. 97-114.

Engineers Joint Council and US Department of Defense (1967), *Thesaurus of Engineering and Scientific Terms: A List of Engineering and Related Scientific Terms and Their Relationships for Use as a Vocabulary Reference in Indexing and Retrieving Technical Information*, Engineers Joint Council and US Department of Defense, New York, NY.

Faceted Application of Subject Terminology (FAST) (2005a), available at: www.oclc.org/research/projects/fast/default.htm

Faceted Application of Subject Terminology (FAST) (2005b), "Project documentation", available at: www.oclc.org/research/projects/fast/international_auth200302.doc (accessed April 2005).

Fast, K., Liese, F. and Steckel, M. (2003), "Facets and controlled vocabularies: an annotated bibliography", available at: www.boxesandarrows.com/archives/Facets_CV/Bibliography.htm (accessed April 2005).

Foskett, A.C. (2000), "The future of faceted classification", in Marcella, R. and Maltby, A. (Eds), *The Future of Classification*, Gower, Aldershot.

Getty Institute (1994), *Art and Architecture Thesaurus*, Oxford University Press, New York, NY, available at: www.getty.edu/research/conducting_research/vocabularies/aat/ (accessed April 2005).

Gödert, W. (1987), "Klassificationssysteme und Online-Katalog (Classification systems and the on-line catalogue)", *Zeitschrift für Bibliothekswesen und Bibliographie*, Vol. 34 No. 3, pp. 185-95.

Gödert, W. (1991), "Facet classification in online retrieval", *International Classification*, Vol. 18 No. 2, pp. 98-109.

Hjorland, B. (2002), "The method of constructing classification schemes: a discussion of the state-of-the-art", in López-Huertas, M.J. and Muñoz-Fernández, F.J. (Eds), *Challenges in Knowledge Representation and Organization for the 21st Century. Integration of Knowledge across Boundaries, Proceedings of the the 7th International ISKO Conference, Granada, Spain, 10-13 July 2002*, *Advances in Knowledge Organization*, Vol. 8, Ergon Verlag, Würzburg, p. 451.

Hyvonen, E., Saarela, S. and Viljanen, K. (2004), "Application of ontology techniques to view-based semantic search and browsing", *Proceedings of the 1st European Semantic Web Symposium, Heraklion, Greece, May 10-12, 2004*, Springer-Verlag, Berlin.

Ingwersen, P. and Wormell, I. (1992), "Ranganathan in the perspective of advanced information retrieval", *Libri*, Vol. 42, pp. 184-201.

ISO/IEC (1999), *Topic Maps ISO/IEC 13250*, ISO/IEC, Geneva, December.

Kaiser, J. (1911), *Systematic Indexing*, Pitman, London.

LaBarre, K. (2004), "Adventures in faceted classification: a brave new world or a world of confusion?", in McIlwaine, I.C. (Ed.), *Knowledge Organization and the Global Information Society, Proceedings of the 8th International Conference of the International Society for Knowledge Organization, University College London, 13-16 July, 2004*, *Advances in Knowledge Organization*, Vol. 9, Ergon, Würzburg, pp. 79-84.

McGuinness, D. (2002), "Ontologies come of age", *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*, MIT, Boston, MA, available at: wwww.ksl.stanford.edu/people/dlm/papers/ontologies-come-of-age-mit-press (accessed April 2005).

McGuiness, D.L. and Noy, N.F. (n.d.), "Ontology development 101: a guide to creating your first ontology", available at: http://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html (accessed April 2005).

McIlwaine, I.C. and Williamson, N.J. (1993), "Future revision of UDC: progress report on a feasibility study for restructuring", *Extensions and Corrections to the UDC*, UDC Consortium, The Hague, pp. 11-17.

McIlwaine, I.C. and Williamson, N.J. (1994), "A feasibility study on the restructuring of the Universal Decimal Classification into a full-faceted classification system", in Albrechtsen, H. and Oernager, S. (Eds), *Proceedings of the 3rd International Society for Knowledge Organization (ISKO) Conference: Knowledge Organization and Quality Management, Copenhagen, Denmark, 20-24 June, 1994*, *Advances in Knowledge Organization*, Vol. 4, INDEKS Verlag, Frankfurt/Main, pp. 406-13.

McIlwaine, I.C. and Williamson, N.J. (1995), "Restructuring of class 61 – medical sciences", *Extensions and Corrections to the UDC*, UDC Consortium, The Hague, pp. 11-66.

Merholz, P. (2001), "Innovation in classification", available at: www.peterme.com/archives/00000063.html (accessed April 2005).

Miksa, F.L. (1998), *The DDC: the Universe of Knowledge and the Post-modern Library*, Forest Press, Albany, NY.

Mills, J. and Broughton, V. (1977), *Bliss Bibliographic Classification*, 2nd ed., Butterworth, London.

Mitchell, J.S. (1996), "Introduction", *Dewey Decimal Classification and Index*, 21st ed., Forest Press, Albany, NY.

NISO (n.d.), "Facet analysis", *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies ANSI/NISO Z39.19-200x*, NISO, Bethseda, MD, (draft version of the thesaurus standard, section 5.3.4).

Otlet, P. and La Fontaine, H. (1905), *Universal Decimal Classification*, 1st ed., Institut International de Bibliographie, Brussels.

Pollitt, A.S., Smith, M.P. and Braekevelt, P. (1996), "View-based searching systems: a new paradigm for information retrieval based on faceted classification and indexing using mutually constraining knowledge-based rules", in Johnson, C. and Dunlop, M. (Eds), *Information Retrieval and Human Computer Interaction (Proceedings of the Joint Workshop of the Information Retrieval and Human Computer Interaction Specialist Groups of the British Computer Society)*, GIST Technical Report G96-2, Glasgow University, Glasgow, pp. 73-7.

Pollitt, A.S., Tinker, A. and Braekevelt, P. (1998), "Improving access to online information using dynamic faceted classification", *Online Information 98: Proceedings of 22nd International Online Information Meeting, London, 8-10 December, 1998*, Learned Information Europe, Oxford, pp. 17-21.

Prieto-Diaz, R. (2002), "A faceted approach to building ontologies", paper submitted to the 21st International Conference on Conceptual Modeling, Tampere, available at: https://users.cs.jmu.edu/prietorx/public/RubenPubs%5B1%5D.html (accessed April 2005).

Priss, U. (2000), "Comparing classification systems using facets", in Beghtol, C., Howarth, L.C. and Williamson, N.J. (Eds), *Dynamism and Stability in Knowledge Organization. Proceedings of the 6th International ISKO Conference, Toronto, Canada, 10-13 July, 2000*, *Advances in Knowledge Organization*, Vol. 7, Ergon, Würtzburg.

Ranganathan, S.R. (1960), *Colon Classification Basic Classification*, 6th ed., Asia Publishing House, New York, NY.

Ranganathan, S.R. (1967), *Prolegomena to Library Classification*, 3rd ed., Ranganathan Series in Library Science, Vol. 20, Asia Publishing House, London.

Slavic, A. and Inês Cordeiro, M. (2004), "Core requirements for automation of analytico-synthetic classifications", in McIlwaine, I.C. (Ed.), *Knowledge Organization and the Global Information Society, Proceedings of the 8th International Conference of the International Society for Knowledge Organization, University College London, 13-16 July, 2004*, *Advances in Knowledge Organization*, Vol. 9, Ergon, Würzburg, pp. 187-92.

Swartout, B., Patil, R., Knight, K. and Russ, T. (2005), "Ontosaurus: a tool for browsing and editing ontologies", available at: http://ksi.cpsc.ucalgary.ca/KAW/KAW96/swartout/ontosaurus_demo.html (accessed April 2005).

Tzitzakas, Y., Analyti, A., Spyratos, N. and Constantopoulos, P. (2004), "An algebraic approach for specifying compound terms in faceted taxonomies", *Information Modelling and Knowledge Bases XV, Proceedings of EJC 2003*, IOS Press, Amsterdam/Washington, DC, pp. 67-87.

Van Dijck, P. (2002),)," XFML Core – eXchangeable faceted metadata language 2002", available at: www.xfml.org/spec/1.0.html (accessed April 2005).

Corresponding author
Vanda Broughton can be contacted at: uczcvbr@ucl.ac.uk

Evaluating evaluation

Introducing a research project on the impact of *Improve Your Library: A Self-Evaluation Process for School Libraries*

Lucy Gildersleeves
School of Library, Archive and Information Studies, University College London, London, UK

Received 22 August 2005
Revised 7 October 2005
Accepted 17 October 2005

Abstract

Purpose – The aim of this article is to investigate the nature of use and impact of the *Improve Your Library* self-evaluation process on school libraries and their integration into whole-school development planning for pupil learning.

Design/methodology/approach – The research is structured in two phases. Phase one seeks to map levels, patterns and experiences of use by those responsible for the library within schools and to establish the nature of support needed for librarians from school library services and training providers. Phase two will track schools in detail to explore the extent of difference the process may have on library integration into the wider school self-evaluation, development priorities and planning, and on library contribution to pupil reading and learning.

Findings – Preliminary investigations indicate that there is considerable diversity in current evaluation practice and in expectations of the *Improve Your Library* toolkits. There is a need for a national vehicle for sharing individual library experience to inform reflective practice.

Research limitations/implications – The research focused on English schools. Successful delivery of the intended case banks of practice examples is dependent on a good range of schools being willing to share their experience.

Practical implications – Training and support needs and networks will be identified for providers and practitioners. Practical experience will be made available to librarians, school management and policy-makers.

Originality/value – This is a timely evaluation of the first unified system of library self-evaluation across English schools within the context of a new approach to whole-school evaluation. It will provide librarians with a means of sharing and comparing experiences in the absence of a formalised benchmarking system for school libraries and contribute to a national picture of school library activity and priorities.

Keywords Schools, School libraries, Information centres, Project evaluation, England

Paper type Research paper

Introduction – the drivers for research into the impact of self-evaluation of school libraries

In July 2004 the Department for Education and Skills (DfES) published a set of self-evaluation toolkits for school libraries in England. These toolkits, *Improve Your Library: A Self-evaluation Process for Primary Schools* (DfES, 2004a) and *Improve Your Library: A Self-evaluation Process for Secondary School Libraries and Learning Resource Centres* (DfES, 2004b) and accompanying support booklet were developed by Sharon Markless and David Streatfield as a result of work done by the DfES School Libraries Working Group (DfES, 2004a). The background and process of this

Aslib Proceedings: New Information Perspectives
Vol. 58 No. 1/2, 2006
pp. 73-88

0001-253X
DOI 10.1108/00012530610648680

development is fully explained in their article in *Update* (Streatfield and Markless, 2005). Essentially, the toolkits (hereafter referred to as *IYL*) provide those responsible for managing a school library (hereafter "librarian") with a set of frameworks, indicators and models which support library evaluation in a way which parallels the existing Ofsted frameworks for school evaluation. This is significant as it offers the opportunity for school libraries to be reviewed for their contribution to teaching and learning in a way that is credible to teachers and respected by those in education. Streatfield and Markless (2005, p. 38) identify a need for school librarians to pool the results of their self-evaluation to enable mutual comparison, sharing of ideas and to provide educational policy-makers with a body of evidence for advocacy and strategy development.

At the School Libraries Forum[1] in November 2004 Phil Jarrett, Ofsted Subject Advisor for English, indicated that inspectors would be considering school use of the *IYL* self-evaluation frameworks, particularly in identifying good practice schools (Jarrett, 2004). The new light-touch, drop-in inspection programme commencing September 2005 would require that schools keep their overall framework documentation up-to-date and available at all times. Forum members present from schools and school library services reported a very varied take-up of *IYL*. Members also noted differences in the ways that school library services were supporting schools in their use of the *IYL* toolkits. A need was identified for systematic research into the use of the frameworks by those responsible for school libraries and the ways in which the library frameworks were being integrated into whole-school self-evaluation and development.

In February 2005, David Bell, Chief Inspector of Schools, published his annual report, including the Ofsted subject reports for English 2003/2004. Bell (2005a) said:

> Senior managers do not take sufficiently close interest in the work of the school library or the impact it has on pupils' learning … Schools rarely ask librarians to report formally on library policy, use of resources and effectiveness. It is rare to find schools that have an effective library development plan.

Concerning primary schools, Bell (2005b) also noted:

> Too often, libraries do not have a significant impact on pupils' learning.

In both subject reports the Chief Inspector detailed a number of significant weaknesses in the school management and the exploiting of libraries as a key resource in pupil learning. He explicitly referred schools to the *IYL* toolkits to evaluate library effectiveness.

This research project therefore seeks to identify the nature of use of the *IYL* toolkits within school libraries, how this self-evaluation process is reported to school management and how it is integrated into wider school evaluation and development planning. The research aims to explore how *IYL* self-evaluation is taken into consideration within the inspection process and reporting, and also whether this has any impact on library integration into the curriculum and the collaborative role of the librarian in pupils' learning and assessment. It seeks to identify how school library services can best support schools in using the toolkits. A key outcome of the research will be the setting up of databases of good-practice examples and of practical-experience cases that will allow schools and school librarians to share and compare ideas, approaches and developments.

Literature review

There is no statutory requirement in England for schools to have a library (or learning resource centre) or for local education authorities to provide school library services. The context for evaluation of libraries in schools is therefore different from that in many countries. The absence of statutory requirement in England means that there are no standards or benchmarks formally laid down, such as exist for public libraries. As a result there is a certain political caution in specifying any minimum parameters for school library resourcing, qualification of school librarians or for defining how school librarians should contribute to pupil learning through teaching and supporting the curriculum and through wider pupil development.

Charles Clarke, then Secretary of State for Education and Skills stated in his forward to the CILIP (2004, p. vii) *Guidelines for Secondary School Libraries*:

> [. . .] the school library is the heart of a school, which itself has learning at its core, and good libraries can empower the learner.

The strength of this statement is validated by a substantial body of research in library impact conducted since the 1990s, particularly in the USA and in Canada. These studies have identified how good school libraries, staffed by qualified librarians, make a significant positive contribution to pupil learning and achievement at both primary and secondary levels. A convenient summary of the findings of studies conducted across fourteen American states is provided by *School Libraries Work!* (Research Foundation, 2004). These identified that a significant increase in pupil reading achievement is influenced by the amount of information literacy teaching provided by librarians, the degree of collaborative teaching, the full-time availability and involvement of school librarians and the quality and regular use of resources by pupils. In particular, research in Pennsylvania found that:

> [. . .] the mere presence of a large collection . . . in the school library is not enough to generate high levels of academic achievement by students. Such collections only make a positive difference when they are part of school-wide initiatives to integrate information literacy into the school's approach to standards and curricula (Lance *et al.*, 2000).

A similar point is noted by Bell (2005a) in his Ofsted subject reports on English. These findings underpin the current concern that evaluation of the library needs to identify not only what resources are provided and how much they are used, but also how this use is embedded into the wider process of pupils' learning. The Pennsylvania point reinforces the need for the library to be recognised as being in partnership with teaching at a strategic level.

Todd and Kuhlthau took investigation of impact further in their Ohio studies, seeking the perceptions of pupils on the difference libraries and librarians made to their curricular achievement, to reading interests and to success outside of school (Todd *et al.*, 2004). Todd has stressed the importance of involving the pupils directly in the research through writing "stories" which allowed them to say what they felt about how the library made a difference for them. This qualitative approach also enabled Todd and Kuhlthau to understand pupil patterns of resource use and where the library added value for pupils in ways which had not been anticipated. This approach has been of particular benefit in identifying ways in which librarians have been able to improve their own position within schools (Todd, 2005). Brophy (2005) has also valued "narratives" as a key research tool for getting inside problems experienced by

practitioners and for identifying significant issues which can then inform future practice.

Closer to home, Williams and Wavell (2001) carried out a study of library impact in Scottish secondary schools during 1999-2001. For Scottish schools there is a statutory requirement to provide libraries, and these have been covered for some years by a self-evaluation process. Williams and Wavell (2001, p. 3 S2.1) found that the school library (SLRC) is:

> [...] seen as having an active and supporting role in achieving [these] educational priorities but for librarians as well as teachers the nature of the potential SLRC contribution remains vaguely defined and raises wider questions about the relationship between the SLRC and the learning experience.

They noted that teachers and librarians did not have a common language for discussion of the pupil learning experience. This is a challenging observation, given the statutory place of the library within the school in Scotland, and suggests that a common language and clear shared understanding may be even more of a problem in the English context.

The CILIP (2002, p. 15) report *Start with the Child*, commenting on the use of evaluation processes, noted the importance of communicating findings to demonstrate value:

> [...] where [a scheme] is evaluated and reported widely, the commitment from partners and stakeholders will follow.

Bell (2005a) has noted with concern that:

> Schools rarely ask librarians to report formally on library policy, use of resources and effectiveness ... It is rare to find schools that have an effective library development plan.

Bell's observations were made in relation to inspections for the academic year 2003-2004 – three years after the introduction of specific guidance for inspecting school libraries which briefs inspectors to assess the corporate view of the use of the library and its impact on quality and standards (Ofsted, 2001). This "corporate view" includes consideration of how well the role of the library and library staff is represented in any whole-school development planning and language and literacy policy. It would seem that too often school libraries remain isolated from the wider school policy and learning strategy, despite awareness that Ofsted will be looking for this corporate view. However, the cyclical nature of inspections may mean that it will take some time for this focus to work through into identified improvements in schools. In his annual report Bell (2005a, b) specifically identified a role for *IYL* in providing the sound evidence base required for judgement and in building the systematic communication of library value into the wider school improvement planning. It is noted, however, that Ofsted is absent from the various education and library bodies which have endorsed *IYL* by "badging" the toolkit documents and that Jarrett (2005) stopped short of expecting routine inclusion of the library in school inspection reports.

The notion of comparing services with identified leaders in the field and with the best of similarly-situated services is an established part of quality methodology. It is used in "best value" assessment and advocated by the DfES and Ofsted as a key part of the "improving performance through self-evaluation" approach. Ofsted is building up a case bank of examples of best-practice schools, concerned with general school

self-evaluation, to support this comparison. It is planned that these examples will include some consideration of best-practice libraries (Jarrett, 2005). However, it would seem useful to have a bank of examples specifically to offer models in respect of *IYL* self-evaluation – which is in line with the original intention of MLA, the Council for Museums, Libraries and Archives.

Such a case bank does exist for Scottish school libraries, provided as part of the support resource for using *Libraries Supporting Learners* (SLAINTE). Markless and Streatfield (2005) have offered a caveat concerning the use of best-practice case banks. While models of excellence may be inspirational, for a number of librarians who feel that their situation is substantially less advanced or that their relationship with school is a barrier to development needs, such best-practice cases may simply seem unrealistic and consequently be demotivating. They argue that librarians need to start from understanding the teaching and learning nature of their school, in order to pitch their approaches and their expectations meaningfully within this context. (For a full explanation of the "four phases of school development" model, see Streatfield and Markless (1994).)

Conclusions from the literature review for designing the *IYL* research project

The body of studies considered reflects the growing interest in setting out clearly the importance of school libraries for teaching, pupil learning and attainment. The *IYL* research project is not itself concerned with establishing library impact on pupil learning. Rather, it is focused on how effectively librarians and schools are using the library self-evaluation process towards enhancing pupil learning and towards better collaborative whole-school understanding of the library's role in this enhancement. It is therefore relevant to consider the impact studies for how they demonstrate that:

- school libraries do make a difference;
- a variety of ways of evaluating provision is needed to make this difference obvious, including analysis of a range of stakeholders' "stories"; and
- this transparency is needed in order to make policy makers, educators and pupils themselves recognise how libraries and the role of the librarian make a positive difference to their teaching and learning.

While this may seem self-evident to many librarians, these are points that Markless and Streatfield have consistently been making for more than a decade and still find library "invisibility" a key problem. This research project therefore aims to identify how far use of *IYL* is actually making library value visible – or the findings of self-evaluation being constructively used to inform school development planning and pedagogical initiatives.

The concerns that inspection of English schools has highlighted and the Chief Inspector's specific recommendation of *IYL* to schools make it relevant for this research to examine:

- How schools perceive the role of library self-evaluation within their overall self-evaluation agenda.
- The extent to which school inspection refers to *IYL* as an evidence base.

- How often, and in what capacity, the library is specifically included in inspection reporting.
- How far Ofsted identifies improvement in the communication and integration of the library into corporate strategy. It is recognised, however, that real impact of *IYL* may take longer than the period of this research project to become evident in Ofsted reporting.

Provision of case examples of excellence for inspiration and benchmarking is an established approach, recognised by education and local government. Such a database, specific to school library self-evaluation, already exists in the Scottish context. This research project therefore aims to set up case banks for English school libraries that will provide examples of excellence, but will also offer a range of examples modelling good practice related to immediate school contexts. This dual approach is intended to give librarians and educators pragmatic and attainable models, as well as best practice to aspire to.

Issues raised by the pre-research scooping discussions and interviews used to shape the research project and inform survey design

A number of key issues has been identified at the preliminary stage, which the first, mapping, phase of the research project seeks to explore and understand for their implications on effective use of the self-evaluation process.

Availability of documentation

The printed documentation (*IYL for Primary Schools, IYL for Secondary Schools* and the *IYL Support Booklet*) was made available for distribution via school library services to schools in their area. The initial print-run was based on demand estimated by the school library services. Take-up to Spring 2005 exceeded the original print-run, necessitating reprints of both the primary and secondary editions of *IYL*. However, DfES deemed it too expensive to reprint the *Support Booklet*. Arguably this should not present a problem to users as all the documentation is also available online for download. However, some librarians have observed (conversations at SLA Conference, 2005) that they found the downloaded version awkward to print and that the benefit of colour coding is lost where school libraries only have black-and-white printers, presenting an unfortunate barrier to encouraging use. It also appears that despite the documentation being first released in July 2004, distribution to schools has been varied, some receiving it at the start of the 2004/2005 academic year but others still waiting as at Easter 2005 (conversations, April-June 2005). Several factors are hypothesised here: that some school library services are preferring to ensure that all their own staff have been trained in self-evaluation before *IYL* is rolled out to schools; that some schools fall into areas not served by a school library service; or that schools may not be deemed part of the school library service catchment and therefore have not been counted into the distribution. This latter may particularly affect independent schools – some interviewees in this position reported that they either had not received the documentation or had had to request it specifically (conversations, April-June 2005).

Availability of training and development of expertise

A New Relationship with Schools: The Next Steps (DfES, 2004c, p. 8) states that training for schools in effective self-evaluation is to be available before September 2005. There

is no indication of which staff are to be trained, but use of *IYL* is not a requirement within the overall self-evaluation and, apart from certain specific cases; headteachers can choose the school's evaluation focus. It seems likely that the priority for training will not therefore fall on librarians, although it is possible that in primary schools the teacher or teaching assistant responsible for library resources may benefit. The focus of the 2005-2006 session Ofsted English subject conferences for teachers and educators was on effective library use within the curriculum and Lemaire (2005a) provided a short overview of library self-evaluation, primarily raising awareness rather than delivering practical training in the *IYL* toolkit use. The importance of integration of library self-evaluation into the wider school self-evaluation was reiterated by Jarrett (2005), but he noted that the Ofsted subject focus would now be shifting away from school libraries.

Observation of workshop sessions for librarians on self-evaluation and feedback from training providers have suggested that many librarians do not as yet feel confident about how to tackle *IYL* or how best to build on the self-evaluation in developing either their own role or the role of the library in school. Markless has noted that her experience of delegates is that they frequently feel overwhelmed by the documentation and the process. They need reassurance that they can – should, in fact – be selective about what to evaluate and confidence to pick out appropriate evidence from the toolkit suggestions (Markless and Streatfield, 2005). Lemaire (2005b), chief executive of the School Library Association (SLA), has observed that the majority of delegates to SLA workshops are secondary school librarians who have either not yet started but want to, or have just started to use *IYL*. She has commented that SLA courses have tended to be targeted to areas where there is no school library service to offer training to schools, a point which reflects the concern expressed (above) about lack of support where schools have no such service.

There are likely to be a number of school librarians who have no access to training for a variety of reasons and who see themselves as lacking the time and confidence to use *IYL*. One from the south-east who decided not to start the process commented: "I thought about doing it but it is so complicated-looking – [it] expects such a lot of evidence which I don't have". It seems probable that this difficulty may be exacerbated in schools where there is no qualified librarian (conversations with librarians, various dates, 2005). For these librarians, qualified or unqualified, it would seem particularly important to find ways of providing training and support.

On a positive note, Lemaire (2005b) recounts the experience of an East London librarian who undertook *IYL*, and as a result found that she was perceived as being the self-evaluation "guru" in her school in a position to support teaching colleagues. This enhanced her status and put the library at the centre of overall school self-evaluation.

Reasons for using IYL

Those school librarians who had started to use *IYL* by June 2005 had chosen to do self-evaluation or had selected their particular Key Question and Strand for a number of different reasons.

Two librarians had chosen to do IYL because they saw it as a way of bonding the library into the school's overall development strategy. One commented that she felt teachers to be "friendly" towards the library, but that nevertheless the library remained "a separate animal". By showing active participation in the school focus on

self-evaluation, they hoped to get the library written into the school's development strategy at a point when the school was setting up its structures in response to *A New Relationship with Schools: Improving Performance*. This states "schools should have a single integrated development plan" (DfES, 2004d, p. 5).

Another had chosen her Key Question because it conformed to a key priority identified by school management for the year. Williams and Wavell (2001, p. 131), summarising the challenges of measuring the impact of school libraries, drew attention to the need for evaluation to reflect the school's stated curricular goals and how the school approaches achieving these goals. This librarian was already an experienced evaluator of her library. She felt that using *IYL* tactically to capitalise on school management's agenda demonstrated that the library was part of that agenda, and did not prevent her from continuing less formally with evaluation of her own library concerns also.

A number of librarians consulted indicated that they had selected a given Key Question or Strand because this was an area the school library was good at or because it represented something for which they already had much of the evidence. These librarians hoped that this would make the process easier to do as a trial starter – and it would be a positive way of reinforcing library input to senior management. Rhona Arthur of the Scottish Library and Information Council advocated this as a good way into using self-evaluation, at a workshop session on Scottish school library self-evaluation at the School Library Association Conference 2005. She argued that generally it was helpful for libraries to start from a position of strength. Where libraries could demonstrate success they tended to influence senior management more. She also reinforced the importance of seeking out the school's agenda (Arthur, 2005).

At this preliminary sounding-out of motivations, a number of librarians identified the potential of the toolkits in "finding a way into working with teachers". One summed this up by hoping that a recognised self-evaluation process would help to put systems in place for better liaison and expectations of collaboration between herself and teachers. (conversations with librarians, various dates, 2005; SLA Conference, 2005). These hoped-for outcomes fit well with Todd's (2005) statement, based on his Ohio research, that:

> [...] when teachers and librarians work together principals see firsthand evidence of your value. When teachers see that you are making a difference in student learning, they become your biggest advocates.

Finally, some had chosen to use a given Key Question because they saw this as an area their library was weak on. As one librarian planning on using 4c ("how well accommodation allows the curriculum to be taught") put it:

> I have been battling to get the resources – I see this as a way of presenting the case to [the senior management team] with all the evidence in front of them (SLA Conference, 2005).

Reasons for not using IYL

For some librarians, the *IYL* framework seems "complicated" and "expects such a lot of evidence which I don't have". One commented:

> I have had a go – but it took a huge amount of effort and time – I don't have this time as a solo librarian, unsupported, and it has been hard to use this evidence to make a difference to the [senior management team] – so I am unlikely to do it again (SLA Conference, 2005).

This ties in with Lemaire's and Markless's observations that librarians starting out on the process sometimes assume that they have to take on all of *IYL* or do all of a Key Question, rather than selecting one Strand or element. The self-evaluation process is intended as iterative and over time methodologies and evidence will be built up, so use is likely to become easier – but only if librarians are not discouraged at the outset. The comment does draw attention to a practical problem that a significant number of librarians face, that of rarely getting beyond basic stock management and promotion and policing of the space, related to a very narrow conception by school management (and occasionally by librarians) of what the role of library and librarian is. This narrow "resource manager" conception is at odds with the profession's understanding of the librarian's role as an active participant in school policy and strategy development in a number of cross-curricular areas and wider development of readers; also with recognition of the librarian as a partner in collaborative teaching of information literacy in and out of the library venue. The narrow view may, however, be reinforced in some instances by school management decisions taken in defining the librarian's role as support staff within current workforce remodelling.

The second point arising from the comment above concerns the benefit of engaging in evaluation. The speaker felt that it was difficult to use the evidence to make a difference. This is fundamental. Evaluation is a tool to inform development towards the library purpose of enabling learners rather than an end in itself. Another librarian observed that she "prefer[s] to do [evaluation] the way I can cope and to target effort onto those staff where I can make a difference" (SLA Conference, 2005).

For others, the problem seems to be that their school is academically highly successful and sees no need to develop its mode of resource use or pupil learning experience. The librarian is then faced with the challenge of identifying how the library is adding value to pupil learning in ways that are also valued by management:

> I moved out of teaching, which was overburdened by assessment and paperwork, into librarianship to be able to continue to work directly with children in school but out of the classroom. I don't want the library to be tarred by the assessment or measuring brush (conversations, April-June 2005).

This was the view of one (qualified) librarian expressed to the author in conversation. As long ago as 1983 Schön (1983 in Haines, 2004) identified:

> [. . .] the capacity to reflect on action so as to engage in a process of continuous learning is a defining characteristic of professional practice.

It is essential for the acceptance by teachers of the professional status of the school librarian, in an educational environment used to evidence-based practice, that librarians are seen to be reflective practitioners. The existing body of research evidence and the performance climate cannot accommodate rejection of self-evaluation. *IYL*, as a toolkit of processes, examples and scenarios, may make it easier for those who are uncomfortable with this to manage it.

However, there is a second, more challenging point at stake here, namely the tension between engaging in evaluation in order to prove library impact in terms of pupil achievement and curricular success (see Lance *et al.*, 2000) and developing the library as a "space" for pupils distinct from the classroom climate. Heeks (2000, p. 122) draws attention to this evaluation dilemma:

> [. . .] if you try to measure [the library] in terms of improvement of academic performance, it becomes just another area where children are assessed.

The concern is relevant as librarians seek to foster greater reading for pleasure and encourage pupils to recognise the library as supporting and inclusive to all as individuals. A significant amount of the recommended evidence and methods for obtaining this within the *IYL* frameworks involves pupil observation, assessment of work or interviews with pupils (e.g. see the guidance materials for Key Questions 1, 2, 4 and 5 provided within *Improve Your Library: A Self-evaluation Process for Secondary School Libraries and Learning Resource Centres: Support Booklet* (DfES, 2004e)) and there is a risk that how evidence is captured will alienate pupils.

Conclusions from the pre-research scooping discussions for shaping of the research project methodology and design of survey instruments

Informal feedback has shown considerable variation in librarian experience of *IYL* access, support and use during the year following publication. This feedback has been dominated by comments relating to secondary school library use. It is also evident that librarians and schools are at different starting points, in terms of perceptions of library and librarian role, expectations of integration into management decision making and collaborative working. Understanding these issues and constraints will be useful for training and support providers. It will also be useful for identifying areas where practice examples can be helpful to librarians embarking on *IYL* Key Questions. Therefore a clear mapping of these factors is needed as a base line for investigation of any impact.

A key point that has emerged at this stage is the range of motivations and expectations librarians have for use of self-evaluation. It is important therefore for this research project to examine these and to relate them to the expectations for self-evaluation of school management. The research must investigate tactics of use by school librarians and the extent to which these deliver on the intended expectations. In doing so, it will look at any patterns which can be established for how Key Questions contribute to better whole-school awareness of the role and contribution of both library and librarian to teaching and pupil learning, and how this is translated into development planning. If the complaint that *IYL* is too time-consuming and difficult to use to make a difference is found to be significant, this has implications for support from school library services and training providers and for positive sharing of experiences via the case banks and other practitioner networks. It could have implications for viability of the self-evaluation process itself. Understanding the relative impact of targeting specific school agendas or exploiting library strengths or weaknesses will be helpful for librarians selecting their approach. Exploring teacher and pupil feelings about their involvement with the library self-evaluation process and how this affects their perceptions of the library will be important for how librarians handle stakeholder involvement. The research project must therefore build on the broad mapping exercise by exploring in depth individual use of *IYL* within a of range schools.

Methodology

This research is structured to run in two phases.

Phase 1 is scheduled to run from June-December 2005 and is concerned with mapping patterns and experiences of use of the toolkits by schools throughout England.

Goals are to:

- Establish a national picture of the nature of use of *IYL*.
- Identify issues experienced by schools and school librarians in using *IYL*.
- Explore the motivations of schools and school librarians in using *IYL*.
- Identify the nature of support and involvement of school library services for *IYL* – to local authority/independent/school library service-subscriber schools.
- Identify training and other support needs noted by schools in respect of using *IYL*.

The mapping phase of the research is being conducted using a series of questionnaires targeted to librarians and to school library services. The questionnaire design has been informed by discussions with librarians concerning their experiences of *IYL* up to June 2005 and informal interviews with practitioners involved in the delivery of support and training to librarians. The questionnaire to librarians in schools is being disseminated via a number of discussion lists, e.g. sln (School Libraries Network), and membership mailings via organisations concerned with school libraries. It is also being sent to all the schools involved in the *IYL* toolkit development pilot. To reach primary schools, where the individual responsible for library provision is likely to be a teacher, the questionnaire is being sent to a sample of schools across England. The selection is constructed by statistical sampling from alphabetic school listings, including maintained, aided and independent schools, within each local education authority. This survey investigation is supplemented by open forum consultation sessions aimed at school librarians and practitioners involved with *IYL*.

The survey of school library services is being conducted via a questionnaire sent directly to each service in England. Informal comment has shown that although *IYL* has been designed for use in English schools, there has been some interest in it in Wales. Accordingly, school library services in Wales are being included in the consultation.

A preliminary report is scheduled for spring 2006 and findings will be discussed at the School Libraries Group Annual Conference 2006.

Phase 2 of the research is planned for 2006-2007 and is concerned with identifying the impact of the self-evaluation process on school libraries and their engagement with school priorities and teaching and learning. It is intended that a varied sample of schools will be followed in detail over the two year period. This sample will be constructed by selection from school types, location and range of library experiences identified in phase one. The schools will be tracked to explore how *IYL* is being used and the extent to which this use is having an impact on library and librarian involvement within the curriculum and contribution to pupil reading and learning.

Goals are to:

(1) Identify how far use of *IYL* is integrated into whole school self-evaluation and development planning.

(2) Explore the extent to which Ofsted assessment takes into account use of *IYL* frameworks during the inspection process and in subsequent reporting.

(3) Evaluate what difference use of IYL makes for school libraries and how IYL is being exploited as an advocacy tool in schools.

(4) Build up and maintain databases of good practice examples accessible by:
- Key Question and Strand;
- self-assessed grading level;
- nature of school; and
- impact achieved.

This evaluative phase of the research will be conducted through a combination of observation visits, interviews, survey documents, research participant "stories" and open group consultation sessions. Stakeholders will include librarians, management and teaching staff in the sample schools, pupils and Ofsted representatives. Comparison will be made with the experiences of Scottish schools using *Libraries Supporting Learners* (SLIC/HMIe, 2005) and its predecessor framework, *Taking a Closer Look at the School Library Resource Centre* (SCCC/SLIC, 1999) to examine what may be learned from their experiences of library self-evaluation for developing the application of *IYL*. Examination of the Scottish libraries best-practice case bank (SLAINTE, n.d.) and the planned Ofsted schools best-practice case bank will be used to inform design of the *IYL* case banks.

It is intended that there will be a phase two interim report and a final report at conclusion of the research.

Where next? A phase three?

An issue that has emerged even at this early stage is the need for easy access to helpful examples, to build confidence and offer useful benchmarking. Provision of case banks is a feature of approaches used by Ofsted and MLA, to inform practitioners and policy-makers. Accordingly, development of appropriate case banks has been identified as a key part of the planned outcomes for this research. However, the value of this provision, including the "best-practice" and the "development phase" cases, will depend entirely on how end-users judge their usefulness. It is not anticipated that the case banks will be sufficiently established within the present research project parameters to explore such patterns of use. It would therefore be helpful to develop a third phase to this research, which would look at who uses the case banks, motivation for use and how this is then exploited. Questions arise here for use of case databases in wider advocacy work and in strategic policy development at government level. It would be helpful also to understand the relative benefits to users of good-practice examples available as models at the point of undertaking initiatives, and case examples that allow users to grow throughout a project. Analysis of preference for the "best-practice" and "development phase" models would be valuable in identifying case bank design for future projects and in refining the case banks arising from this research.

Another aspect that has become apparent at this preliminary stage is the potential of *IYL* as a model for school library evaluation outside of England. The interest from school library services in Wales has already been commented on above and included within phase one and two parameters. However, Lemaire (2005a, b) notes interest from the Hungarian School Library Association in developing a similar self-evaluation framework, and has reported enthusiasm from school librarians in Moscow to learn about *IYL*. It is hoped that the findings from the present research will be of use for any

such development, but it is beyond the present scope to investigate any impact of these frameworks. However, a further research project might compare experiences and impact of parallel processes in other countries, testing the links between educational and cultural contexts and the benefits of self-evaluation in school libraries.

Conclusion

The *IYL* research project is timely for a number of reasons. It recognises that the status of school libraries in England in 2005 is extremely variable and that a significant number of schools still do not have a full understanding of how important a part their library and qualified librarian could play in pupil learning, development and achievement. It recognises that this variation, the problems of a common language for teachers and librarians surrounding these issues as identified by Williams and Wavell (2001) and differing experience levels of librarians mean that there are different starting points and motivations for librarians embarking on self-evaluation. Librarians and trainers have both raised issues of access to *IYL* and training, and user confidence. The research will examine these to identify the level of take-up of *IYL* and barriers to this, and the nature of support needed from providers.

Development of the frameworks sought to set the library securely within the educational evaluation context and Ofsted has recommended their use. The research will show how far school management picks this up in seeking to integrate the library into wider evaluation. It will also address how Ofsted carries through its recommendation to schools by actively looking for application of *IYL* at inspection or in identifying best-practice schools.

The findings of the pre-research consultations have indicated some of the concerns and the range of expectations librarians have from self-evaluation. These have been useful in informing the construction of the surveys to librarians, schools and school library services that form part of the first phase of the research. The findings confirm the need for some of the key areas of investigation proposed. In particular, where librarians are using *IYL* to bond the library into whole school development planning or to tie it to a key evaluation focus chosen by management, it will be important to discover how far this approach is in fact embedding the library into the wider planning. Equally it will be important to capture the extent to which attitudes of senior management and teachers are influenced by library engagement in this self-evaluation process. Likewise it will be valuable to discover whether the toolkits enable librarians to present development cases more effectively or to build better partnerships with teaching staff than through earlier methods. Librarians consulted have chosen to exploit Key Questions in areas of strength and of weakness. The research will identify what patterns of use are emerging, whether particular Key Questions or Strands seem to achieve greater impact and whether in fact there is a difference in development success in choosing to focus on library strengths or weaknesses. The research will also seek to identify if there are particular approaches that suit schools in a given phase of development.

Through the case banks it is intended that librarians, educators and policy-makers will have access to examples that can provide a form of benchmarking, offer models for exploiting self-evaluation and be of use in advocacy. A limitation may be obtaining sufficient examples during the research period to populate the case banks with a good range of examples. However, it is intended that population will continue to be built on

and it may be possible to refine the banks, through future investigation into how the case examples are used.

It is evident also that others are looking at the *IYL* model to inform developments in other countries. It is hoped that this research will be able to contribute to better understanding of the issues involved in using such self-evaluation processes and the nature of the difference they achieve for school libraries. Similar future investigation into the impact of any potential other self-evaluation toolkits would also enhance this research.

Finally, although this research in no way seeks to demonstrate the impact of school libraries in England on pupil learning, it is hoped that better understanding of the impact of self-evaluation on library engagement with school priorities and teaching and learning might contribute helpfully towards any such future studies.

Note

1. The School Libraries Forum is a body drawing together representation from a range of organisations, agencies and government departments concerned with school libraries, education, publishing and reader development, with a view to enhancing every aspect of children's education through the provision of effective school libraries. Chairmanship is shared by CILIP and the School Library Association

References

Arthur, R. (2005), "Libraries supporting learners", paper presented at workshop session, School Library Association Annual Conference, Knowledge Is Power: Creating the Information-Literate School, University of Surrey, Guildford, 17-19 June.

Bell, D. (2005a), *Ofsted Subject Reports 2003/04 – English in Secondary Schools*, HMI 2429, Ofsted, London, available at: www.ofsted.gov.uk/publications.annualreport0304/subject_reports/secondary/english.htm (accessed 1 August 2005).

Bell, D. (2005b), *Ofsted Subject Reports 2003/04 – English in Primary Schools*, HMI 2413, Ofsted, London, available at: www.ofsted.gov.uk/publications.annualreport0304/subject_reports/primary/english.htm (accessed 1 August 2005).

Brophy, P. (2005), "Researching practice, practising research", paper presented at Umbrella Conference, University of Manchester, Manchester, 30 June-2 July.

Chartered Institute for Library and Information Professionals (CILIP) (2002), "Start with the child", *Report of the CILIP Working Group on Library Provision for Children and Young People*, CILIP, London.

Chartered Institute for Library and Information Professionals (CILIP) (2004), *Guidelines for Secondary School Libraries*, Facet, London.

Department for Education and Skills (DfES) (2004a), *Improve Your Library: A Self-evaluation Process for Primary Schools*, DfES, London.

Department for Education and Skills (DfES) (2004b), *Improve Your Library: A Self-evaluation Process for Secondary School Libraries and Learning Resource Centres*, DfES, London.

Department for Education and Skills (DfES) (2004c), *A New Relationship with Schools: The Next Steps*, DfES, London.

Department for Education and Skills (DfES) (2004d), *A New Relationship with Schools: Improving Performance through School Self-evaluation*, DfES, London.

Department for Education and Skills (DfES) (2004e), *Improve Your Library: A Self-evaluation Process for Secondary School Libraries and Learning Resource Centres: Support Booklet*,

DfES, London, available at: www.teacdhernet.gov.uk/schoollibraries and www.informat.org/schoolibraries/index/htm (accessed 1 August 2005).

Haines, M. (2004), presentation given at the UCRG/CoFHE Conference, 5-8 April.

Heeks, P. (2000), "Assessing services", in Elkin, J. and Kinnell, M. (Eds), *A Place for Children: Public Libraries as a Major Force in Children's Reading*, British Library Research and Innovation Report No. 117, Library Association Publishing, London.

Jarrett, P. (2004), briefing comments made to School Libraries Forum, CILIP, 18 November.

Jarrett, P. (2005), "Good school libraries: making a difference to pupils' learning", paper presented at the Ofsted Subject Dissemination Conferences for English, London and Sheffield, 17-18 October.

Lance, K., Curry, R., Marcia, J. and Hamilton-Pennell, C. (2000), *Measuring up to Standards: The Impact of School Library Programs and Information Literacy in Pennsylvania Schools*, Pennsylvania Citizens for Better Libraries, Greensburg, PA, available at: www.statelibrary.state.pa.us/libraries/lib/libraries/measuring up.pdf (accessed 1 August 2005).

Lemaire, K. (2005a), "Finding the evidence: school libraries and self-evaulation", paper presented at the Ofsted Subject Dissemination Conferences for English, London and Sheffield, 17-18 October.

Lemaire, K. (2005b), conversation with Kathy Lemaire, CEO of School Library Association, SLA, Swindon.

Markless, S. and Streatfield, D. (2005), conversation with Sharon Markless and David Streatfield, King's College London.

Office for Standards in Education (Ofsted) (2001), "Inspecting school libraries and learning resource centres", *Guidance for Inspectors*, Appendix D, Ofsted Update No. 36, August, available at: www.ofsted.gov.uk/publications/docs/update36/update36-08.htm (accessed 1 August 2005).

Research Foundation (2004), *School Libraries Work!*, Scholastic, available at: www.scholastic.com/librarypublishing (accessed 1 August 2005).

Scottish Consultative Council on the Curriculum/Scottish Library and Information Council (SCCC/SLIC) (1999), *Taking a Closer Look at the School Library Resource Centre: Self-evaluation Using Performance Indicators*, SCCC/SLIC, Hamilton, available at: www.slainte.org.uk/files/pdf/slic/schoolpisx.pdf (accessed 1 August 2005).

Scottish Library and Information Council/Her Majesty's Inspectorate for Education (SLIC/HMIe) (2005), *Libraries Supporting Learners*, How Good Is Our School? Series, No. 19, Livingstone, available at: www.hmie.gov.uk/documents/publication/hgioslsl.pdf (accessed 1 August 2005).

SLA Conference (2005), conversations with librarians at School Library Association Annual Conference, Knowledge Is Power: Creating the Information-Literate School, University of Surrey, Guildford, 17-19 June.

SLAINTE: Information & Libraries Scotland (n.d.), "Case studies", available at: www.slainte.org.uk/learners/casestudies.htm (accessed 1 August 2005).

Streatfield, D. and Markless, S. (1994), *Invisible Learning? The Contribution of School Libraries to Teaching and Learning*, British Library Research and Development, Library and Information Research Report No. 98, British Library R&D, London.

Streatfield, D. and Markless, S. (2005), "Putting school libraries back into the picture", *Library + Information Update*, Vol. 4 No. 6, pp. 36-8.

Todd, R. (2005), keynote speech at School Library Association Annual Conference, Knowledge Is Power: Creating the Information-Literate School, University of Surrey, Guildford, 17-19 June.

Todd, R. and Kuhlthau, C. (2004), *Student Learning through Ohio School Libraries: The Ohio Research Study*, Ohio Educational Media Association (OEMA), Columbus, OH.

Williams, D. and Wavell, C. (2001), *Impact of the School Library Resource Centre on Learning*, Library and Information Commission Research Report No. 112, Robert Gordon University School of Information and Media, Aberdeen.

Corresponding author
Lucy Gildersleeves can be contacted at: l.gildersleeves@ucl.ac.uk

The current issue and full text archive of this journal is available at
www.emeraldinsight.com/0001-253X.htm

The rendering of humanities information in a digital context

Current trends and future developments

Susan Hockey
School of Library, Archive and Information Studies, University College London, London, UK

Received 7 June 2005
Revised 8 August 2005
Accepted 11 August 2005

Abstract

Purpose – The purpose of this paper is to consider how digital resources might best be created and how the digital medium might best be exploited to serve the needs of research and teaching in the humanities.

Design/methodology/approach – The history of text-based humanities computing is traced to illustrate the complexity of humanities sources and the requirements for their representation in digital form.

Findings – The paper finds that digital information is completely different from any other kind of information, being dynamic, extensible and manipulable in many different ways. It is not static and it can be manipulated in many different ways. The Extensible Markup Language (XML) has been widely adopted in the humanities particularly for full text sources encoded according to the Text Encoding Initiative application. XML permits the display, analysis and manipulation of fine detail within a document. It also facilitates the chunking of information and the integration of data and metadata leading to new forms of publication that exploit the digital medium to the full.

Research limitations/implications – The research environment proposed in this paper challenges current modes of working with a more innovative approach to the nature and functionality of digital representations of humanities sources.

Originality/value – The paper looks at the development of humanities computing and should be helpful for anyone embarking on or managing a humanities digital project.

Keywords Information management, Resources, Extensible Markup Language, Research methods, Internet

Paper type General review

Introduction

Digital resources are now part and parcel of every day work in the humanities and the virtual world of the internet is becoming the focal point for research and teaching. But exactly what are these digital resources in the humanities? How can they be designed and developed to best serve the needs of the humanities? What are the challenges that face information professionals and scholars as the shift to this new way of working continues to accelerate? This paper examines some of these issues from the perspective of a designer and user of resources. It looks at the current landscape and speculates on some future developments that could lead to completely new ways of working and new forms of publication.

The nature of digital information

The arguments made here are built on the understanding that digital information is completely different from any other format. First, the digital medium can be used to

Aslib Proceedings: New Information Perspectives
Vol. 58 No. 1/2, 2006
pp. 89-101

0001-253X
DOI 10.1108/00012530610648699

represent all other forms of information including text, numbers, images, audio and video. A digital resource may consist of a representation of only one of these but more often, as evidenced by many web sites, it now represents several of them. Each has its own technical structure requiring different processing software. In a multimedia resource, the components are glued together usually by some text which indicates links between the components and access to those links. Tools for manipulating text are therefore fundamental for any kind of resource.

Second, and more importantly for the arguments to be made here, digital information exhibits very different characteristics from those found in physical objects. It is not static. It can be revised and updated almost continuously. It is processable in many more ways than physical objects that users can view, read or listen to, but not easily manipulate in any other way. Processability includes searching where the benefits of electronic information retrieval are very obvious. It also includes analysis of various kinds such as the counting of features within the resource, mathematical operations on numeric data, and zooming and enhancing of images. On a different level, processability includes the presentation of the resource in different views and displays. These displays may present the entire resource or, more often, subsets of it selected according to criteria specified by the user. Subsets may then be aggregated with information taken from different resources to create new resources dynamically.

Inherent to all of this is the flexibility of digital information. With appropriately designed resources it ought to be possible to create a research environment where resources of many different kinds can be mined and analysed and the results of these analyses embedded in new kinds of publications that take full advantage of the potential afforded by the digital medium. We can see the beginnings of this in the world wide web where links between pieces of information in the form of web pages can be made with ease – but only by the creator of the page. The idea of hypertext and linking between pieces of information predates the web by over 40 years. Bush's (1945) proposed Memex machine was intended to model the way that humans follow links between pieces of information. The term "hypertext" was coined by Nelson in the 1960s as part of the plan for Xanadu, a projected docuverse which is rather similar, but more sophisticated than what was finally implemented as the world wide web in the late 1980s (Nelson, 1987). The explosion in use of the web is testament to its simplicity and ease of use, but, as we shall see later, the technology underlying the web of today inhibits some of the potential of the digital medium.

The nature of humanities research

In the humanities much research concerns the study of primary sources. These sources might be texts in manuscript, printed or spoken form, collections of works of art, historical records, film and other forms of performance. In fact almost anything can be an object of study, but the one thing that most humanities sources have in common is complexity. Textual material may be literary sources in the form of prose, verse, and drama, or historical documents in the form of charters, newspapers, letters, or diaries, or spoken texts (conversations, political speeches, radio broadcasts), or dictionaries (both historical and modern). Readings of parts of manuscripts may be unclear, sections might be missing completely or other manuscripts of the same text might contain different versions. Spelling is not consistent, especially in early texts. Spoken

language is often ungrammatical and incomplete. Names can appear in different forms and it is often unclear whether they refer to the same person. Dates are another problematic feature. For example, how can a date such as "possibly last year" or "in the reign of Nero" be put into chronological order? The material may be in several languages and writing systems. Humanities sources also have a very long life cycle, hundreds of years in some cases. Provenance, both of the original physical object and of any digital representation of it, is therefore also important. Furthermore, the same source can be studied by different researchers for different purposes, for example, linguistic or historical or sociological investigations.

Research output takes different forms. It could be the publication of some sources in the form of an edition that also contains notes, indexes and other tools to help the reader understand the sources. It could be an article which discusses and analyses the sources, promoting the scholar's view of what the source and features within it might represent. Traditionally it is a monograph that perhaps can be considered as a very long article divided into sections (chapters). In the non-digital world methodology is rarely an issue as it consists of locating sources, reading or viewing them, taking notes and writing. Interpretation is the key component of the research output, the new knowledge that is created by the research. The notes in a scholarly edition and the content of articles and monographs provide the scholar's interpretation of the source. These interpretations illuminate and explain features and items within the sources. They make connections between them giving the rationale for the connections. The interpretations and explanations may well then be challenged by other scholars working on the same material, who publish new and different views of the same source.

Early developments in humanities computing

Reference works in the form of word indexes, dictionaries and encyclopaedias are another form of research output in the humanities and it is here that the potential of the computer in assisting with research was first recognized. The use of digital resources in the humanities began in 1948 when Father Roberto Busa started work on his *Index Thomisticus* in Italy. He aimed to compile a concordance (alphabetical list of words with contextual information) of the works of St Thomas Aquinas and some related authors, approximately 11 million words in total. He sought help from IBM and began to create electronic representations of all his texts and to develop computer programs to create the index (Busa, 1980). The first volume of a series of printed concordances appeared in 1974 (Busa, 1974), but the digital text was re-used to create an electronic version ("*cum hypertextibus*") (Busa, 1992).

Busa's lead was followed by others, not only for the compilation of reference works, but also for other kinds of analysis which are based on alphabetical lists and counts of word forms. Gradually a nucleus of people interested in literary and linguistic computing began to get together to discuss common problems and future directions (Bessinger and Parrish, 1965; Wisbey, 1972; Aitken *et al.*, 1973; Jones and Churchhouse, 1976; Ager *et al.*, 1979; Hockey, 1980). This group consisted mostly of humanities scholars and computer programmers. It lacked input from the information professions and was consequently less aware of standards for describing and documenting information. Other scholars in the humanities, notably historians, also began to compile digital data sets (Denley and Hopkin, 1987; Denley *et al.*, 1989; Mawdsley *et al.*,

1990; Greenstein, 1994). Until the early 1990s, the primary purpose of most of this work was analysis and manipulation. The digital resource was created according to the needs of the scholar who directed the project. It was analysed and manipulated by standard programs such as OCP for text analysis or SPSS or SAS for numeric data. The results of the analysis were then published by conventional needs and the digital resource was often then abandoned somewhere in the university computing centre.

Methodology came to the fore. Research was needed on how to carry out research using this new and completely different medium. These early practitioners began to identify and tackle problems in representing and analysing their material. Can a multilingual manuscript in Greek, Arabic and Latin be represented digitally in such a way that the text in each language can be indexed separately and according to its own alphabetical order? How can the different forms of the same word (for example, go, going, gone, went) be listed under the headword go where they would be found in a dictionary? Father Busa recognized the need for this process, which became known as lemmatization. He always maintained high standards for his work and provided manual assistance to his software in order to achieve it. In non-text-based disciplines attention turned to data structures (Greenstein, 1991). For example, what would be the most appropriate digital structure for representing an archaeological site with all its finds and contexts? In a conventional database system these data might be squeezed into a linear or multi-tabular format, but scholars were concerned that this would most likely lead to some loss of information. Data modelling for a relational database assumes knowledge of the relationship between the elements in the data, whereas the major aim of a historical research project is often to find those relationships. Putting the elements into a relational database thus imposes a level of interpretation at the data creation rather than at the analysis stage of the project.

Structured markup and the representation of humanities information

Research in what became known as humanities computing began to concentrate on representation issues. This was driven partly by the need to represent fine detail within complex sources but also by a growing realization that other people might want to use the same digital resource, possibly for a different purpose. The Standard Generalized Markup Language (SGML) (International Organization for Standards, 1986) was adopted as the basis for a major international project called the Text Encoding Initiative (TEI), which began at the end of 1987. In 1994 it published its *Guidelines for the Electronic Text Encoding and Interchange* (Sperberg-McQueen and Burnard, 1994). This was the first serious attempt to codify the features that humanities scholars might want to examine within a digital text. The TEI suggests methods for encoding these features but it is extremely flexible. Very few encoding features are mandatory and most of these are concerned with the metadata that describes the electronic text, rather than features within the text itself. SGML was chosen as the basis for the TEI encoding scheme because of its flexibility and machine independence. It appeared to provide the best solution for longevity and reusability of the digital text. The Extensible Markup Language (XML) derived from SGML and is built on the same principles (Yergeau *et al.*, 2004). The TEI project developed an XML version as soon as it was able (Sperberg-McQueen and Burnard, 2001). This encoding scheme has formed the basis of many projects in the humanities and is widely used today (Text Encoding Initiative, n.d.).

XML meets the needs of many application areas. An XML file is ASCII or Unicode and is thus completely machine and software independent. The markup consists of tags embedded in the document, identifying the features within the document. Technically XML is a methodology for designing markup schemes. The XML document designer chooses which features to encode and the relationships between them. Markup tags are nested within in each other. Theoretically anything can have markup associated with it from the level of a huge text down to a fragment of a letter. Any feature which is encoded can thus be picked up and manipulated by a processing program.

XML-based encoding facilitates many different kinds of processing. With the aid of stylesheets, the same text can be displayed on screen or printed in many different formats. Subsets of the text can be processed separately by the selection of certain XML-encoded features (elements), for example, abstracts in a collection of articles. The TEI devised a very flexible method of encoding the structural divisions within a text (chapters, sections, verses, acts, scenes, speeches, etc.), making it easy, for example, to extract or examine all the speeches by Ophelia in Hamlet. Within these sections a variety of more detailed features can be encoded. Using the attributes associated with XML elements, suggested readings for unclear texts can be given as part of the unclear encoding. A standard form of spelling can be associated with a variant one, making it possible to index all the variant ones together. Names can be encoded with attributes identifying the person or indicating that the name is that of a place or even that of a ship. Dates can also carry attributes which can be used to put them into some kind of chronological order. A method is even provided to encode the part of speech for each individual word, which means that the verb "lead" is not retrieved when the metal "lead" is required.

The advent of digital libraries: a shift of emphasis to delivery and publication

The arrival of the world wide web in the early 1990s changed perspectives in humanities computing as it did in many other disciplines. It made it possible to publish material in a completely new way. Attention thus shifted away from analysis and towards the delivery and publication of sources (Hockey, 2004). Not only was access much easier (for anyone who had access to a computer), but also the shape and functionality of a publication became very much more flexible. Material no longer had to fit on to the linear structure of printed pages and it was no longer restricted by the physical size of a book. Furthermore, the technology for digital imaging became more widely available. In text-based work in the humanities this meant that a picture of a manuscript could be delivered as well as a transcription. Multiple manuscripts of the same text could be provided with links between them or software tools to present them side by side. The possibilities for annotation were greatly enhanced. For example, in a printed documentary edition a short biography of an individual in the text would appear once as a footnote and then be cross-referenced by future mentions even when these occur in subsequent volumes. In the digital environment the biography is a piece of information referenced by a hypertextual link every time that the individual is mentioned.

The early 1990s also brought the introduction of a new term "digital library". The availability of large-scale funding for digital library research, especially through the

US National Science Foundation (Griffin, 1998), attracted large projects, but many of these were focussed on the delivery of very large amounts of fairly raw scientific data or journal articles, rather than the smaller amounts of fine detail in primary sources often required by the humanities. In the UK the ELib programme funded a larger number of smaller projects (Rusbridge, 1995), some of which had participation from bodies outside higher education. It was also at this time that libraries and archives began to get interested in the delivery of digital resources themselves rather than pointers to these resources in the form of catalogue records or finding aids. Publishers were also beginning to experiment with electronic journals, but almost all of these efforts were concentrated on the delivery of resources rather than on the analysis and manipulation of them. The digital environment began to be seen more as a communications medium than as a computational medium. Of course there are many advantages to this, convenience for the user being foremost, but most electronic journals are still at the stage of delivering what one might call electronic print.

Perhaps we can do better. Although journals are not the primary focus of this paper, two areas warrant discussion. The first is the linking of research sources to a journal article. A few experiments have been conducted in this area, for example in the journal *Internet Archaeology* (1996). In one method of use, the article merely references the relevant data source so that the user can bring it up in another window. In another the electronic article provides a link to the source so that users can access the data, repeat the analyses of the data and re-use that data for their own research purposes. The linking of data (whether it is text, numbers, images or other formats) to the analysis and interpretation provides the user with a greater insight into the research and greater possibilities for the creation of new knowledge based on his or her own interpretation of the source. The second would be simpler to implement if the journal article was appropriately encoded. When a quote is to be taken from an article, it should surely be possible for the user to highlight the quote in the text of the article in one window, then select just one function to insert the quote into the article he or she is writing and have the reference generated automatically for the bibliography. Other possibilities for electronic journals are outside the scope of this paper, but some re-thinking of practices and procedures and a recognition that print is better considered as a by-product of an electronic text, not the ultimate end-product, could lead to a more imaginative and thus more useful use of the medium.

The limitations of HTML

HTML is the first exposure to markup for many people, at least to markup that is more easily visible than that embedded in Microsoft Word. In comparison with richer markup schemes based on XML, HTML is rather a curious mixture of tags. They appear to be designed to facilitate display and to interact with the browser. HTML is a mixture of structural and presentational tags, but the structural ones do not work well for text that is not prose or tabular in format. In fact they are intended for presentation. A $<$p$>$ tag tells the browser to insert a blank line and then start on a new line. Lines of poetry are often represented in HTML as $<$blockquote$>$ which they are clearly not. Structural elements do not have any identifiers associated with them. We can find a $<$p$>$ but we do not know which one it is or where it occurs in the text. HTML tables are clunky to say the least and their use to control layout as well as to hold tabular

information is likely to confuse any but the most sophisticated software for processing tabular material.

We have seen that linking (making connections) is a major part of the research process in the humanities, but the linking functions in HTML are also rather simple. In most circumstances when the user clicks on a link a new page is opened, replacing the one containing the link. The new page can be opened in another window, enabling the user to view both pages side by side. With a link to a named anchor, the new page opens with the anchor position at the top of the screen, but the entire page is still loaded. Links can only be inserted in an HTML document by the creator of the document. A user can create a link to the document from his or her own page, but cannot directly address part of another web page.

XML provides more sophisticated linking which mirrors more closely what researchers might want to do (DeRose *et al.*, 2001) More often than not, it is only portions of a document that are needed. XML links make use of the XML-based structures within a document so that, for example, it is possible to link to the third chapter in a document that contains many chapters. Only the third chapter is then accessed, not the entire document as in an HTML named anchor link. XML links can also access a structural unit which begins with or contains certain text and can be used with software that also identifies other features such as author or title of the entire work that contains the specific structural unit. The XML links also identify a chunk of information not a position in the text as HTML links do. The user can then take this chunk and process it in some way. In effect these kinds of links facilitate the chunking of information and help to weed out unwanted material. A collection of XML links can be stored externally and used to mark out different pathways through a collection of material without changing any of the markup within the documents. The TEI developed a syntax for extended pointers (Sperberg-McQueen and Burnard, 1994, Ch. 14) which uses these mechanisms, reflecting the strong emphasis on linking and connections in the humanities. These kinds of links give us an insight into a digital environment of the future where new kinds of publications can be created by charting pathways through digital objects stored in various locations and adding to the pathways the scholar's own interpretation of the objects and the reasoning behind the chosen pathways.

Metadata issues

The relationship between metadata and the resource which it describes, also begins to be called into question. Most catalogue records for digital resources now provide a link to the resource, but only at the level of HTML links. The user must go to the resource to find out how to use it. In most current and conventional systems metadata and resource are thus separate, but as we move into an environment where content is in digital form, is there any reason why the metadata and the resource are held in different formats and systems with only superficial links between them? It would make more sense to integrate the two. In an XML-based environment the same processing tools could be used for the metadata and for the content. Portions of the content could be derived, accompanied by the metadata which are relevant for them.

The distinction between metadata and data is already blurred in some kinds of humanities resources, for example, manuscript descriptions which provide detailed information about the content and palaeography of the manuscript. These descriptions

may themselves carry metadata about their creation and maintenance. Two features are important for metadata for humanities digital resources. The first is provenance, giving details about how the resource was originally compiled and created and its lifecycle since then. The second concerns the technical structure of the resource which should provide enough information for the user to determine whether software chosen for its functionality will actually work with the resource.

The TEI is believed to be the first serious attempt to provide metadata in the same format as the data. Every TEI document must contain a header which consists of a number of XML elements (Sperberg-McQueen and Burnard, 1994, Ch. 5; Giordano, 1994). Few elements are mandatory, but in the file description section, the bibliographic information for the electronic text and for the source from which it was transcribed or derived must also be given. Encoding and transcription policies can be specified in the encoding description section. For example, the user can specify whether errors in the original text have been corrected, whether hyphens at the ends of lines have been retained and whether spelling has been standardized. Among other things, the profile description section provides a way of giving demographic or other information about the participants in a conversation which has been transcribed. Finally, the revision history records changes made to the electronic text, either by the original creator or by subsequent users who have added markup to the text.

The LEADERS Project: integrating data and metadata

The Encoded Archival Description (EAD) is an XML-based method of encoding archival finding aids (Encoded Archival Description, 2002). It is now widely used throughout the world but in most cases in a similar manner to a catalogue. It helps a user locate sources but the user must then visit the archive to work with the sources. Where the EAD does point directly to digital content, it is normally only at the level of a simple link. Since both the EAD and TEI are XML-based, it made sense to investigate how they might be integrated more closely to provide a working environment more akin to what researchers might need. This was the focus of the Linking EAD to Electronically Retrievable Sources (LEADERS) Project which was carried out at SLAIS from late 2001 to early 2005 with funding from the Arts and Humanities Research Council (LEADERS Project, 2004).

LEADERS has created a toolkit which integrates not only EAD and TEI (for transcriptions of sources), but also NISO Mix, a standard for specifying the technical structures of digital images, and the Encoded Archival Context (EAC), a new standard for archival contextual information such as biographical details about persons mentioned within the archive. Part of the research in LEADERS was an investigation of the overlap which occurs between some of the information in the EAD and the TEI header (Sexton *et al.*, 2003). This overlap was rationalized and the LEADERS XML schema uses the EAD as its shell with elements from the other standards incorporated as extensions. An XML stylesheet can then be used to create standalone TEI documents by mapping some of the EAD elements on to TEI header ones.

LEADERS is a set of tools using web services technologies. These are very flexible and anyone with some programming knowledge can add functionality or derive new versions or views of the digital sources. The project built a demonstrator using samples from two collections at UCL, The George Orwell Archive and the College Archive (LEADERS Toolkit @ Sourceforge.net, 2004). The Orwell material contains

letters, notes and diaries. The College Archive material is very different in nature. It consists of minutes of meetings and student registers. Both sample collections exist in transcription and digital image form in LEADERS. They are intended to illustrate some of the functionality in LEADERS and are too small for serious use. The user can search on any feature that has been indexed through its XML markup. The demonstrator offers several different displays. In one the transcription appears in the left-hand pane with appropriate elements from the metadata in the right-hand pane. Links from the transcription go to EAD records for persons and for organizations. The transcriptions also show how cancellations (text which is crossed out in the original) can be handled. Various options are provided for which of the metadata elements are to appear. In another the digital image is shown with other metadata elements. In a third the digital image and transcription are shown together.

The displays are controlled by XML stylesheets which the LEADERS implementer creates. These can easily be changed to suit the needs of different projects. LEADERS is completely modular and so new software could be added, for example to allow the user to request certain elements in the display, or to do further processing such as select items encoded as dates within the documents and sort them into chronological order. The screen colours and layout are controlled by cascading style sheets and so a completely different screen design could be used by a LEADERS implementer without the need to interfere with the functionality of the software.

The decision on what features to encode cannot be taken lightly. If a feature is encoded it can be retrieved more easily, but the more encoding there is the more expensive the document creation is. Part of the research in LEADERS was a study of archive users. The results of a questionnaire in six archives and record offices, including the National Archives which returned by far the largest number of questionnaires, indicated strongly that most users were personal interest users interested in individuals, families or organizations, reflecting the widespread interest in family history (Sexton *et al.*, 2004). Academic and occupational users tended to focus more on topic-based research. An implementation of LEADERS for a collection relating to maritime history might want to encode the names of ships. That for military history might want to encode battlefields. Items for indexing are selected by a stylesheet which operates on XML elements defined by the implementer and so almost anything can be indexed. In order to emphasize the potential of XML-based markup, we decided not to allow a free text search in the demonstrator, but this would also be possible in a LEADERS implementation.

The Perseus Project: flexible processing of chunks of data

Many other humanities projects are based on XML, but one of the longest running demonstrates much more of the real potential of the digital medium. The Perseus Project, now based at Tufts University, began at Harvard in the late 1980s (Crane, n.d.). Now described as an "Evolving Digital Library", it was initially designed as a multimedia database of Ancient Greek to be used as a teaching tool. It now contains several collections, but the Greek texts best illustrate what it can do. Many aids are provided for the reader of a Greek text. A click on any word brings up a morphological (grammatical) analysis of the word. Each word is also linked to its entry in a major Greek dictionary where the user can see how that word is used elsewhere. The dictionary also has links from the quotations for the entry to the texts from which those

quotations are taken. Links go from names of places to an interactive map of Ancient Greece. Site plans are provided for archaeological sites many of which are also illustrated by photographs taken by the Perseus team. The library also contains images of coins, sculptures, vases and other artefacts. The contextual information associated with these also has links to places, persons described in the Perseus encyclopaedia and to textual sources.

From the beginning Perseus has used structured markup. It began with SGML and converted to XML as soon as was feasible. The Perseus database consists of very many XML files and a set of software tools that have been designed and written by the Perseus team (Smith *et al.*, 2000). The markup was designed to enable the links to be created automatically and on the fly. The very first publication from Perseus predated the web and was implemented in Hypercard for the Macintosh. The markup design enabled the project to move to the web swiftly. Other collections have now been added to Perseus, some in different disciplines, but all have been able to build on the existing infrastructure.

Perseus is an illustration of what a flexible and extensible digital library can really do. In effect it consists of chunks of information that can be processed and linked in many different ways. Much of the early contextual material in Perseus was written by the Perseus team, largely because of copyright restrictions caused by a lack of understanding on the part of copyright holders about what Perseus was trying to do. Now that publishers are more attuned to the digital medium, it should be possible to create similar resources which link to the appropriate entries in standard reference works rather than to have to rewrite existing knowledge.

Strategies for the future

Digital technology is definitely changing the way we work and, in the author's view, there needs to be a much longer and harder look at its potential. After 500 years of print we are moving into an environment that is very different. Most information now starts its life in digital form, usually created by Microsoft Word. From the perspective of a long-term computer user, it seems odd that the world's most frequently used electronic writing tool is designed to create print as an end in itself, not to generate print as a by-product of a more useful document format. But this is where we find ourselves now and it will take some time for attitudes to move on from this. Existing procedures and systems will have to be rethought as the volume of digital information increases. Perseus, LEADERS and other projects in the humanities are raising questions that did not exist before. Many of these projects are experimental, and it is not at all clear how their resources will be processed in the future, but with XML-based technologies we have begun to understand something about how to hold on to the intellectual investment in the content of the resource.

Reusability has become a crucial issue. Many early humanities resources were created by individual scholars for their own purposes. They lacked a means of describing these purposes and the intellectual rationale behind them. Resources therefore got lost, perhaps not physically but in a sense of usability, let alone usefulness. A collection of numeric or coded information is of little use if nobody knows what the numbers or codes represent. Re-usability also implies the possibility of adding information to a resource, or of taking subsets of it and perhaps creating a new resource by adding information to the subset. This calls into question exactly when

something becomes a new resource and perhaps returns to the author's earlier arguments about the chunking of information. Chunks should be derivable dynamically with appropriate metadata associated with them automatically. Perseus has now taken a step in this direction with the announcement in May 2005 of a new facility for developers to extract XML fragments of well-formed primary sources for use in their own systems.

The material in LEADERS was all hand-coded, something that is possible for a small-scale demonstrator project, but encoding is costly and we need to find new ways of dealing with it. Automatic systems for adding part of speech and other fairly simple forms of linguistic annotation are now in widespread use in corpus linguistics (Garside *et al.*, 1997). Can these be applied to humanities sources to add different types of markup that is likely to be more interpretive? Data mining techniques might also be applied to humanities sources but these would perhaps be more appropriate for historical databases, where much of the information is more factual, than for a literary text where authors are sometimes deliberately ambiguous in their intent.

Within the UK we are fortunate in that there is now some support for infrastructure and fundamentals for digital resources in the humanities. The Arts and Humanities Data Service (AHDS) is funded by the Joint Information Systems Committee and the Arts and Humanities Research Council as a "UK national service aiding the discovery, creation and preservation of digital resources in and for research, teaching and learning in the arts and humanities" (Arts and Humanities Data Service, 2003-2005). It operates as a distributed service with centres for history, archaeology, literature, languages and linguistics, visual arts and performing arts. The AHDS functions much more than as a data archive. It works hard to raise awareness of the issues and challenges associated with producing produced digital resources. Its guides to good practice are widely cited and it is addressing some of the intellectual challenges of metadata for digital resources. In late 2003 the Arts and Humanities Research Council established an Initiative for ICT in Arts and Humanities Research Programme (2005a). This programme is concentrating more on methodology. It has established a methods network which will examine and promote new methodologies and it has also funded a series of strategy projects which will not only help to chart the existing landscape better, but will also produce exemplars or generic tools which can be adapted and adopted by other disciplines.

Funding has also been secured for participation by the arts and humanities in the e-science programme (ICT in Arts and Humanities Research Programme, 2005b). Here it should become possible to begin to address some of the issues raised in an AHRB seminar on e-science held in spring 2004 where some of the requirements identified for the humanities fit well with the vision of a global humanities digital library. These included access to: distributed resources of data, computing power, storage space and applications; integration of remote heterogeneous data; automatic capture of metadata; access and finding tools, shared vocabularies and ontologies; and virtual networks (Anderson, 2004). With the work done under these initiatives and elsewhere we might just be taking some steps to make humanities digital resources that begin to exploit the full potential of the medium.

References

Ager, D.E., Knowles, F.E. and Smith, J. (Eds) (1979), *Advances in Computer-aided Literary and Linguistic Research: Proceedings of the 5th International Symposium on Computers in Literary and Linguistic Research*, University of Aston, Birmingham.

Aitken, A.J., Bailey, R.W. and Hamilton-Smith, N. (Eds) (1973), *The Computer and Literary Studies*, Edinburgh University Press, Edinburgh.

Anderson, S. (2004), "E-science (e-research) expert seminar: report on proceedings", available at: www.ahrcict.reading.ac.uk/activities/e_science.pdf (accessed 16 October 2005).

Arts and Humanities Data Service (2003-2005), available at: www.ahds.ac.uk (accessed 16 October 2005).

Bessinger, J.B. and Parrish, S.M. (1965), *Literary Data Processing Conference Proceedings*, IBM, White Plains, NY.

Busa, R. (Ed.) (1974), *Index Thomisticus: Sancti Thomae Aquinatis operum omnium indices et concordantiae*, Frommann-Holzboorg, Stuttgart.

Busa, R. (1980), "The annals of humanities computing: the Index Thomisticus", *Computers and the Humanities*, Vol. 14 No. 2, pp. 83-90.

Busa, R. (Ed) (1992), *Thomae Aquinatis opera omnia: cum hypertextibus in CD-ROM*, Editoria Elettronica Editel, Milano.

Bush, V. (1945), "As we may think", *Atlantic Monthly*, July.

Crane, G. (Ed.) (n.d.), "The Perseus Digital Library", available at: www.perseus.tufts.edu/ (accessed 16 October 2005).

DeRose, S., Maler, E. and Orchard, D. (Eds) (2001), *XML Linking Language (XLink) Version 1.0*, World Wide Web Consortium, available at: www.w3.org/TR/xlink/ (accessed 16 October 2005).

Denley, P. and Hopkin, D. (Eds) (1987), *History and Computing*, Manchester University Press, Manchester.

Denley, P., Fogelvik, S. and Harvey, C. (Eds) (1989), *History and Computing II*, Manchester University Press, Manchester.

Encoded Archival Description (2002), "Official EAD Version 2 website", available at: www.loc.gov/ead/ (accessed 16 October 2005).

Garside, R., Leech, G. and McEnery, A. (1997), *Corpus Annotation: Linguistic Information from Text Corpora*, Addison-Wesley Longman, Harlow.

Giordano, R. (1994), "The documentation of electronic texts using text encoding initiative headers", *Library Resources and Technical Services*, Vol. 38 No. 4, pp. 389-401.

Greenstein, D.I. (Ed.) (1991), *Modelling Historical Data: Towards a Standard for Encoding and Exchanging Machine-readable Texts*, Max-Planck-Institut für Geschichte, Göttingen.

Greenstein, D. (1994), *A Historian's Guide to Computing*, Oxford University Press, Oxford.

Griffin, S. (1998), "NSF/DARPA/NASA digital libraries initiative", *D-Lib Magazine*, July/August, available at: www.dlib.org/dlib/july98/07/griffin.html (accessed 16 October 2005).

Hockey, S. (1980), *A Guide to Computer Applications in the Humanities*, Duckworth, London.

Hockey, S. (2004), "The history of humanities computing", in Schreibman, S., Siemens, R. and Unsworth, J. (Eds), *A Companion to Digital Humanities*, Blackwell Publishing, Oxford, pp. 3-19.

ICT in Arts and Humanities Research Programme (2005a), available at: www.ahrcict.reading.ac.uk/ (accessed 16 October 2005).

ICT in Arts and Humanities Research Programme (2005b), "E-science", available at: www.ahrcict.rdg.ac.uk/activities/e-science/ (accessed 16 October 2005).

International Organization for Standards (1986), *Information Processing – Text and Office Systems – Standard Generalized Markup Language (SGML)*, International Organization for Standards, Geneva.

Internet Archaeology (1996), available at: http://intarch.ac.uk (accessed 16 October 2005).

Jones, A. and Churchhouse, R.F. (Eds) (1976), *The Computer in Literary and Linguistic Studies (Proceedings of the 3rd International Symposium)*, University of Wales Press, Cardiff.

LEADERS Project (2004), available at: www.ucl.ac.uk/leaders-project/ (accessed 16 October 2005).

LEADERS Toolkit @ Sourceforge.net (2004), available at: http://leaders.sourceforge.net/ (accessed 16 October 2005).

Mawdsley, E., Morgan, N., Richmond, L. and Trainor, R. (Eds) (1990), *History and Computing III: Applications in Research and Teaching*, Manchester University Press, Manchester.

Nelson, T.N. (1987), *Literary Machines*, Edition 87.1, Sausalito Press, Sausalito, CA, (self-published).

Rusbridge, C. (1995), "The UK electronic libraries programme", *D-Lib Magazine*, December, available at: www.dlib.org/dlib/december95/briefings/12uk.html (accessed 16 October 2005).

Sexton, A., Turner, C., Hockey, S. and Yeo, G. (2003), "TEI, EAD and integrated user access to archives", in Anderson, J., Dunning, A. and Fraser, M. (Eds), *Digital Resources for the Humanities 2001-2: An Edited Selection of Papers*, Office for Humanities Communication, London, pp. 243-62.

Sexton, A., Turner, C., Yeo, G. and Hockey, S. (2004), "Understanding users: a prerequisite for developing new technologies", *Journal of the Society of Archivists*, Vol. 25 No. 1, pp. 33-49.

Smith, D.A., Mahoney, A. and Rydberg-Cox, J.A. (2000), "Management of XML documents in an integrated digital library", *Markup Languages: Theory and Practice*, Vol. 2 No. 3, pp. 205-14.

Sperberg-McQueen, C.M. and Burnard, L. (Eds) (1994), *Guidelines for Electronic Text Encoding and Interchange (TEI P3)*, Association for Computers and the Humanities, Association for Computational Linguistics, and Association for Literary and Linguistic Computing, Chicago, IL and Oxford.

Sperberg-McQueen, C.M. and Burnard, L. (2001), *Guidelines for Electronic Text Encoding and Interchange (TEI P4)*, 2nd ed., Association for Computers and the Humanities, Association for Computational Linguistics, and Association for Literary and Linguistic Computing, Chicago, IL and Oxford.

Text Encoding Initiative (n.d.), "Yesterday's information tomorrow", available at: www.tei-c.org/ (accessed 16 October 2005).

Wisbey, R.A. (Ed.) (1972), *The Computer in Literary and Linguistic Research*, Cambridge University Press, Cambridge.

Yergeau, F., Bray, T., Paoli, J., Sperberg-McQueen, C.M. and Maler, E. (Eds) (2004), *Extensible Markup Language (XML) 1.0*, 3rd ed., available at: www.w3.org/TR/2004/REC-xml-20040204 (accessed 16 October 2005).

Corresponding author
Susan Hockey can be contacted at: s.hockey@ucl.ac.uk

Received 6 May 2005
Revised 14 August 2005
Accepted 12 September 2005

Interpreting the image: using advanced computational techniques to read the Vindolanda texts

Melissa Terras
School of Library, Archive and Information Studies, University College London, London, UK

Abstract

Purpose – This paper aims to provide an overview of the development of a computer system designed to aid historians in the reading of the stylus tablets from the Roman fort of Vindolanda. It proposes outlining the different stages in developing the system, and giving the preliminary results.

Design/methodology/approach – The paper provides a literature review regarding Vindolanda, stylus tablets, and the process of reading an ancient document. Knowledge elicitation techniques are used to model explicitly expert processes used to read an ancient document. A corpus of character forms and lexicostatistics is gathered. An advanced cognitive imaging system utilising artificial intelligence techniques is implemented to produce plausible interpretations of the document.

Findings – This paper describes the developmental stages undertaken to construct a system that can read in images of an ancient document and produce plausible interpretations of the document, to aid the historians in the lengthy process of reading an ancient text. In carrying out the development, an explicit representation of how experts approach and reason about damaged and deteriorated texts was formulated, and a large corpus of letter forms and linguistic data were captured. Preliminary results from the resulting computer system are presented which demonstrate the usefulness of the technique, although more work is needed to develop this into a stand-alone computer system.

Research limitations/implications – The study is focused on the Roman stylus tablets from Vindolanda, near Hadrian's Wall, although the technique could be extrapolated to cover other types of ancient documents from any period.

Practical implications – It is demonstrated that using techniques from artificial intelligence and cognitive psychology can result in further explicit understanding of humanities expert processes, which allow computational systems to be constructed. The resulting computational system is a tool for the humanities expert, which carries out a task in a similar manner, allowing for faster reasoning time and quicker hypotheses development.

Originality/value – The paper presents the first known system to intake an image of an ancient text and output a plausible interpretation of the text in a reasonable time frame, assisting the papyrologist in resolving ambiguities in the damaged and abraded text.

Keywords Historical research, Knowledge processes, Artificial intelligence, Image processing

Paper type Research paper

Thanks must go to the supervisors of the author's doctoral thesis, from which this paper stems: Professor Alan Bowman (Centre for the Study of Ancient Documents, University of Oxford), and Professor Mike Brady (Department of Engineering Science, University of Oxford). The GRAVA system was adopted and adapted with the consent and help of Dr Paul Robertson (formerly of the Robots Research Group, Department of Engineering Science, University of Oxford, and now Research Scientist of Massachusetts Institute of Technology Computing Science and Artificial Intelligence Laboratory) and this research could not have come to fruition without his help and generosity. Thanks must also go to the anonymous papyrologists' guinea pigs who gave considerable time and effort to allow this study to take place.

Aslib Proceedings: New Information Perspectives
Vol. 58 No. 1/2, 2006
pp. 102-117

0001-253X
DOI 10.1108/00012530610648707

1. Introduction

The ink and stylus texts from Vindolanda are an unparalleled source of information regarding the Roman army and Roman occupation of Britain for historians, linguists, palaeographers, and archaeologists. The visibility and legibility of the handwriting on the ink texts can be improved through the use of infrared photography. However, due to their physical state, the stylus tablets (one of the forms of official documentation of the Roman army) have proved almost impossible to read.

This paper describes a system designed to aid historians in reading the stylus texts: in the process developing what appears to be the first system developed to aid experts in reading an ancient document. This system brings together years of research in image analysis, artificial intelligence (AI) architectures, papyrology, and palaeography to make a complete AI system that intakes a signal and outputs a symbol. The system is not an "expert system" that automatically "reads" and provides a transcription of the texts, rather it is a papyrologist's assistant that mobilises disparate knowledge, such as linguistic and visual clues, and uses these to speed up the process by which an expert can arrive at the most likely interpretation of a text.

Before designing and building any tools to aid papyrologists in the reading of texts, it is a necessary requirement first, to ask: "just what does a papyrologist do when trying to read and understand an ancient text?" A proposed model of how experts read ancient documents is presented, and this representation was subsequently used as a basis for the development of the computer system which can aid historians in the reading of the Vindolanda texts by "reading" texts in the same manner that they do themselves. The architecture of the system was developed from a system previously designed as a proof of concept to analyse aerial images, and the system was trained on linguistic and palaeographic corpus data gathered from previously read Vindolanda texts.

It can be demonstrated that the system takes in an image of unknown text as input and outputs plausible interpretations of that text by propagating various solutions to the problem and presenting the expert with the most likely results. This aids the papyrologist in sorting through differing hypotheses while trying to read an ancient text, which is often a lengthy and difficult process. The system presented here is currently being developed into a stand-alone desktop system, and it is envisaged that the tools will be applicable to different linguistic systems should there be data sets to train the system on.

Section 2 of this paper provides background information regarding the Vindolanda texts, and Section 3 discusses related research undertaken on the stylus tablets and stresses the need for the development of a cognitive visual system to aid in the reading of the tablets. Section 4 provides an overview of the knowledge elicitation techniques utilised in order to construct an explicit model of the processes used to read an ancient text, Section 5 details the gathering of corpus information and the construction of models to train the system, and Section 6 describes the architecture used to construct the system and presents results from the system in action. A summary of this research is given in Section 7 which demonstrates the contribution this research has made towards constructing a stand alone system to aid papyrologists in their task.

2. The Vindolanda texts

The discovery of the tablets at Vindolanda[1], a Roman Fort built in the late 80s AD near Hadrian's Wall at modern day Chesterholm, has provided an unparalleled

resource regarding the Roman occupation of northern Britain and the use and development of Latin around the turn of the first century AD. Textual sources for the period in British history from AD 90 to AD 120 are rare, and the ink and stylus tablets are a unique and extensive group of documents providing a personal, immediate, detailed record of the Roman Fort at Vindolanda from around AD 92 onwards (Bowman and Thomas, 1983, 1994, 2003; Bowman, 1999).

The ink tablets, carbon ink written on thin leaves of wood cut from the sapwood of young trees, have proved the easiest to decipher. In most cases, the faded ink can be seen clearly against the wood surface by the use of infrared photography, a technique used frequently in deciphering ancient documents (Bearman and Spiro, 1996). The majority of the 600 writing tablets that have been transcribed so far contain personal correspondence, accounts and lists, and military documents (Bowman and Thomas, 1983, 1994, 2003).

The 200 stylus tablets found at Vindolanda appear to follow the form of official documentation of the Roman army found throughout the Empire (Turner, 1968; Fink, 1971; Renner, 1992). It is suspected that the subject and textual form of the stylus tablets will differ from the writing tablets as similar finds indicate that stylus tablets tended to be used for documentation of a more permanent nature, such as legal papers, records of loans, marriages, contracts of work, sales of slaves, etc. (Renner, 1992), although the linguistic aspects of the tablets will be similar as they are contemporaneous documents from the same source, probably written by the same scribes (see Figure 1).

Note: The incisions on the surface can be seen to be complex, whilst the wood grain, surface discoloration, warping, and cracking of the physical object demonstrate the difficulty papyrologists have in reading such texts

Figure 1.
Stylus tablet 836, one of the most complete stylus tablets unearthed at Vindolanda

Manufactured from softwood with a recessed central surface, the hollow panel of the stylus tablets was filled with coloured beeswax. Text was recorded by incising this wax with a metal stylus, and tablets could be re-used by melting the wax to form a smooth surface. Unfortunately, in nearly all surviving stylus tablets[2] the wax has perished, leaving a recessed surface showing the scratches made by the stylus as it penetrated the wax[3]. In general, the small incisions are extremely difficult to decipher. Worse, the pronounced wood grain of the fir wood used to make the stylus tablets, staining and damage over the past 2,000 years, and the palimpsestic nature of the re-used tablets further complicate the problem; a skilled reader can take several weeks to transcribe one of the more legible tablets, while some of the texts defy reading altogether. Prior to this research, the only way for the papyrologists to detect incisions in the texts was to move the text around in a bright, low raking light in the hope that indentations would be highlighted and candidate writing strokes become apparent through the movement of shadows, although this proved frustrating, time consuming, and insufficient in the transcription of the texts.

3. Research and the Vindolanda texts

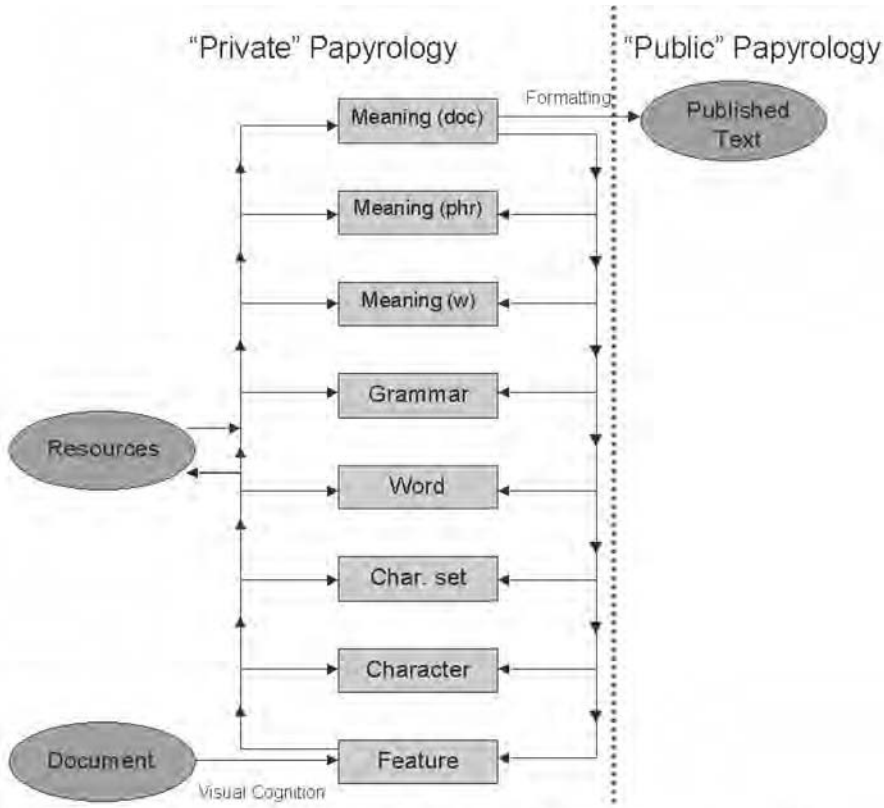
Novel imaging techniques have been developed at the Department of Engineering Science, University of Oxford, to analyse these texts (Bowman *et al.*, 1997; Schenk, 2001; Molton *et al.*, 2003; Brady *et al.*, 2005; Pan *et al.*, 2004). These techniques involve scrutinising the document surface, using image-processing techniques to provide new information that the human eye cannot see. However, these techniques do not provide a means of sorting through this information to provide the most likely interpretation of the data. Therefore, it was necessary to develop a computer system to aid the historians in the reading and interpretation of these images themselves, to speed up the reading process: the system is designed to work in combination with these image-processing algorithms and the experts themselves.

4. The process of reading an ancient document

In order to identify the tools that could be built to aid the papyrologists in their transcription of the Vindolanda tablets, it was first necessary to try to gain an understanding of what the papyrology process actually entails. Although the readings generated from ancient documents provide one of the major primary information sources for classicists, linguists, archaeologists, historians, palaeographers, and scholars from associated disciplines, surprisingly little research has been carried out regarding how an expert constructs meaning from deteriorated and damaged texts (Terras, 2002)[4]. Techniques borrowed from the field of knowledge elicitation (McGraw and Harbison-Briggs, 1989; Waterman, 1986), a structured process used to gain understanding of the behaviour of experts in order to train expert systems, were used to gather quantitative and qualitative information about how papyrologists work, resulting in an in-depth understanding of the ways different experts approach and reason about damaged and abraded texts. First, as with all knowledge acquisition tasks, the domain literature was researched, and any associated literature was collated. Three experts were then identified who were working on the ink and stylus texts, and who were willing to take part in this investigation. The experts were observed while going about their tasks, and unstructured interviews were undertaken, where the experts described their domain, and the individual processes and techniques that they preferred. More structured interviews were then undertaken, when the experts were

asked to describe particular facets of their task, such as the identification of letter forms and the role of grammar, word lists, and external historical and archaeological resources in the reading of the documents. A series of think aloud protocols (TAPs) were then undertaken (a technique adopted from experimental psychology, where the expert is urged to utter every thought that comes to mind while undertaking a specified task) and the experts were given structured tasks to complete. These sessions were recorded, transcribed, and analysed using content analysis techniques (Terras, 2002, 2005). These exercises demonstrated that the experts use a recursive reading mechanism which oscillates between different levels, or modules, of reading, and the process was rationalised into defined units, to develop a connectionist model of how papyrologists approach and start to understand ancient texts.

The proposed model of how experts read ancient text is presented in Figure 2. An expert reads an ancient document by identifying visual features, and then incrementally building up knowledge about the document's characters, combinations of characters, words, grammar, phrases, and meaning, continually proposing



Note: Public papyrology (Youtie, 1963) refers to the published reading of texts in a common format, after the reading has taken place. Private papyrology is the implicit process the experts undertake when reading a text

Figure 2.
The proposed model of how experts read an ancient text

hypotheses, and checking those against other information, until she/he finds that this process is exhausted. At this point a representation of the text is prepared in the standard publication format. At each level, external resources may be consulted, or be unconsciously compared to the characteristics of the document. Although a simple representation, the model shows the overall scope of the process of reading an ancient text: further, more detailed procedural analysis regarding every level (or "agent") in this system is presented in Terras (2002).

5. Gathering corpus information

In order to train a computational system to function in the way that the model, above, does, it was necessary to collate information regarding the type of letter forms and language used at Vindolanda in such a way that it could be incorporated into the system. The Vindolanda ink tablet corpus is the sole major contemporaneous linguistic resource to the Vindolanda stylus tablets, and this was used to generate a corpus of textual data and letter forms regarding Latin of this period. The palaeographic data were then manipulated to produce a set of "average" characters which the system uses for comparison with unknown stroke input data.

5.1 Textual corpus

The Vindolanda ink corpus used comprised the 230 Latin texts published in Bowman and Thomas (1994), plus 56 new texts that had been read in preparation for the publication of the next volume of the Vindolanda texts (Bowman and Thomas, 2003). There were 27,364 characters (excluding space characters) in total, comprising 6,532 words, or word fragments. In the corpus there were 2,433 unique word tokens (1,801 words appeared only once, the rest were repeated). This provided an adequate corpus on which to base any conclusions about the language used at Vindolanda[5].

Additional linguistic analysis of the Vindolanda ink texts was undertaken (Terras, 2002) using a common Corpus Linguistics program for lexical analysis, WordSmith (www1.oup.co.uk/elt/catalogue/Multimedia/WordSmithTools3.0/), in order to generate lexicostatistics such as word lists, letter frequency, and bigraph frequency. This provided statistics regarding the language used at Vindolanda, which could then be integrated into the system described here.

5.2 Image corpus

The ink and stylus tablets from Vindolanda use a form of handwriting commonly referred to as Old Roman Cursive (ORC). There is a paucity of documents from this period of the Roman Empire, and so the development of the character forms in ORC is the subject of much academic debate (see Bowman and Thomas, 1983, 1994, 2003). In order to train the system, a selection of characters from the Vindolanda ink texts previously read by Bowman and Thomas were encoded using a markup scheme and a tool developed to annotate aerial images to train AI systems (Terras and Robertson, 2004). This resulted in an XML representation of each character on a stroke-by-stroke basis that captured information about stroke length, position, direction, and relation to other strokes. In total, 1,506 individual characters from the ink tablets, and 180 characters from the stylus tablets were annotated, providing a training set for the system.

Characters were annotated first, by drawing around the outline of a character using a computer mouse, and assigning it a character label. Individual strokes were then

traced, and numbered by selecting the option from a drop down menu. Stroke ends were then identified and labelled. Finally, stroke junctions were noted and assigned labels. An example of this is shown in Figure 3, using the letter S from the start of ink tablet 311 as an example.

These annotations are preserved in an XML file that textually describes the annotated image. The placing of the strokes is preserved by noting the co-ordinates of each feature. The annotation of the corpus represents approximately 300 hours' worth of work, and is the largest palaeographic corpus of ORC in existence. The corpus can also be used to analyse the ORC letter forms (Terras, 2006).

5.3 Generating models from the corpus

The primary aim of building such a corpus was to provide data with which to drive an AI system, which could input unknown text and output probable interpretations. To match unknown to known characters, it is necessary to have models of those known characters to compare the unknown data to. Annotating the character forms in the way

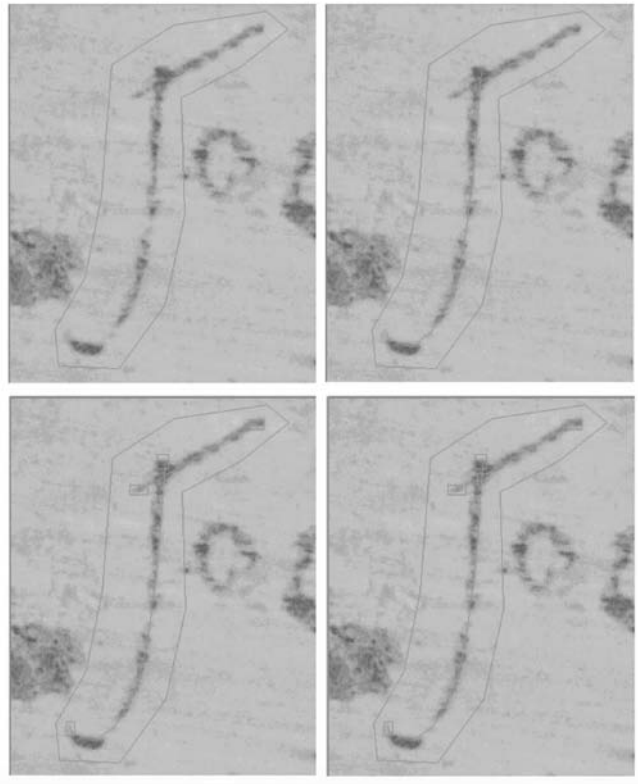


Figure 3.
Steps taken in annotating a letter

Note: The outline is traced at first, followed by individual strokes, stroke endings, and stroke meeting junctions. Further information about the markup used to annotate these characters can be found in Terras and Robertson (2004)

described above meant that vector co-ordinates for all the characters had been captured within the XML data, and so character models for each type of character could be generated.

A character model is defined as a probability field that indicates the likely placing of one or more strokes of a two-dimensional character, producing a general representation of a character type. Unknown characters can then be compared to a series of these models, and the probability that they are an instance of each one calculated, the highest probability indicating a match.

On a conceptual level, the (stroke-based) character model is constructed by taking an image of an individual character, finding its bounding box (identifying the rightmost x co-ordinate, and the leftmost x co-ordinate, and the highest and lowest y co-ordinates), and transforming this into a standardised (21 by 21 pixel) grid. The stroke data are blurred slightly to produce a generalised model which allows specific instances to be compared to it more successfully[6]. Each standardised representation is accumulated onto a generalised matrix for each character type: resulting in a generalised representation of each type of character. These are subsequently used as the models to which unknown characters are compared. An example of how these steps combine to generate a character model is given in Figure 4, where a small corpus which contains three "S" characters is used to generate a character model of an S.

Character models of every character contained within the dataset were generated, as shown in Figure 5.

These models were used to train the AI system to effectively "read" in stroke data of unknown text and output possible, reliable, interpretations of that text. By combining these stroke models with the linguistic data and advanced AI techniques, a system was constructed that was able to replicate the process the historians went through to output possible interpretations of the Vindolanda texts.

6. Developing an AI-based system

AI has repeatedly shown that interpreting our world requires bringing to bear a great deal of world knowledge. The process of reading an ancient document, especially documents that are in a very bad state of disrepair, such as the Vindolanda stylus tablets, is an especially good example of the need to mobilise a great deal of a priori knowledge. It is clear that the remnants of writing on the tablets themselves contain insufficient information to recover the original written text, but by applying sufficient linguistic knowledge of character shapes, word and letter frequency, and grammatical information, the tablets can sometimes be read. Implementing a system that can automate the process requires an architecture that can fuse the many different kinds of knowledge in order to arrive at an estimate of the most probable interpretation. Numerous architectures have been proposed to solve interpretation problems: the HEARSAY system (Erman *et al.*, 1980) was developed for speech recognition and more recently hidden Markov models (HMM) have been used in natural language processing (NLP) (Charniak, 1993) but these approaches have drawbacks. Blackboard systems can be hard to control, and HMMs can be too restrictive to incorporate complex and diverse kinds of knowledge (Robertson, 1999, 2001; Robertson and Laddaga, 2002, 2004).

To construct a computational model of the papyrology process, we adopted, and adapted, the Grounded Reflective Adaptive Vision Architecture (GRAVA) architecture, developed by Dr Paul Robertson[7], which was built to provide a flexible backbone on

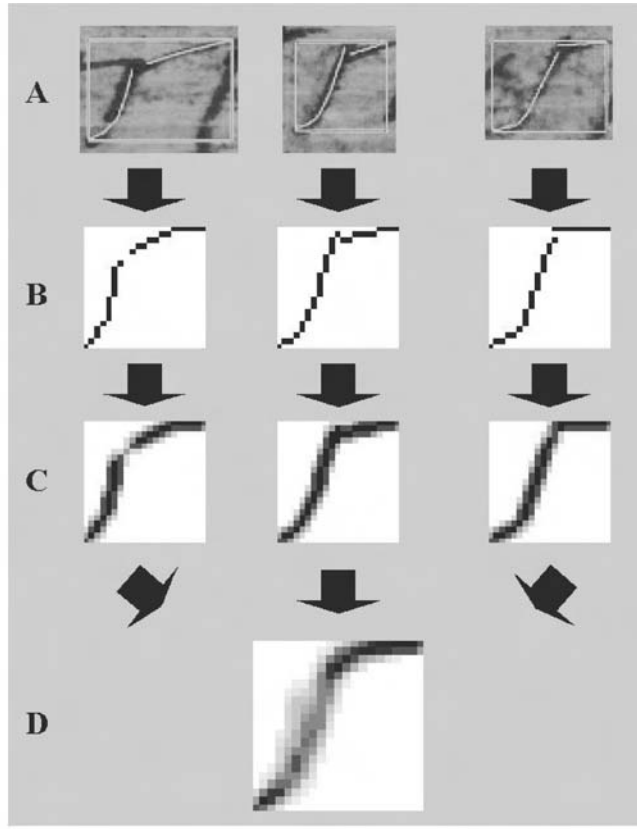
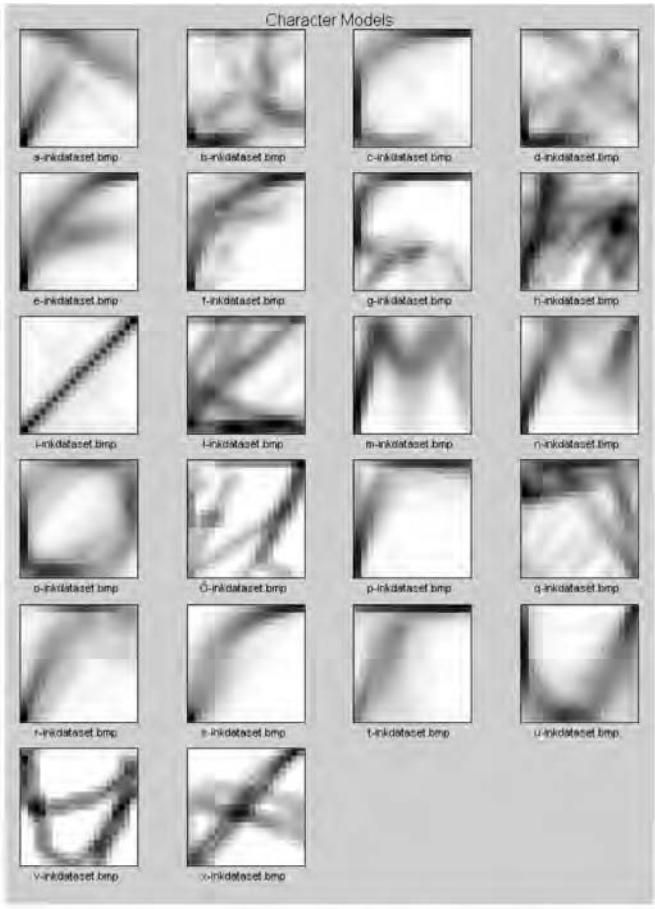


Figure 4.
The generation of a character model from a small corpus

Note: Three letter S's are identified, and bounding boxes drawn around them (A). The stroke data is then transformed into a 21 by 21 pixel grid (B). A Gaussian Blur is applied (C). The composite images generate a character model (D). The darker the area, the higher the probability of the stroke passing through that pixel. In this way, the probabilities of the stroke data occurring are preserved implicitly in the character models

which image interpretations problems can be solved. The GRAVA system was originally developed to analyse aerial satellite images (Robertson, 1999, 2001). The GRAVA architecture uses minimum description length combined with Monte Carlo Select algorithms to find a fair way of comparing data across semantic levels, and is implemented in YOLAMBDA, a dialect of LISP. To test his system, Robertson constructed a small program which could effectively "read" a nursery rhyme, utilising data regarding letter form, letter frequency, and word frequency, to propagate the most likely interpretation of the image of the text. This system was adapted considerably to be able to work with the Vindolanda data, and the architecture and mathematical underpinnings of the adapted system developed for this project are described in detail in Robertson *et al.* (2005), Terras and Robertson (2005), and Terras (2006). It is enough, here, to show a model of the system architecture, which demonstrates that it replicates



Note: The darker the area, the higher the certainty that a stroke will pass through that individual section of the character box. The letter M, for example, can be clearly seen, suggesting that it is made in a similar fashion each time it is used. The letter H is messier, but this is partly due to the number of letter Hs available to make the character models: the more letters available, the higher the accuracy of the character model. There is no letter j, k, w, y or z, as they were not part of the Old Roman Cursive alphabet

Figure 5.
Character models generated from the training set of the ink text corpus

the lower modules, or "agents" of the papyrology process, using AI techniques to sort through the data to propagate the most likely interpretation of the image.

All agents work with data generated from the Vindolanda corpus, as discussed in section 5: the character agent uses the set of character models plus data regarding the frequency of letters in the Vindolanda corpus, while the word agent utilises a list of words generated from the documents read at Vindolanda so far.

The basic schematic of the system is shown in Figure 6. To input an image of an unknown document, this image has to be annotated (either manually or automatically, using the image processing algorithms) to identify the key

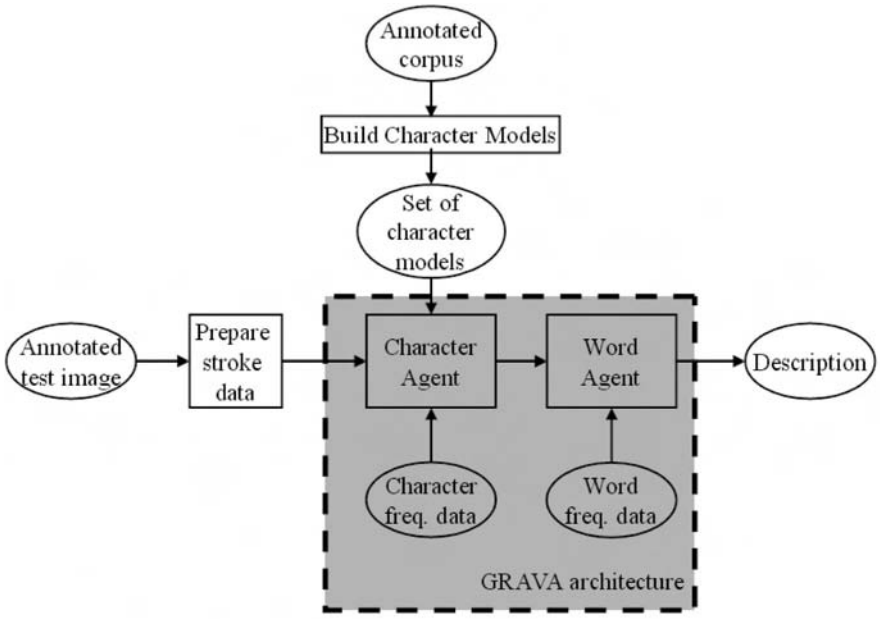


Figure 6.
Basic schematic of system

Note: Robertson's GRAVA architecture is highlighted to indicate the processes which are carried out as part of the final run of the system

features, and the stroke data prepared for comparison (for more details see Robertson *et al.* (2005)). The stroke data are then passed on to the character agent, which compares this data to the character models (generated previously from the annotated corpus), and calculates a description length based on which character it is most likely to fit, and the frequency of characters which appear in the corpus ("i" is much more likely to appear than "q", for example). One of these characters is selected by a random process, and is passed, with its description length, to the word agent. As characters are passed to the word agent, they form a string of characters, these resulting "words" are then compared to known words in the corpus, and a description length is calculated for how well they fit known words from Vindolanda. The description length from the character agent and the word agent is then added, giving a global description length for that run of the system. Subsequent runs of the system generate different global description lengths, as the random process in selecting which letters are passed to the word agent assures that different combinations of characters that may fit are put forward as possible interpretations of the problem. When the description lengths of these separate runs of the system are compared, the minimum description length gives the overall most likely interpretation of the test image. The system has effectively sorted through the different permutations of the data, and provided the experts with the most likely solution to the problem.

6.1 The system in action

When implemented in practice, the system outputs the most likely interpretation of the texts, as can be seen in Figures 7 and 8. An annotated image is entered into the system, and after eight iterations, the system outputs the correct interpretation of the text, keeping note of the alternative possible hypothesis it has also thrown up.

More remains to be done to make this a useful tool for the papyrologists. The word lists contained within the system need to be expanded, to allow for fragmentary sections of words that may not appear in the existing corpus. The process of annotating a document and feeding it into the system is time consuming, and requires an understanding of the annotation tools themselves. The system will require development to become a stand-alone desktop application, as it currently depends on the knowledge engineer and engineering scientist being present to run it (this development is being undertaken at the Department of Engineering Science and the Centre for the Study of Ancient Documents at Oxford University, in collaboration with the School of Library, Archive and Information Science at University College London). Nevertheless, the system is a working tool that can take in images of texts and output plausible interpretations of those texts, which is speeding up the process of reading the Vindolanda stylus texts by propagating realistic interpretations of complex data. The

Note: *Ussibus puerorum meorum* means "for the use of my boys", referring to some cloaks and tunics for Clodius Super's *pueri*

Figure 7.
A section of ink tablet 255, which reads "*ussibus puerorum meorum*", used as input

```
Grounded Reflective Agent Vision Architecture (GRAVA) Version 2.0.
Yolambda listener pushed. Type :exit to return to GRAVA.

=> ... load the system and the data ...

=> (runCycles 25)
iteration 0 DL=440.220794 interpretation=( ... ((2482 252) (2517 250))) ... )
iteration 1 DL=64.085075 interpretation=(  u r s i b u s puerorum  m n o a u m  )
iteration 2 DL=49.374412 interpretation=(ussibus puerorum  m n o r u m  )
iteration 3 DL=48.831413 interpretation=(ussibus puerorum  m n o a u m  )
iteration 5 DL=47.816696 interpretation=(ussibus puerorum  m e o a u m  )
iteration 8 DL=36.863136 interpretation=(ussibus puerorum meorum  )
iteration 25

=> :exit
```

Note: The system outputs each interpretation which is deemed to be more successful than the last, and so builds up a list of plausible interpretations. The last solution matches the correct solution to the problem

Figure 8.
Output from first successful run on the section of 255, indicating that the system resolves upon the correct interpretation of the image

system is also being developed to work with the image processing algorithms, so that particularly damaged and complex incisions of unknown text can be automatically fed into the system (Robertson *et al.*, 2005). Other types of graphical and linguistic textual systems could also be analysed using this technique, provided the statistical data regarding letter forms, word frequency, and letter frequency could be obtained.

Conclusion

This research has discussed the development of a computer system to aid historians in the reading of ancient texts, from initial demonstration of the problem, work with experts to ascertain their needs, the gathering of relevant information and statistical data, and identifying, adopting and adapting a technology which enables the disparate knowledge used to read such texts to be mobilised. As such it is a proof of concept. More development needs to be undertaken before this system becomes a stand-alone application, but its success is grounded on the fact that it has such a rich dataset to draw information from, and that working closely with the papyrologists from the outset has enabled the construction of a system that behaves in a similar manner to the experts themselves, making this a useful and intuitive tool which speeds up the time taken to read an ancient text.

There were further unexpected rewards from this research. The data set of ORC letter forms has proved useful in determining palaeographic differences between the characters which appear on the stylus and ink texts (Terras, 2006). The modelling of the papyrology process is the first detailed analysis of how experts who read ancient texts operate (Terras, 2002, 2006), allowing an insight into how texts which are so important to our understanding of history and society can actually be deciphered and read. From a computer science point of view, the Vindolanda texts provided real world data to test the implementation of a system that had just been applied to a simplistic problem (the nursery rhyme data). The technologies used have demonstrated that information from differing semantic levels, such as image and word data, can be compared in a reasonable time frame utilising our comparison technique and this could be applied to comparing any sort of differing data in computer systems (up until now, it has been time consuming and computationally expensive to compare and contrast differing semantic data computationally).

This research is also a demonstration of fruitful collaboration between differing academic tribes and territories: the use of advanced computational techniques in a humanities discipline benefiting both the scholars in the humanities and the computer and engineering scientists. Future work will entail delivering this system as an application for papyrologists to use, aiding the experts in their complex and time-consuming task.

Notes

1. See Bidwell (1985) for further information regarding Vindolanda, Birley (1999) for an account of the discovery of the tablets, and Bowman (2003) for an introduction to the content of the Vindolanda texts.
2. It is estimated that around 2,000 of such tablets exist outside Egypt (Renner, 1992).
3. Only one stylus tablet, 836 (see Figure 1), has been found so far with its wax intact. Unfortunately this deteriorated during conservation, but a photographic record of the waxed

tablet remains to compare the visible text with that on the re-used tablet. (The figure shown here is of the tablet after the wax was lost).

4. Aalto (1945), Youtie (1963, 1966) and Bowman and Tomlin (2005) are the only discussions found (as yet) which try to describe what the papyrology process actually entails, with some higher level discussion available in Turner (1968).
5. The representativeness of a corpus has been much discussed in the field of Corpus Linguistics (Kenny, 1982; Biber, 1983; Oostdijk, 1988; Biber, 1990), the consensus being that "small is beautiful" and that "there is every reason to make maximal use of these corpora [of 2,000 word length] for analysis of linguistic variation until larger corpora become readily available" (Biber, 1990, p. 269) The major corpus of American English, *The Standard Corpus of Present Day Edited American English* (known as the Brown Corpus, see Francis and Kucera (1979)) uses 2,000 word samples to represent its various genres. Zipf's "law" (Zipf, 1965), which predicts an "ideal" set of frequencies (and hence probabilities) for lexical items given a particular vocabulary size, maintains that a dictionary size of 1,000 words will give a percentage cover of 56.220 per cent of the language, although this is based on an analysis of English language texts. In the case of this research, the corpus used is 100 per cent of the known corpus of Latin for this period.
6. That is to say, the stroke data are convolved with a Gaussian Blur operator to reduce over-fitting.
7. Formerly of the Robots Group, Department of Engineering Science, University of Oxford. Now Research Scientist of the Computing Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology.

References

Aalto, P. (1945), "Notes on methods of decipherment of unknown writings and languages", *Studia Orientalia, Edidit Societas Orientalis Fennica*, Vol. XI No. 4.

Bearman, G.H. and Spiro, S. (1996), "Archaeological applications of advanced imaging techniques", *Biblical Archaeologist*, Vol. 59 No. 1, pp. 56-66.

Biber, D. (1983), "Representativeness in corpus design", *Literary and Linguistic Computing*, Vol. 8 No. 4, pp. 243-57.

Biber, D. (1990), "Methodological issues regarding corpus-based analysis of linguistic variation", *Literary and Linguistic Computing*, Vol. 5 No. 4, pp. 257-69.

Bidwell, P.T. (1985), *The Roman Fort of Vindolanda at Chesterholm, Northumberland*, Historic Buildings and Monuments Commission for England, London.

Birley, R.E. (1999), "Vindolanda IV: Fasc. IV", *Writing Materials*, Roman Army Museum Publications, Carvoran.

Bowman, A.K. (1999), "The Vindolanda Writing-Tablets 1991-1994", *Atti del XI Congresso Internazionale di Epigrafia Greca e Latina, Roma, 18-24 Settembre*, Vol. I, pp. 545-51.

Bowman, A.K. (2003), *Life and Letters on the Roman Frontier, Vindolanda and Its People*, 2nd ed., British Museum Press, London.

Bowman, A.K. and Thomas, J.D. (1983), *Vindolanda: The Latin Writing Tablets (Tab. Vindol. I.)*, Society for Promotion of Roman Studies, London.

Bowman, A.K. and Thomas, J.D. (1994), *The Vindolanda Writing Tablets (Tab. Vindol. II)*, British Museum Press, London.

Bowman, A.K. and Thomas, J.D. (2003), *The Vindolanda Writing Tablets (Tab. Vindol. III)*, British Museum Press, London.

Bowman, A.K. and Tomlin, R.S.O. (2005), "Wooden stylus tablets from Roman Britain", in Bowman, A.K. and Brady, M. (Eds), *Images and Artefacts of the Ancient World*, Oxford University Press, Oxford.

Bowman, A.K., Brady, J.M. and Tomlin, R.S.O. (1997), "Imaging incised documents", *Literary and Linguistic Computing*, Vol. 12 No. 3, pp. 169-76.

Brady, M., Pan, X., Terras, M. and Schenk, V. (2005), "Shadow stereo, image filtering and constraint propagation", in Bowman, A.K. and Brady, M. (Eds), *Images and Artefacts of the Ancient World*, Oxford University Press, Oxford, pp. 15-30.

Charniak, E. (1993), *Statistical Language Learning*, Massachusetts Institute of Technology Press, Cambridge, MA.

Erman, L.D., Hayes-Roth, F., Lesser, V.R. and Reddy, D.R. (1980), "The HEARSAY-II speech understanding system: integrating knowledge to resolve uncertainty", *Computing Surveys*, Vol. 12 No. 2, pp. 213-53.

Fink, R.O. (1971), "Roman military records on papyrus", *Philological Monographs of the American Philological Association*, No. 26, American Philological Association, Philadelphia, PA.

Francis, W.N. and Kucera, H. (1979), *Brown Corpus Manual*, Brown University, Providence, RI, available at: http://helmer.aksis.uib.no/icame/brown/bcm.html (accessed 25 February 2005).

Kenny, A. (1982), *The Computation of Style: An Introduction to Statistics for Students of Literature and Humanities*, Pergamon Press, Oxford.

McGraw, K.L. and Harbison-Briggs, K. (1989), *Knowledge Acquisition: Principles and Guidelines*, Prentice-Hall International Editions, London.

Molton, N., Pan, X., Brady, M., Bowman, A.K., Crowther, C. and Tomlin, R. (2003), "Visual enhancement of incised text", *Pattern Recognition*, Vol. 36 No. 4, pp. 1031-43.

Oostdijk, N. (1988), "A corpus linguistic approach to linguistic variation", *Literary and Linguistic Computing*, Vol. 3 No. 1, pp. 12-25.

Pan, X., Brady, M., Bowman, A.K., Crowther, C. and Tomlin, R.S.O. (2004), "Enhancement and feature extraction for images of incised and ink texts", *Image and Vision Computing*, Vol. 22 No. 6, pp. 443-51.

Renner, T. (1992), "The finds of wooden tablets from Campania and Dacia as parallels to archives of documentary papyri from Roman Egypt", *Proceedings of the 20th International Congress of Papyrologists, Copenhagen, 23-29 August*.

Robertson, P. (1999), "A corpus based approach to the interpretation of aerial images", *Proceedings of 7th International Congress on Image Processing and its Applications (IEE IPA99), Manchester*, Vol. 2, pp. 527-31.

Robertson, P. (2001), "A self adaptive architecture for image understanding", DPhil thesis, Department of Engineering Science, University of Oxford, Oxford.

Robertson, P. and Laddaga, R. (2002), "Principal component decomposition for automatic context induction", in Ishii, N. (Ed.), *Artificial and Computational Intelligence, Proceedings of the IASTED International Conference, Tokyo, Japan*, pp. 243-50.

Robertson, P. and Laddaga, R. (2004), "The GRAVA self-adaptive architecture; history; design; application; and challenges", *24th International Conference on Distributed Computing Systems Workshops (ICDCS 2004 Workshops)*, *Institute of Electrical and Electronics Engineers Computer Society*, *Piscataway, NJ*, pp. 298-303.

Robertson, P., Terras, M., Pan, X., Brady, J.M. and Bowman, A.K. (2005), "Image to interpretation: an MDL agent architecture to read ancient roman texts" (submitted for publication).

Schenk, V.U.B. (2001), "Visual identification of fine surface incisions", DPhil thesis, Department of Engineering Science, Oxford University, Oxford.

Terras, M. (2002), "Image to interpretation: towards an intelligent system to aid historians in the reading of the Vindolanda texts", DPhil thesis, Department of Engineering Science, University of Oxford, Oxford.

Terras, M. (2005), "Reading the readers: modelling complex humanities processes to build cognitive systems", *Literary and Linguistic Computing*, Vol. 20 No. 1, pp. 41-59.

Terras, M. (2006), *Image to Interpretation: An Intelligent System to Aid Historians in Reading the Vindolanda Texts*, Oxford University Press, Oxford, in press.

Terras, M. and Robertson, P. (2004), "Downs and acrosses, textual markup on a stroke based level", *Literary and Linguistic Computing*, Vol. 19 No. 3, pp. 397-414.

Terras, M. and Robertson, P. (2005), "Image and interpretation: using artificial intelligence to read ancient Roman texts", *Human IT*, Vol. 7 No. 3, pp. 1-56.

Turner, E.G. (1968), *Greek Papyri, An Introduction*, Clarendon Press, Oxford.

Waterman, D.A. (1986), *A Guide to Expert Systems*, Addison-Wesley, Reading, MA.

Youtie, H.C. (1963), "The papyrologist: artificer of fact", *Greek, Roman and Byzantine Studies*, Vol. 4, pp. 19-32.

Youtie, H.C. (1966), "Text and context in transcribing papyri", *Greek, Roman and Byzantine Studies*, Vol. 7, pp. 251-8.

Zipf, G.K. (1965), *The Psycho-biology of Language*, Massachusetts Institute of Technology Press, Cambridge, MA (originally published in 1935).

Further reading

Schenk, V.U.B. and Brady, M. (2003), "Visual identification of fine surface incisions in incised Roman stylus tablets", paper presented at the ICAPR 2003, International Conference in Advances in Pattern Recognition, Calcutta, 10-13 December.

Corresponding author

Melissa Terras can be contacted at: m.terras@ucl.ac.uk

Received 13 April 2005
Revised 15 September 2005
Accepted 20 September 2005

Improving the relevance of web menus using search logs: a BBCi case study

Paul Huntington and David Nicholas
Centre for Information Behaviour and the Evaluation of Research (Ciber), School of Library, Archives and Information Studies, University College London, London, UK

Abstract

Purpose – The paper seeks to propose a method for selecting menu items based on an analysis of user-entered search terms. Menu pages inform users what is coming next and what questions are going to be answered by an information communication technology service. Menus need to reflect user needs. The paper aims to argue that users reveal the scope of their information needs by the words used in their search expressions and these can be analysed to inform menu titles.

Design/methodology/approach – The paper presents an analysis and classification of user search expressions that are automatically collected by the server. The paper examines the search expressions of about 1,000 users of the BBC site related to search expressions on diabetes.

Findings – The search expressions were classified, analysed and compared with the diabetes menu of three health sites: NHS Direct (www.nhsdirect.nhs.uk); BBC health (www.bbc.co.uk); and Diabetics UK (www.diabetes.org.uk). Finally, a six-point menu is derived.

Practical implications – The practical implication of this paper is development of relevant web menus based on user information needs as revealed in search expressions entered by users.

Originality/value – This is the first explanation of how search logs can be used to construct menu lists. Previously menus have been designed at worst to suit producers and site designers based on the information that they have available and at best on interviews with small usability or focus groups who are not necessarily users.

Keywords Search output, Information searches, Internet marketing, Worldwide web

Paper type Research paper

Introduction

The making available of information content on the web does not mean or create its usage. Site menus are of particular importance to the users' flow through to content as they are indicators of site content and give users an idea of how likely their query will be answered by the site. Menu item prominence is information immediacy. It tells the user early on that the information system can answer queries on a topic. We have found evidence in our own studies (Nicholas *et al.*, 2004) that many users will tend to terminate their session after just viewing one screen, perhaps leaving because they are not finding links and menu signposts for the information that they are looking for. What may in fact be happening is that if a user cannot see a menu signpost to what it is that they are looking for on the home page (immediacy) they will leave. That is, the user does not have a good reason to believe that the site can meet their information need and hence leaves. There is evidence to support this. Nicholas and Huntington (2005) found that over one-third (38 per cent) of users of health sites said they "often"

Aslib Proceedings: New Information Perspectives
Vol. 58 No. 1/2, 2006
pp. 118-128

0001-253X
DOI 10.1108/00012530610648716

left a site if they could not see what they were looking for on the page landed at and 56 per cent "often" left if they could not see what they were looking for within the first two screens looked at.

Menu pages offer a particular type of information. They are informing users what is coming next and what questions are going to be answered by the service. Virtually everything that is available on a site is out of immediate sight so users look at menus as indicators of content. The difficulty is how to decide what goes on this menu. Traditionally this has really been a hit or miss affair. At worse site designers or managers best guess a menu based on available content and what they subjectively think users would be interested in. For better-informed sites use is made of usability studies and focus groups, though these studies rarely include users and their remit is often just to judge the adequacy of a stated range of menu items or menu-related tasks.

The authors argue that search logs, which are a collection of user search expressions, offer an effective way of assessing and selecting menu items. Users enter search words, in their own language, without prompting and that reflect a genuine information need, to find information that they are looking for. This group is hence the best group for investigating information needs because they are actual users participating in a real process of searching for information. The search terms entered by users reveal their information need. Furthermore, the fact that these statements of information need are collected by server computers automatically, routinely, and at very little cost, argues that the information source offers a potent insight into searching and user behaviour. In fact the difficulty posed by the information source is the quantity of the data. We are moving from analysing a limited number of search expressions related to a usability group asked to image an information need to details of thousands of search expressions of users engaged in realising an actual information need. User information needs should be the foundation for constructing user-relevant menus. This paper analyses user search expressions, as a form of revealing information need, to develop a user-centred menu.

This research was funded by the Department of Health and forms part of ongoing research at UCL into the use and development of online health information.

Aims, objectives and scope of the study

This research analyses and classifies web search logs and extracts themes from this information. The subject themes are compared to existing menus and are used to construct and derive a menu. The research is limited to search expressions on diabetes.

Previous literature

Menus are recognized as key to a site. In understanding the success of a web-based system, navigation, information design, and access are identified as key usability factors (Preece, 2000). However, the design and construction of navigation menus for web sites are too often left to the web designer. A problem with developing menu labels based on this practice is that it is subjective, albeit by a professional designer, and is likely to result in loss of menu coverage. Zhang *et al.* (2004) investigates web-menu layouts and uses formal concept analysis (FCA) – a mathematical model to assist in the design and automatic generation of a navigation hierarchy for a set of web documents. This procedure examines for the clustering of attributes. The clustering algorithm determines which collection of attributes forms a coherent entity. The

problem of this approach is that the subject starting point is still based on subjective decisions and what information the site has and not on user information needs.

A typical method for investigating information needs regarding menu layout and construction is to observe a small group of participants as they work through a specific number of set-tasks. Eisenberg (2005) argues that this approach examines the site's interface and process barriers that keep visitors from accomplishing a conversion task. Typically usability studies look at terminology usage and adequacy of menu structure and so on. Frick *et al.*(1999) investigated performance differences between web-based navigation models using 44 college students using information-finding tasks and found no significant differences between the models. Ebenezer's (2003) usability evaluation of the recently launched South London and Maudsley NHS Trust library web site is an example of this. Usability tests were employed to explore how users' cognitive understanding of information presentation relates to the interfaces they interact with. Further examples of this are found in Hennig's (1998) examination of the Bose Corporation intranet site, and studies by Neilsen (2002). The problem of usability tests is that by their nature they are small-scale, are generally not populated with users with a genuine information need, are directed by a series of pre-defined usability or set tasks, the results are impacted and dependent on the test environment and are expensive. Neilson (Eisenberg, 2005) noted that usability test participants work harder on tasks in a test session than they would at home and concluded that usability tests cannot measure persuasive momentum and individual motivation. Usability tests aim to simulate reality in a laboratory setting with regard to specific set tasks and hence the results are limited to how well the simulation can be said to exist and limited by the boundaries of the set tasks.

There has been little previous research linking search transaction log data with site menu construction though point of site entry has been previously researched. Van der Geest (2002) found that point of arrival can be used to redesign site structure. If many visitors arrive on a page that is hidden three layers deep in the site then he argues the ordering of the information should be changed and the link to that information should be positioned at a higher level. Further, Rozic-Hristovski *et al.* (1999) studied site access, available by a log analysis, to evaluate a developing medical library web site in Slovenia. Both these studies used the page landed at as to some extent revealing the user's information need and hence arguing a greater digital prominence for that page. The study proposed here differs in that only search logs are examined. Other studies (e.g. D'Alessandro *et al.*, 1998) have used log data to examine the origin of users (by IP address) and their information behaviour. In a more search expression related study Goodrum and Spink (1999) conducted a specific analysis of 1.2 million queries for images. They further identified these queries as reflecting information needs but, however, only related this to the number of different search terms used. They did not go on to use this information for the critical construction of a menu, perhaps because the search expressions were too general and were not directed at a particular topic. Jansen *et al.* (2000) in a related study and again with expressions covering a range of topics reported that over half of the terms used were used only once. Further, Ozmutlu *et al.* (2004) looked at search logs and did a country comparative time-based study and study fluctuations over day and number of words in a query. Wolfram *et al.* (2001) classified a random sample of 2,414 queries from 1997 and 2,539 queries from 1999 into 11 broad subject categories to assess how web searching topics changed from 1997 to 1999.

Nicholas *et al.* (2002) found a relationship between service usage and menu prominence. The study found that both access to and "use", as measured by users returning to the site, and use and user statistics, of a site declined significantly over the survey period and these declines match almost exactly changes in the positioning of the service on the menus. As the service becomes more difficult to access as its sign posting becomes ever more removed from the opening menu, the proportion of new visitors as a percentage of all users declines. New users were not coming through because of the increasing difficulty of finding the service. This research underlines the importance of menus as a signpost of what is coming next and where this information is absent the use of a service will decline.

Methods

Logs of user search queries using BBCi were supplied by the BBC (www.bbc.co.uk) for 12 days in March 2003; two sequences were supplied for analysis: 1st-6th and 9th-15th. This sequence recorded 4,048,137 search queries made by 891,129 users across 1,035,514 sessions: a testament to the sheer popularity of the BBC web site and search facility. Lines related to diabetes were extracted. In all search transaction logs of 1,004 users were included. This covered 1,838 searches and 384 different search expressions. SPSS was used for all data selection, classification and analysis.

The log file relates to search queries entered by users. The file provides the following fields:

(1) Field 1: machine id (cgiper13).
(2) Field 2: process number.
(3) Field 3: process number.
(4) Field 4: date.
(5) Field 5: time.
(6) Field 6: search expression entered.
(7) Field 7: number of BBCi best links returned – best link may be an external or a BBC link.
(8) Field 7: operation time.
(9) Field 8: cookie.
(10) Field 9: scope of search – a subject category which represents a specific area within which the user initiated the search. A scope search confines the results returned within a specified number of directories (urls).
(11) Field 10: tab of search user defined to specify a search within the BBC or the world wide web (WWW).

There are potentially a number of interesting fields here. BBCi best links, these are short cuts to what BBC has identified as best links, are returned in response to a user search expression, although not all users are returned with a best link. The return of a BBCi best link may reflect the effectiveness of a user search expression. Also of interest is the "Scope of Search" field. This is a subject category which represents a specific area within which the user initiated the search. A scope search confines the results returned within a specified number of subject directories as this suggests that a user has landed in and chosen to limit their search to a specific area.

Results

To understand the analysis best it is probably best to highlight one user's interaction with the search engine. This particular person used the search facility 13 times over a 20-minute period between 11.32 p.m. and 11.53 p.m. on 12 March (Table I). The user's decision to return to the search time and time again was taken relatively quickly. Generally there was a gap of between 10 to 30 seconds between the re-framed search queries and argues that the user had quite quickly decided that nothing relevant was forthcoming from the links on offer.

The user first searched for diabetogenic, then repeated this search, presumable because they did not find anything, four times switching between tab of search options searching the WWW and allbbc (all BBC areas). This user then added to the search expression and searched for diabetogenic pregnancy, repeating this search three times, first for allbbc and then twice for the WWW. There followed a break of about 15 minutes. This break might reflect the user reading a link returned by the search engine; alternatively the user might have gone elsewhere or just decided to have a break. The user returned to the search facility at 11.52 pm and repeated six searches on contraception diabetes mellitus. The first two searches the user miss-spelt contraception. The user alternated the searches between WWW and allbbc. The user was not returned a BBCi best link in their session. In addition the user only used the facility once, did not return, and searched for a single topic, diabetes.

Table II gives the top 20 search expressions used. The top 20 expressions accounted for a high 58 per cent of all search expressions used. However, most users looked for pages on this topic by just typing in the single word diabetes: 39 per cent typed in diabetes, 4 per cent Diabetes and 1 per cent DIABETES. In all users used 384 search expressions to find topics related to diabetes. It seems from Table II that about half of user search expressions were just looking for anything on diabetes and typed in the word diabetes while about half were looking for more specific information related to diabetes.

The top 20 expressions (Table II) give an idea of the diabetes topics looked for. They include a surprising range of topics from feline diabetes, social aspects, to searches for a variety of types of diabetes. There were also searches on diabetes and diet, aspirin, and an interest in finding out about other web sites. Each of the topics in the main accounted for less than 1 per cent of search expression use and this in part reflects the

009ffdee867161faf9b4a5710d4d083e				12 March 2003	
21:32:59	Diabetogenic	0	0.00	www	homepage
21:33:08	Diabetogenic	0	0.00	allbbc	homepage
21:33:08	Diabetogenic	0	0.00	www	homepage
21:33:11	Diabetogenic	0	0.00	allbbc	homepage
21:35:36	Diabetogenic pregnancy	0	0.00	allbbc	
21:35:36	Diabetogenic pregnancy	0	0.01	www	
21:36:57	Diabetogenic pregnancy	0	0.01	www	
21:52:35	Contraception diabetes mellitus	0	0.01	allbbc	
21:52:35	Contraception diabetes mellitus	0	0.01	www	
21:52:52	Contraception diabetes mellitus	0	0.01	allbbc	
21:52:52	Contraception diabetes mellitus	0	0.01	www	
21:53:05	Contraception diabetes mellitus	0	0.00	allbbc	
21:53:18	Contraception diabetes mellitus	0	0.00	www	

Table I.
Micro analysis of user 1

Search expression	%
diabetes	39.1
Diabetes	4.2
type 2 diabetes	1.4
DIABETES	1.3
gestational diabetes	1.1
diabetes	1.1
articles on social aspects of diabetes in children	1.0
diabetes UK	1.0
type 2 diabetes	1.0
diabetes.	1.0
causes of diabetes	0.7
www.diabetes.org.uk	0.7
Travelling with Diabetes	0.6
diabetes mellitus	0.5
diabetes diet	0.5
aspirin and diabetes	0.5
feline diabetes	0.5
American diabetes association	0.5
diabetes symptoms	0.5
hockaday diabetes	0.4
Top 20 accounts for % of all expressions	57.7

Table II. The top 20 search expressions used to find topics related to diabetes

subject range and different word combinations used by users to search for their topic. Though clearly from the above groupings are suggested.

The 384 different search expressions were classified and grouped into 19 broad subject categories. The groupings and categories were refined and reflect a theme analysis that was completed on the search terms. This involved selecting a theme and looking at the number of search expressions that could be classified into the theme. The results are given in Figure 1 and give the percentage of unique user search expressions falling into each group. By unique we mean that double or more searches by a user using the same terms were excluded. Figure 1 represents the allocation into subject groupings of all search expressions but did not include expressions that included the single term diabetes. In all 84 per cent of search expressions were allocated into subject groups while 16 per cent could not be allocated. A total of 18 per cent of user search expressions looked for either contacts and or web sites; 16 per cent looked for different types of diabetes; 10 per cent related to diabetes and life stage (this grouped search expressions related to diabetes and pregnancy, children, diabetes late in life and so on). A total of 9 per cent related to diabetes and diet and types of diet; 6 per cent to different treatments for diabetes; 5 per cent included the word symptoms; 4 per cent to insulin and sugar; 3 per cent diagnosis; 3 per cent complications and 3 per cent to interactions. Causes of diabetes, travel, exercise, alcohol and smoking, statistics, obesity, aspirin and prevention each made up 2 per cent or less.

It was decided to compare the identified diabetes subject groupings as revealed by the search logs to the menu subject list available on three diabetes information services. The three services selected were: Diabetes on BBC health (www.bbc.co.uk), Diabetes on NHS Direct Online (www.nhsdirect.nhs.uk) and Diabetes UK (www.

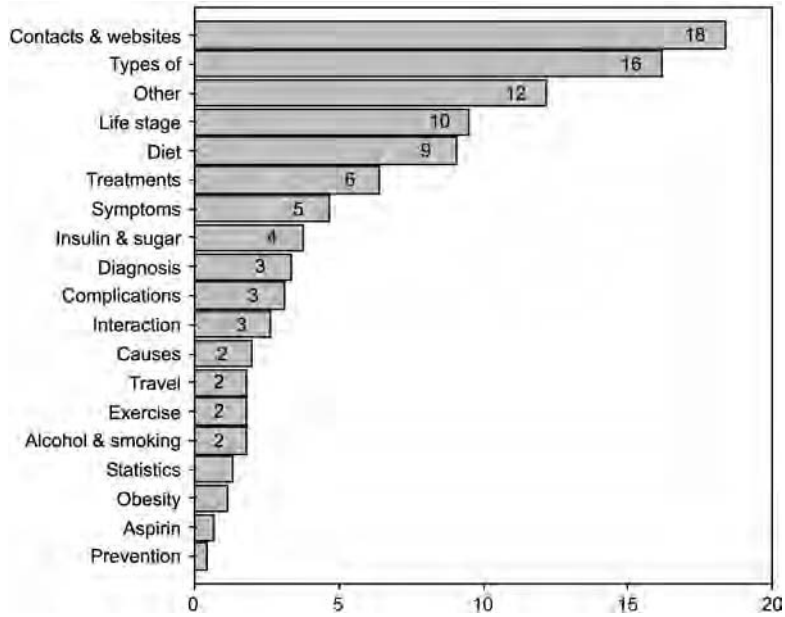


Figure 1.
The percentage of unique user search expressions falling into 19 search categories

diabetes.org.uk). The menu items relating to diabetes on these sites (as at 25 November 2004) are given in Table III.

Diabetes as a menu item was not available from the home page of the NHS Direct web site but this page was accessed by going to a "Health Encyclopaedia" link and then to diabetes. The NHS Direct diabetes homepage included links to ten topics. These were Introduction, Symptoms, Causes, Diagnosis, Treatment, Complications, Prevention, Selected links, Info Partners and Audio Clips. The coverage of likely searches by these menu options can be estimated by comparing matches to the broad 19 search categories identified above. Six of the menu links – Symptoms, Causes, Diagnosis, Treatment, Complications and Prevention – can be related to user search expressions. The six NHS Direct menu topics, however, only meet about

NHS Direct Online	BBC	Diabetes UK (homepage)
Introduction	About diabetes	What is diabetes?
Symptoms	Treatment	Managing diabetes
Causes	Living with diabetes	How we help
Diagnosis	Prevention approach	Get involved
Treatment	Links and organisations	Diabetes research
Complications		For healthcare professionals
Prevention		
Selected links		
Info partners		
Audio clips		

Table III.
Diabetes home pages of three health sites

a third of likely diabetes type of enquiries, as identified in Figure 1. The percentage of search expressions accounted for by each menu item was: symptoms 5 per cent, causes 2 per cent, diagnosis 3 per cent, treatments 6 per cent, complications 3 per cent, prevention 1 per cent, and selected links 18 per cent. The NHS DO diabetes menu covers about 38 per cent of user subject search expressions and does not give an obvious indication to content, for example, related to types of diabetes, diabetes and life stage and diet.

Diabetes as a menu item did not feature on the BBC home page nor did diabetes feature as a menu option from the health home page. A diabetes menu section was found by clicking on the A-Z index then scrolling down to health and then to diabetes. The diabetes home page included five menu items to content: About diabetes; Treatment; Living with diabetes; Prevention approach; and Links & Organisations. In comparing these five items, which in themselves are fairly broad, to the 19 search subject categories, the menu covers about 38 per cent of search queries. Again there are no obvious links to content such as types of diabetes, life stage and diet.

Diabetes UK had six menu options on the home page these were: What is diabetes; Managing diabetes; How we help; Get involved; Diabetes research and For healthcare professionals. Again these menu options are particularly wide and do not appear to address the directed subjects that users are likely to go to a diabetes site to find. There was no obvious link to types of diabetes, life stage and diabetes and diet. Perhaps site designers are less inclined to state the obvious.

Figure 1 classified approximately 300 unique user search expressions into 19 subject categories, these categories represent broad subject topics that users search for. These categories were then compared to diabetes menu options actually offered to users by three health sites. In each case the menu items, that direct users to content, covered less than half, by expression usage, of the topics searched for by users. None of the three sites, for example, addressed the user information needed to know about types of diabetes, diabetes related to a life stage and diabetes and diet, though 35 per cent of user search expressions were directed to looking for content on just these three topics. There were additional topics such as travel and diabetes, exercise, smoking and obesity that none of the sites directly covered in the opening diabetes menu. Users with a directed information need are likely to question whether any of these sites could meet their information need in these areas. Menu items link user information needs to content that meets this need and a poor menu is likely to result in users feeling confused and frustrated as to how the site could meet their need. Recent research (Nicholas and Huntington (2005)) found that over one-third (38 per cent) of users of health sites said they "often" left a site if they could not see what they were looking for on the page landed at, further a half (47 per cent) said that they sometimes did this. Clearly the cost in terms of site usage of a poorly designed menu is high.

Using the 19 subject categories based on user search expressions (Figure 1) as a starting point it was decided to construct a six-point menu that attempted to cover most of these categories. The six menu items are listed in Table IV along with the likely percentage of user queries covered.

This menu reflects generally what users seeking for diabetes information are looking for. It is based on actual search expressions entered by users and reflects the information needs of users.

Conclusion

Search logs and server transaction logs provide a real time insight into actual user interrogation of searching and navigating to site content. Here we have attempted to show how search logs can be used to monitor and assess item selection for navigational menus. Menus are of particular importance as they are indicators of content and give users an idea of how likely their query will be answered by the site. Menus are enabling and are the bridge or link between user information need and supplier content. It is where demand intention meets supply intention and getting it wrong interrupts the users' flow through to content. Previous studies (Nicholas and Huntington (2005)) on users have shown that over one-third (38 per cent) of users of health sites said they "often" left a site if they could not see what they were looking for on the page landed at and 56 per cent "often" left if they could not see what they were looking for within the first two screens looked at. Providing an adequate menu in part addresses the issue and user concern for effective indicators to content. In any event menu and menu changes should be assessed and the team has argued elsewhere (Huntington *et al.*, 2004; Nicholas *et al.*, 2002) that an effective way to monitor the impact of a change of menu, or any site change, is to analyse actual usage statistics available from the server log. This procedure dubbed "change and see" monitors for small changes in actual user navigational and usage patterns resulting from small changes in menu and site design. Ultimately menu and site changes should aim to increase traffic and direct users effectively to content and this is best assessed by examining actual user behaviour. We hope in the future to explore all the related issues raised in this paper by further examining if usage patterns do, as predicted, increase as a result of applying the results of analysing search expressions, as a form of revealing information need, to develop user-centred menus.

This study examined search logs as a form of revealing information need and proposes that this can be used as a base to construct a user centred menu. It examined 384 user unique search expressions related to diabetes. A total of 44 per cent had just typed in the single word diabetes and these searches were not particularly directed, while 56 per cent had typed in a directed search expression. An analysis of the directed searches showed that 84 per cent of these could be classified into one of 18 subject categories. These subject categories were compared to the menu coverage of three sites providing health information on diabetes: NHS Direct Online, BBC Health and Diabetes UK. In each case it was found that the site diabetes menu page covered only about one-third of searches as indicated by the subject category analysis of actual user

Table IV.
Six-point menu structure based on search expressions entered by users

Menu items	Percentage of queries covered
Introduction	
Types of diabetes, causes, diabetes and pregnancy, children and by age	30
Symptoms, diagnosis and treatments	14
Diabetes and lifestyle and interactions (diet, travel, interactions, exercise, alcohol and smoking, aspirin)	20
Staying fit and well with diabetes (prevention, complications, insulin and sugar)	8
Contacts and web sites	18

searches. Finally, a suggested six-item menu based on reflecting the 18 subject categories was proposed.

References

D'Alessandro, M.P., D'Alessandro, D.M., Galvin, J.R. and Erkonen, W.E. (1998), "Evaluating overall usage of a digital health sciences library", *Bulletin of the Medical Library Association*, Vol. 86 No. 4, pp. 602-9.

Ebenezer, C. (2003), "Usability evaluation of an NHS library web site", *Health Information and Libraries Journal*, Vol. 20 No. 3, pp. 134-42.

Eisenberg, B. (2005), "Prioritize usability testing and web analytics", available at: www.clickz.com/experts/crm/traffic/article.php/3483671 (accessed 11 April 2005).

Frick, T., Kisling, E., Cai, W., Min Yu, B., Giles, F. and Brown, J.P. (2004), "Impact of navigational models on task completion in web-based information systems", AECT 1999 presentation to the Research and Theory Division: paper no. 439, available at: http://education.indiana.edu/~frick/aect99/rtd439.html (accessed 15 January 2004).

Goodrum, A. and Spink, A. (1999), "Visual information seeking: a study of image queries on the world wide web", *Proceedings of the 62nd Annual Meeting of the American Society for Information Science, Washington, DC, October*, pp. 665-74.

Hennig, N. (1998), "Going forward: usability testing the web site", *Internet Librarian*, 4 November, available at: www.hennigweb.com/presentations/il98/ (accessed 5 March 2001).

Huntington, P., Nicholas, D. and Warren, D. (2004), "Digital visibility and its impact upon online usage: case study a health web site", *Libri*, Vol. 54 No. 4, pp. 158-68.

Jansen, B.J., Spink, A. and Saracevic, T. (2000), "Real life, real users and real needs: a study and analysis of users' queries on the web", *Information Processing and Management*, Vol. 36 No. 2, pp. 207-27.

Neilsen, J. (2002), "Web usability for senior citizens: 46 design guidelines based on usability studies with people age 65 and older", available at: www.useit.com/alertbox/20020428.html (accessed 11 April 2005).

Nicholas, D. and Huntington, P. (2005), *Digital Health Information Consumers and the BBC Website (bbc.co.uk): Users and Usage of Non-dedicated Health Sites*, UCL, London.

Nicholas, D., Huntington, P., Williams, P. and Dobrowolski, T. (2004), "Re-appraising information seeking behaviour in a digital environment: bouncers, checkers, returnees and the like", *Journal of Documentation*, Vol. 60 No. 1, pp. 24-39.

Nicholas, D., Huntington, P., Williams, P. and Gunter, B. (2002), "Digital visibility: menu prominence and its impact on use of the NHS Direct information channel on Kingston Interactive Television", *Aslib Proceedings*, Vol. 54 No. 4, pp. 213-21.

Ozmutlu, S., Spink, A. and Ozmutlu, H.C. (2004), "A day in the life of web searching: an exploratory study", *Information Processing and Management: An International Journal*, Vol. 40 No. 2, pp. 319-45.

Preece, J. (2000), *Online Communities: Designing Usability, Supporting Sociability*, Wiley & Sons, New York, NY.

Rozic-Hristovski, A., Todorovski, L. and Hristovski, D. (1999), "Developing a medical library website at the University of Ljubljana, Slovenia", *Program*, Vol. 33 No. 4, October, pp. 313-25.

Van der Geest, T. (2002), "Evaluating a web site with server data", *Document Design*, Vol. 1 No. 2, pp. 131-2.

Wolfram, D., Spink, A., Jansen, B.J. and Saracevic, T. (2001), "*Vox populi*: the public searching of the web, available at: http://jimjansen.tripod.com/academic/pubs/jasist2001/jasist2001a.html (accessed 12 April 2005).

Zhang, G.Q., Shen, G., Staiger, J., Troy, A. and Sun, J. (2004), "FcAWN: concept analysis as a formal method for automated web-menu design file", available at: http://newton.eecs.cwru.edu/~gqz/papers/web-menu.pdf (accessed 24 November 2004).

Corresponding author
Paul Huntington can be contacted at: p.huntington@ucl.ac.uk

Triangulating qualitative research and computer transaction logs in health information studies

Pete Williams
School of Library, Archive, and Information Studies, University College London, London, UK, and
Barrie Gunter
Centre for Mass Communications Research, University of Leicester, Leicester, UK

Received 28 April 2005
Revised 14 August 2005
Accepted 22 August 2005

Abstract

Purpose – The aim of this paper is to outline a triangulated methodology for studying usage of electronic health information systems which combines the quantitative data accrued from computer logs with qualitative data from in-depth interviews and observation.

Design/methodology/approach – The appropriate methods and inherent issues are reviewed from the literature, with an emphasis on qualitative research. The work of the authors is then highlighted, showing how qualitative methods can inform log analysis.

Findings – The paper suggests from the review that it is not only possible but also extremely fruitful to combine quantitative and qualitative data to interpret user behaviour.

Originality/value – The methods used by the group, known as "deep log analysis", are innovative, and the attempt both to discuss these and to provide concrete examples from this research provides its originality.

Keywords Qualitative research, Interviews, Data analysis

Paper type Research paper

Introduction

The Ciber research group at University College London are undertaking a series of studies of the impact of new technology on various strategic groups. The first of these looked at the online news media (Nicholas, Huntington, Williams, Lievesley, Dobrowolski and Withey, 1999; Nicholas, Williams, Martin and Cole, 1999; Williams and Nicholas, 1999). Other research has looked at the Internet in an educational context (Williams, 1999, 2003), with current work (Gunter *et al.*, 2001, 2003; Nicholas, Huntington, Williams and Blackburn, 2001, Nicholas, Williams and Huntington, 2001, Nicholas *et al.*, 2003; Williams *et al.*, 2001, 2003) concentrating on various health information systems and their use by members of the public (i.e. the "health consumer"). The early research of the group concentrated on in-depth interviews – over 100 journalists, for example, were interviewed as part of the media research. Web log statistics were first used by the group in a study of internet newspaper readership (Nicholas , Huntington, Williams, Lievesley, Dobrowolski and Withey, 1999), and have been used in the studies cited above to build up a picture of global

Aslib Proceedings: New Information Perspectives
Vol. 58 No. 1/2, 2006
pp. 129-139

0001-253X
DOI 10.1108/00012530610648725

usage patterns. This article describes how interviews can enhance and enrich log data, and discusses the theoretical frameworks within which the group undertakes its work.

Computer transaction logs

Transaction logs of web or other electronic information systems are automatically generated records of user activity, and are usually taken from the server side of the interaction. There is much activity, apart from the writers' own research, currently in the field of information science in capturing and analysing such logs. Pomfrett (1999) studied records of transactions from one database, while Zawitz (1998) studied one particular web site – external use of the Bureau of Justice Statistics site. Search engines can also be monitored (Jansen *et al.*, 1998) to make generalisations about search strategies. It is also possible, however, to track web use from the client terminal. Choo *et al.* (1998) installed a "Web tracker" application on searchers' computers, which was activated whenever a browser was opened. It recorded use of browser commands ("back", "forward", "add bookmark", etc.), and URLs of pages requested.

While log data are useful in themselves, they do not give us any indication of reasons for the patterns elicited, nor do they, of course, suggest why non-users elect to decline to employ various systems. Most of the research does recognise this. Taking the above examples, Choo *et al.* (1998), interviewed users of those machines from which they extracted log data, and Pomfrett (1999) and her team also interviewed users of bibliographic databases they were monitoring. Zawitz (1998) declared that "web activity statistics cannot answer all of our questions about website use" and went on to add that, "additional sources of information in combination with web statistics are needed to provide a more comprehensive portrait of the effectiveness of our websites". In short, log results suggest areas of further research, better carried out by *ad hoc* survey or qualitative methodologies. In addition to yielding purely descriptive data, that can represent useful profiles of real-time user activity, logs could be said to generate, at best, tentative hypotheses to be tested using survey or qualitative methods.

Qualitative methods

As Silverman (1999, p. 8) states, "the methods used by qualitative researchers exemplify a common belief that they can provide a "deeper" understanding of social phenomena than would be obtained by purely quantitative data". It may be considered unusual for interviews to be used in the second stage of data gathering, although they are often used to enrich material from questionnaire responses. This is because of the general inductive nature which is the norm in qualitative data gathering, where interviews and observation are used to tentatively explore a phenomenon before any formal hypothesis-forming stage is reached (Hammersley, 1992). Often interviews study a particular happening in its infancy (Marshall, 1997). Glasser and Strauss (1967) coined the term "grounded theory", where data are gathered by a number of qualitative methods, and subject to step-by-step organisation, producing finally a number of conceptual categories and the emergence of hypotheses "grounded" in the data, rather than pre-conceived by the researchers.

In the context of online information provision, the territory is not entirely virgin, as we have a good idea of the "information landscape" by virtue of the log statistics, and can thus also test tentative hypotheses. However, qualitative data can only explore and make suggestions, informing, along with the logs, survey data gathering. These explorations are, in the case of research into the use/non-use of information systems, concerned not only with simple questions about the use of and opinions about the particular system or service, but also with the full context in which an information need arises and is met. Moore (2000, p. 121, original emphasis) claims that "quantitative data is very good at telling you what is happening, qualitative provides an insight into the deeper question of *why?*"

Qualitative research is particularly appropriate in the study of information needs in the wider context of people's environments and situations. With an expertise in log analysis, it would be easy to concern ourselves solely with the readily available and real-time statistical data which can be accrued from logs. To a limited extent, some of the contextual issues can be examined using survey methodology. These include issues related to ease of use, reason for system use and so on. In other words, data that reveal, through self-reports, the inner thinking of online service users.

Previous work by the authors, for example, has examined "health outcomes", where questionnaire responses have indicated that more than one in four people said that web-based information proved an alternative to seeing the doctor (Nicholas, Huntington, Williams and Blackburn, 2001a). Only in-depth interviewing, however, can really examine the relationship and interplay between the "user", the information, the condition and the doctor. Questions that can be asked one-to-one include details on the situation of the information seeker (it may be that the person had little inclination anyway to visit the doctor), the type of condition suffered and its history, prior information or knowledge on the subject, experiences of past visits to the doctor, and perceptions about outcomes. The power of triangulated log and self-report data can reside in the ability to discover whether, for example, people who display a stronger inclination to substitute online information about health and medical conditions for face-to-face consultations differ in their patterns of online information searching.

Clearly, a tremendous amount of rich and contextual data can be gathered by talking to people about their experiences. Rogers and Bouey (1996) point out that the most utilised data collection method in qualitative research study is the interview. Curasi (2001, p. 362) summarises neatly the essence of interviewing specifically and qualitative work in general:

> In-depth interviewing is administered to better understand the experiences, opinions and interpretations of characteristics of a phenomenon. Rather than attempting to grasp the quantities or measurements of the phenomenon, these methods are used to develop a better understanding of how (people) interpret and experience some situation, process or event.

Qualitative research is ideally suited to the "user studies" arm of information research. Savolainen (1992) noted that information studies can be approached from "two opposite directions, that of production or that of utilisation of information". Research starting from the "production" end of the chain has been described as an "intermediary" (Savolainen, 1992) or "system"-centred approach (Sugar, 1995). Looking from the viewpoint of the information user, that is, from the seeking (whether active or passive) and utilisation end of the process, is termed a "user-centred"

approach. In the system's world, users are passive receivers of objective information (Sugar, 1995). In contrast, user-centred approaches study services and systems as seen by users rather than service providers or information creators/gatherers. It asks "how they define needs in different situations, how they present these needs to systems, and how they use what the system offers them" (Sugar, 1995). The user-centred model of sense-making (Dervin *et al.*, 1982, Dervin and Nilan, 1986) goes so far as to posit the user not only as an active player, but as actually constructing information (Shields and Dervin, 1993).

Dervin is one of a number of commentators who point out that use of an information system does not necessarily mean either that it was the only or first choice for information searching, or that it was of any help. She complains that many information studies fall into the trap of repeating "the failures of surveys and studies . . . [which do not acknowledge] the gap between how administrators/experts describe users and publics and the realities of what users and publics [really] think and do", (Dervin, 1999, p. 37). In keeping with this, the author(s) wish to avoid the common mistake of discussing the system with (only) the users and asking simply if the information was easy to find, useful, and so on. A more fruitful approach may be to establish what the real needs are of those for whom the information and system was devised, and to explore how these are being or can be met. The log statistics, clearly, do provide an indication of content popularity, and this can provide a useful reflection of differences in information-seeking behaviour in different environments (the group has carried out comparative log analyses of health kiosks, for example, in medical and non-medical locations). However, qualitative research is needed to inform the debate about appropriate information content and presentation.

Previous research by the group, particularly work on the use and impact of the Internet in the media (Nicholas, Williams, Martin and Cole, 1999) has confirmed previous assertions (e.g. Nicholas, 1996, 2000; Wilson, 1990) that often the perspectives of information professionals and researchers are not those of the user community. Dervin (1992, p. 63) has consistently gone further, asserting that even "user-oriented" studies are "predicated on the idea that the system is the essential order and the person/user bends to it rather than the other way round". Audunson (1999) points out that even researchers who eschew the "systems" approach still predicate their work on a means-end view. This, according to Audunson (1999, p. 71), "makes it difficult to explain anomolies, where information use seems to have no instrumental use and value". He cites a 1989 publication by March and Olsen who found, according to Audunson (1999, p. 71), that "managers and other decision makers frequently ask for information without using it, (and) still ask for more". The implication of this with regard to computer logs is clear - accessing pages is no guarantee of resulting benefit or even use.

It is vital, therefore, that in researching user needs, that is, those for whom the service - be it an information kiosk, web site or digital television initiative – is targeted must be free to describe their situations, needs and experiences, "in terms of their own frames of reference . . . on their own terms" (May, 1997, p. 68). Only in this way can such idiosyncrasies, anomalies and irrationalities be explored. In order for this to be realised, the authors employ a loose semi-structured interview format. Such interviews "allow people to answer more on their own terms than the standardised interview

permits, but still provide a greater structure for comparability over that of the focused [unstructured] interview" (May, 1997, p. 111). They still, however, allow the researcher to work with an open-ended interview framework. This is important in the work of the authors, as much of the qualitative work concerns attempts to interpret data accrued from the log analysis. There is a similarity between this approach and that employed by ethnographers (see, for example, Soloman (1997), for an ethnographic study undertaken in an information science context). In ethnography, interviews tend to be used to follow up on observations, which in turn, can vary in terms of whether the researcher is a participant or non-participant in the situation being observed. In the case of "deep log analysis", the "observation" is the log record.

The emphasis the authors place on the views and perspectives of the interview participants themselves also lends itself to the current movement in information science research to examine information seeking in the "roundness" (Soloman, 1999) or context of people's daily lives and problems. It is incumbent on any research attempting to examine how information technology is integrated (or not, as the case may be) into the lives and natural information seeking habits of "ordinary people", as the group's work tries to do, to recognise the informal, often haphazard ways in which information is accumulated. Savolainen (1995), for example, has studied information seeking "in everyday life". His recent work (Savolainen, 1999) describes how the internet fits into the patterns of people's informal information seeking; and Spink *et al.* (1997) and Chatman (1992) have examined the information needs of various "ordinary" people, such as retired women, low income families, etc.

The qualitative research of the Ciber group

The qualitative research the group are interested in generally looks at:

- How and why an information need arises, and if it is seen as such in the mind of the respondent.
- How the respondent begins solving the problem, and the role which information plays in this process.
- Where the system chosen for study fits into the array of sources consulted (be they informal/formal, professional or lay users).
- The degree to which it is used, and for what purpose.
- The factors in play resulting in use of electronic systems – privacy, convenience, accessibility, etc.
- Any associated problems with using electronic sources.
- Opinions about the presentation, content, functionality, and ease of use.
- Perceptions of the system as a whole and its appeal.
- The relevance and usefulness of the information and the extent to which it answers the question or solves the problem.
- Any issues of non-use and reasons (if any) for this.
- Any resulting behaviour changes or other longer term impact (in the health field this might include dietary, level of re-consultation, etc.).

Clearly, some of these issues can be approached by recourse to survey methodology – behaviour change, for example. Self-reporting of behaviour can be done via pre-coded, fixed response option questions in large scale surveys or probed in greater detail in-depth interviews which can be run one on one or with focus groups. The latter used mainly open-ended responding, whereas surveys used self-report closed-ended responding. Open-ended responding permits respondents the option of articulating their views with relatively few constraints on the nature of their responding. Self-reported closed-ended questions are based on the researcher's best judgement about the answers most likely to be given.

The areas mentioned above cover a great deal of the ground opened up by log analysis and can help to interpret use of particular pages and topics. With the case of kiosks and digital television systems these data can be cross-referenced by various socio-economic and other demographic variables. However, there are also specific points that can be addressed in interviews that arise specifically from peculiarities or specificities of the data. One example data logs have revealed from our current work investigating touch-screen health information systems placed in medical and non-medical locations, is that there appears to be a low take up of use among the 55-74 year age group (Nicholas *et al.*, 2000, 2004). This may, of course, be due to several factors. For present purposes, what is more instructive is that even when the elderly use kiosks they tend to look at fewer pages than other users, so that the majority hardly ever go beyond the three or four menu screens that need to be negotiated to arrive at an information page.

This finding raises some important questions that require answering through survey work and further research. It is important to say, however, that the idea is not to ask respondents simplistic questions such as (taking the first point above) "why don't you look at more pages?" First, of course, the person being interviewed may not be one of those people who look only at the menu screens. Second, asking someone to justify a particular activity of which they may not even have been aware runs the risk of yielding poor quality data. In keeping with the approach of conducting user-led interviews, respondents were asked exactly the kinds of questions outlined above, namely, those concerning information needs in general, how the system fitted into their array of information-seeking activities, and how the information they found online had helped them. Answers helped reveal why kiosks are under-used by some age groups, and why some users do not access sufficient pages to arrive at any personally beneficial information. People said, for example, they do not use the system much because it requires them to remain on their feet for a long time, because they keep activating unwanted pages due to poor hand-eye co-ordination. It is common for people to lose their manual dexterity with age (Hoot and Hayslip, 1983; Williamson *et al.*, 1997), making it difficult to use even a touch-screen. Some commentators (e.g. Demiris *et al.*, 2001) have attempted to design web pages bearing in mind users' functional impairments and inexperience with computers. There is also an issue of feeling under time pressure when searching online technology platforms in public locations where other users may be waiting their turn.

Another interesting finding is that children are by far the biggest users of health information kiosks (Nicholas, Williams and Huntington, 2001b). Interviews with health professionals and practice managers revealed not only a stark contrast

between the views about children using the kiosk of doctors and nurses, but also between medical practitioners generally and managers. More importantly, perhaps, was that no one really had any idea of the extent of children's usage or the reasoning behind it. Armed with log data showing the extent of use, the researchers realised the necessity of interviewing parents with children[1]. Again, the emphasis was on the contextual aspect of system use – the information need, the interplay, in this case – between parent and child, the benefit gained by children of using the kiosk. Without the log information, this part of the research would not have had such a high priority.

As a final example, log data have shown that people seek different topics at the various kiosk locations – more surgical pages are viewed in hospitals while more general pages are consulted in information, community centres and supermarkets. While this might be intuitive, it is worth exploring in depth. What factors are in play here – is it that the health professionals advise patients what topics to search for in medical locations? Or are they simply more focused on a specific problem or condition? Is there a cultural factor involved in seeking less condition-specific pages at non-medical locations? Perhaps it is not considered appropriate to call up pages on surgical operations during the middle of a shopping outing. It is only by sitting for a considerable time with people and going over with them their situations, motivations, needs, perceptions and attitudes that these issues can be fully examined.

There were other differences between platforms, e.g. kiosks versus internet versus DiTV, in terms of types of content searched and most popularly searched topics which beg questions about how different information technologies may be perceived by users and about the social contexts of their use. These cannot be answered by log data, but are clearly signalled by such data. Follow-up interviews are needed to explore online health users' perceptions of different online health information sources and technology platforms to understand why such differences occur. First, however, more log analyses are needed to confirm the early findings of the 2000-2003 research with these platforms (see, for example, Nicholas *et al.*, 2000, 2003, 2004; Huntington *et al.*, 2002; Williams *et al.*, 2003; Gunter *et al.*, 2003).

Conclusion

It is important to note first, in conclusion, that in communications research (as well as information research) there has historically been an antagonism or a tension between social scientific research using quantitative empirical methods based on numerical representation and measurement of social phenomena and research with a hermeneutics/humanities basis that prefers qualitative, descriptive methods. Increasingly, communications scholars are coming to acknowledge that both perspectives have something to offer. Operational definitions of phenomena based on standardised measures allow one to explore broad patterns of behaviour over time and to ascertain their causes or explanations. Such approaches allow one to model behaviour, derive explanations for it, and to predict future performance. In order to achieve such results, however, researchers must intervene with naturally occurring phenomena and impose artificial measurement regimes upon them. Those approaches that give prominence to non-interventionist methods such as observation and analysis of conversations (e.g. focus groups and open-ended depth interviews) focus on

naturally occurring phenomena but do not possess the tools to uncover cause-effect relations or to derive rules of behaviour that are needed to predict future performance. A triangulated approach recognises the strength of combining closed and open-ended forms of data analysis.

Log data permit non-interventionist analysis or real-time behaviour, but represent behaviour in a limited fashion. Log data are descriptive of behaviour but do not provide explanations for behaviour patterns. Such data can raise questions about behaviour patterns, however, that can be examined further by utilising other methods – both quantitative and qualitative. Self-report data can be obtained from users of online information systems that may be structured or unstructured and represented numerically or verbally. Numerical representation of self-report data that deal with personal characteristics or opinions, beliefs, motives, interests and factual knowledge can be combined with essentially quantitative log data on a large population scale to model the behaviour of different online user sub-groups. Verbally represented self-report data allow online service users to express their experiences in their own terms and to discuss these experiences in much more detail than standardised questionnaire measures normally permit. While such verbal data cannot be readily integrated into quantitative modelling exercises, they can enhance our understanding of specific barriers and problems online service users have identified in relation to specific services aspects and applications. Hence they provide a parallel data stream alongside quantitative data yielding a more comprehensive portrait of online service usage and of the nature of the user population.

Note

1. Children themselves were not interviewed, as the original ethical permission granted by the Multi-centred Research Ethics Committee to which the researchers applied, did not include the involvement of children.

References

Audunson, R. (1999), "Can institutional theory contribute to our understanding of information seeking behaviour?", in Wilson, T.D. and Allen, D.K. (Eds), *Exploring the Contexts of Information Behaviour: Proceedings of the 2nd International Conference on Research in Information Needs, Seeking and Use in Different Contexts, Sheffield University, 13-15 August, 1998*, Taylor Graham, London, pp. 67-81.

Chatman, F. (1992), *The Information World of Retired Women*, Greenwood Press, London.

Choo, C.W., Detlor, B. and Turnbull, D. (1998), "A behavioral model of information seeking on the web – preliminary results of a study of how managers and IT specialists use the web", paper presented at 1998 ASIS Annual Meeting, available at: http://choo.fis.utoronto.ca/fis/respub/asis98/ (accessed 21 March 2002).

Curasi, C.F. (2001), "A critical exploration of face-to-face interviewing vs computer-mediated interviewing", *International Journal of Market Research*, Vol. 43 No. 4, pp. 361-75.

Demiris, G., Finkelstein, S.M. and Speedie, S.M. (2001), "Considerations for the design of a web-based clinical monitoring and educational system for elderly patients", *Journal of the American Medical Informatics Association*, Vol. 8 No. 5, pp. 468-72.

Dervin, B. (1992), "From the mind's eye of the user: the sense-making qualitative-quantitative methodology", in Glazier, J.D. and Powell, R.R. (Eds), *Qualitative Research in Information Management*, Libraries Unlimited, Englewood, CO, pp. 61-84.

Dervin, B. (1999), "Chaos, order, and sense-making: a proposed theory for information design", in Jacobson, R. (Ed.), *Information Design*, MIT Press, Cambridge, MA, pp. 35-57.

Dervin, B. and Nilan, M. (1986), "Information needs and uses", in Williams, M. (Ed.), *Annual Review of Information Science and Technology*, Vol. 21, Knowledge Industry Publications, White Plains, NY.

Dervin, B., Jacobson, T. and Nilan, M. (1982), "Measuring aspects of information seeking: a test of a quantitative/qualitative methodology", in Burgoon, M. (Ed.), *Communication Yearbook 6*, Sage, Beverly Hills, CA, pp. 419-45.

Glasser, B. and Strauss, A. (1967), *The Discovery of Grounded Theory*, Aldine Press, Chicago, IL.

Gunter, G., Huntington, P., Williams, P. and Nicholas, D. (2001), "Health advice on the TV: early opinions of users", *Online and CD Notes*, Vol. 14 No. 9, pp. 4-8.

Gunter, B., Nicholas, D., Huntington, P. and Williams, P. (2003), "Digital interactive television: health information platform for the future", *Aslib Proceedings*, Vol. 55 Nos 1/2, pp. 346-56.

Hammersley, M. (1992), *What's Wrong with Ethnography? Methodological Explorations*, Routledge, London.

Hoot, J.L. and Hayslip, B. (1983), "Microcomputers and the elderly: new directions for self sufficiency and lifelong learning", *Educational Gerontology*, Vol. 9 Nos 5/6, pp. 493-9.

Huntington, P., Nicholas, D., Williams, P. and Gunter, B. (2002), "Characterising the health information consumer: an examination of the health information sources used by digital television users", *Libri*, Vol. 52 No. 1, pp. 16-27.

Jansen, B.J., Spink, A., Bateman, J. and Saracevic, T. (1998), "Real life information retrieval: a study of user queries on the Web", *SIGIR Forum*, Vol. 32 No. 1, pp. 5-17.

Marshall, P. (1997), *Research Methods: How to Design and Conduct a Successful Project*, How To Books, Plymouth.

May, T. (1997), *Social Research: Issues, Methods, Progress*, OUP, Buckingham.

Moore, N. (2000), *How to Do Research*, 3rd ed., Library Association, London.

Nicholas, D. (1996), "Information systems versus information users", *Managing Information*, Vol. 3 No. 6, pp. 26-33.

Nicholas, D. (2000), *Assessing Information Needs: Tools, Techniques and Concepts for the Information Age*, Aslib, London.

Nicholas, D., Huntington, P. and Williams, P. (2003), "Three years of digital consumer health information: a longitudinal study the touch screen health kiosk", *Information Processing and Management*, Vol. 39, May, pp. 479-502.

Nicholas, D., Huntington, P. and Williams, P. (2004), *Digital Consumer Health Information and Advisory Services in the UK: A User Evaluation and Sourcebook*, City University, London, six-volume report submitted to the Department of Health, March.

Nicholas, D., Williams, P. and Huntington, P. (2000), "Digital health information: case study, the information kiosk", *Aslib Proceedings*, Vol. 52 No. 9, pp. 315-30.

Nicholas, D., Williams, P. and Huntington, P. (2001), "Health information kiosk use in health organisations: the views of the health professionals", *Aslib Proceedings*, Vol. 53 No. 9, pp. 368-86.

Nicholas, D., Huntington, P., Williams, P. and Blackburn, P. (2001), "Digital health information and health outcomes", *Journal of Information Science*, Vol. 27 No. 4, pp. 265-76.

Nicholas, D., Williams, P., Martin, H. and Cole, P. (1999), *The Media and the Internet*, Aslib, London.

Nicholas, D., Huntington, P., Williams, P., Lievesley, N., Dobrowolski, T. and Withey, R. (1999), "Developing and testing methods to determine the use of web sites: case study newspapers", *Aslib Proceedings*, Vol. 51 No. 5, pp. 144-54.

Pomfrett, S. (1999), "Types of electronic journal users", paper presented at the SuperJournal Conference, Birkbeck College, London, 21 April, available at: http://irwell.mimas.ac.uk/sj/confpomfret.htm (accessed 5 October 2003).

Rogers, G. and Bouey, E. (1996), "Collecting your data", in Tutty, L.M., Rothery, M. and Grinnell, R.M. Jr (Eds), *Qualitative Research for Social Workers: Phases, Steps, and Tasks*, 4th ed., Allyn & Bacon, Boston, MA, pp. 50-87.

Savolainen, R. (1992), "The sense-making theory – an alternative to intermediary-centred approaches in library and information science", in Vakkari, P. and Cronin, B. (Eds), *Conceptions of Library and Information Science*, Taylor Graham, London, pp. 149-64.

Savolainen, R. (1995), "Everyday life information seeking: approaching information seeking in the context of 'way of life'", *Library and Information Science Research*, Vol. 17 No. 3, pp. 259-94.

Savolainen, R. (1999), "The role of the internet in information seeking: putting the networked services in context", *Information Processing & Management*, Vol. 35 No. 6, pp. 765-82.

Shields, V. and Dervin, B. (1993), "Sense-making in feminist social science research: a call to enlarge the methodological options of feminist studies", *Women's Studies International Forum*, Vol. 16 No. 1, pp. 65-81.

Silverman, D. (1999), *Doing Qualitative Research: A Practical Handbook*, Sage, London.

Soloman, P. (1997), "Discovering information behaviour in sense making", *Journal of the American Society for Information Science*, Vol. 48 No. 12, pp. 1097-138.

Soloman, P. (1999), "Information mosaics: patterns of action that structure", in Wilson, T.D. and Allen, D.K. (Eds), *Exploring the Contexts of Information Behaviour*, Taylor Graham, London, pp. 150-75.

Spink, A., Jaeckel, M. and Sidberry, G. (1997), "Everyday life information seeking of low income African-American households: Wynnewood Healthy Neighbourhood Project", in Schwartz, C. and Rorvig, M. (Eds), *Proceedings of the 60th Annual Meeting of the American Society for Information Science, November 1997, Washington DC*, Information Today, Medford, NJ.

Sugar, W. (1995), "User-centred perspective of information retrieval research and analysis methods", in Williams, M. (Ed.), *Annual Review of Information Science and Technology*, Vol. 30, Information Today, Medford, NJ, pp. 77-109.

Williams, P. (1999), "Net generation: the experiences, attitudes and behaviour of children using the internet for their own purposes", *Aslib Proceedings*, Vol. 50 No. 9, pp. 315-22.

Williams, P. and Nicholas, D. (1999), "The migration of the news to the net", *Aslib Proceedings*, Vol. 51 No. 4, pp. 122-34.

Williams, P., Nicholas, D. and Huntington, P. (2001), "Walk in to (digital) health information: the introduction of a digital health information system at an NHS walk-in centre", *Online and CD Notes*, Vol. 14 No. 2, pp. 4-7.

Williams, P., Nicholas, D. and Huntington, P. (2003), "Home electronic health information for the consumer: user evaluation of a DiTV video-on-demand service", *Aslib Proceedings*, Vol. 55 Nos 1/2, pp. 64-74.

Williamson, K., Bow, A. and Wale, K. (1997), "Older people and the internet", *Link-Up*, March, pp. 9-12.

Wilson, T. (1990), "Object or participant? The information user in information research", *Communications from the Centre for Library Research*, Vol. 3, special issue: Swedish library research, pp. 5-15.

Zawitz, M. (1998), *Web Statistics – Measuring User Activity*, available at: www.ojp.usdoj.gov/bjs/pub/ascii/wsmua.txt (accessed 12 June 1999).

Further reading

Williams, P. (2002), "The learning Web: the development, implementation and evaluation of internet based undergraduate materials for the teaching of key skills", *Active Learning in Higher Education*, Vol. 3 No. 1, pp. 40-53.

Corresponding author

Pete Williams can be contacted at: peter.williams@ucl.ac.uk

Three problems in logic-based knowledge representation

Rob Miller
School of Library, Archive and Information Studies, University College London, London, UK

Received 28 June 2005
Revised 8 August 2005
Accepted 12 September 2005

Abstract

Purpose – The purpose of this article is to give a non-technical overview of some of the technical progress made recently on tackling three fundamental problems in the area of formal knowledge representation/artificial intelligence. These are the Frame Problem, the Ramification Problem, and the Qualification Problem. The article aims to describe the development of two logic-based languages, the Event Calculus and Modular-E, to address various aspects of these issues. The article also aims to set this work in the wider context of contemporary developments in applied logic, non-monotonic reasoning and formal theories of common sense.

Design/methodology/approach – The study applies symbolic logic to model aspects of human knowledge and reasoning.

Findings – The article finds that there are fundamental interdependencies between the three problems mentioned above. The conceptual framework shared by the Event Calculus and Modular-E is appropriate for providing principled solutions to them.

Originality/value – This article provides an overview of an important approach to dealing with three fundamental issues in artificial intelligence.

Keywords Logic, Artificial Intelligence, Knowledge capture

Paper type General review

1. Introduction and context

This article is intended to give a brief non-technical overview of some of the technical research in which I have been involved in the last ten years or so in a particular area of mathematical logic and formal knowledge representation, largely at SLAIS, in collaboration with researchers at Imperial College London, the University of Cyprus, Harvard University in Massachusetts, and the IBM T.J. Watson Research Center in New York.

The term "knowledge representation" is used to refer to very different things in different academic contexts. In the field of artificial intelligence (AI), usually viewed as a sub-field of computer science, it has come to mean the use of mathematical logic to produce formal models of mental structures of information and associated inference processes, with the longer term aim of using such models as specifications for machines able to simulate aspects of intelligent reasoning. Within this endeavour, a large amount of research effort has been spent on attempting to represent "commonsense" knowledge and reasoning formally – in other words, attempting to model the information and reasoning processes we use to go about our daily lives in our everyday environment. This has proved very difficult. It turns out, for example, to be much easier to specify how a machine could simulate reasoning about playing chess, with its well-defined rules, strategies and variables, than to specify a comprehensive, flexible and adaptive routine for making a cup of tea (including, for

Aslib Proceedings: New Information Perspectives
Vol. 58 No. 1/2, 2006
pp. 140-151

0001-253X
DOI 10.1108/00012530610648734

example, information and general planning strategies enabling you to decide what do if you have run out of tea bags and the supermarket is closed, etc.).

The idea of modelling the human reasoning process via symbolic logic can, of course, be traced back to ancient Greece, and has had various proponents throughout the ages (including, notably, the librarian, mathematician and philosopher Leibniz). But the modern revival of interest in this endeavour, spurred on by its potential practical realisation via electronic computers, is largely due to the American computer scientist and logician John McCarthy. From the 1950s to the 1980s, McCarthy wrote a number of seminal papers laying out a research agenda for logic-based AI, starting with a 1959 article called "Programs with common sense" (McCarthy, 1959). One of McCarthy's central themes has been general reasoning about cause and effect, in particular via the notion of action. This is because such concepts are essential in the representation of key human reasoning processes such as planning. In 1969 McCarthy and Hayes (1969) proposed a logical ontology and schema for representation of such knowledge called the Situation Calculus, so called because its key ontological commitment is to view the world as progressing discretely from situation to situation via the deliberate execution of actions by an agent or agents. Initial work on the Situation Calculus has since broadened into the AI sub-field of Reasoning about Action, and it is into this area that the research described in the rest of this article fits.

2. Event Calculus and the frame problem

2.1. Classical Event Calculus

The Event Calculus is a logical framework for reasoning about action similar in some respects to the Situation Calculus, but with a different ontological basis. It was first proposed by Bob Kowalski and Marek Sergot (at Imperial College London) in the form of a "logic program" (Kowalski and Sergot, 1986). Murray Shanahan (also at Imperial) then recast it in more conventional classical logic terms (see, e.g. Shanahan, 1997), and this classical formulation has been refined and extended in various ways by myself and others (see Miller and Shanahan (2002) for a technical overview). The key ontological difference between the two calculi is in the treatment of time. Whereas in the Situation Calculus time progresses only by virtue of agents performing actions, in the Event Calculus the flow of time is represented independently, and action occurrences, or "events", are explicitly embedded in this independent temporal structure. This makes the Event Calculus especially suitable for narrative reasoning, that is, representation of and reasoning about events that have actually happened or will actually happen, as opposed to the kind of hypothetical reasoning about arbitrary possible sequences of actions that was the original motivation for the Situation Calculus.

As with most logic-based schemes, an Event Calculus description of an aspect of the world (or "domain") is often referred to as a theory or axiomatisation, and individual statements within this are referred to as axioms or formulae. An Event Calculus theory comes in two parts – the domain-independent axioms which express general commonsense principles or "truths", and the domain-dependent axioms which contain specific facts about what causes what, what has happened or been observed when, etc. What does an Event Calculus domain-independent axiom look like? Here is an axiom intended to partially express the notion that things don't change without a cause:

$$\forall f t_3[\text{HoldsAt}(f t_3) \leftarrow \exists t_1(\text{HoldsAt}(f t_1) \wedge t_1 < t_3 \wedge \neg\, \exists a, t_2[\text{Happens}(a, t_2) \wedge t_1 \leq t_2 < t_3 \wedge \text{Terminates}(a, f, t_2)])]. \quad (1)$$

How can we make intuitive sense of this string of symbols? We can start by recalling the meanings of the various quantifiers, connectives and mathematical notation used within it. $\forall$ means "for any", $\leftarrow$ means "if", $\exists$ means "there exists", $\wedge$ means "and", $<$ means "less than", $\neg$ means "not", and $\leq$ means "less than or equal to". The words HoldsAt, Happens and Terminates are predicates – names for relations that hold between their respective arguments (the symbols that follow the predicate in parentheses). The single lower case letters f, t_1, etc. are variables – place holders that can be replaced by any individual entity of the right sort. HoldsAt(F,T) means "the property F is true at time T", with fluent being the technical name in the Event Calculus for a property that can be true or false at different times (such as "F"). Happens(A,T) means "the action or event A occurs at time T", and Terminates(A,F,T) means "an occurrence of event or action A at time T would cause fluent F to become false".

Putting all this together, the axiom above reads "for any fluent 'f' and time 't_3', 'f' is true at 't_3' if 'f' was true at some earlier time 't_1' and no event or action 'a' occurred at any time 't_2' between 't_1' and 't_3' that would have caused 'f' to become false". In other words, "things don't change from true to false without a cause". A twin of this axiom states that "things don't change from false to true without a cause". Together these two axioms encapsulate the Event Calculus's notion of persistence. Other axioms capture the general notion of causality in a similar style, again in terms of the predicates HoldsAt, Happens, Initiates and Terminates.

But logic does not rely on the reader's intuition for its meaning. Classical logic, from which this variant of the Event Calculus is constructed, has a rigorous notion of semantics based on the idea of a model. (This is a technical use of the word model, very different from the idea of a "representation" implied by its everyday usage. Since we are using both meanings in this article, we will distinguish the former from the latter by continuing to italicise the technical term.) Each model of a theory is an assignment of a fully specified concrete relation to each predicate used within it, such that the relations taken in combination are compatible with every axiom of the theory. The meaning of a theory is the sum total of its models, and a logical statement is entailed from a theory exactly if it is impossible to construct a model of the theory in which the statement is untrue. Much of the technical work in this research area involves reasoning about models, and we will return to this notion in later sections of this article.

Domain-dependent axioms describe particular causal or observed facts about the aspect of the world under consideration. For example, if building an axiomatisation of our everyday knowledge about cars, we might formulate the "causal law" that turning the ignition key causes the engine to run, provided there is petrol in the car and the battery is charged, as follows:

$$\forall[\text{Initiates}(\text{TurnKey}, \text{EngineRunning}, t) \leftarrow (\text{HoldsAt}(\text{Petrol}, t) \wedge \text{HoldsAt}(\text{BatteryCharged}, t))]. \quad (2)$$

and record specific observations that there was petrol in the car at time 2 and the key was turned at time 5 as:

$$\mathrm{HoldsAt}(\mathrm{Petrol}, 2) \wedge \mathrm{Happens}(\mathrm{TurnKey}, 5). \quad (3)$$

In a realistic domain, there are likely to be many hundreds of such statements.

2.2. The frame problem and default reasoning

Early work using the Situation Calculus to describe the effects of actions hit a seemingly fundamental problem – the Frame Problem. This is that, with unmodified use of classical logic, to eliminate unintended models of a theory in which actions have more effects than those explicitly specified, we would have to include axioms stating what an action does not cause as well as what it does. This problem manifests itself in the Event Calculus, and every other logical framework for reasoning about actions, as well. For example, in the context of the "car" domain above, we would have to include statements such as ¬ Initiates(TurnKey, DoorsOpen, *t*) (recall that ¬ means "not") to assert that turning the key does not open the car doors, does not open the windows, does not turn on the windscreen wipers, does not change the colour of the car, does not change the gear, etc, etc. Not only is this impractical – a theory containing hundreds of causal laws would potentially have to be expanded to a theory containing tens of thousands of axioms – but it does not appear to properly reflect the flexibility of our own reasoning processes. In particular, people seem to have the capacity to assume non-effects of actions (and other kinds of information) by default, unless and until given information to the contrary. But no such capacity is built into the rigid and precise mechanism of classical mathematical logic.

At first sight the Frame Problem might appear to be a fairly substantial nail in the coffin of the notion that logic can be used to model human reasoning, and indeed some respected academics have argued this. But in fact the Frame Problem has been a rich source of inspiration in the continuing development of modern logic, and in particular in the development of default or non-monotonic logics. Once again McCarthy (1980) helped lead the way with his idea of circumscription. Circumscription is one of several ways of expressing a preference between different models of a theory, and defining an operator that will automatically and succinctly transform the theory into a stronger theory whose models are precisely the most preferred models of the original. In the context of the Frame Problem, a particular circumscription policy can be defined so as to prefer models in which there is least potentially caused change, so that, for example, in the "car" domain a model in which turning the key does not cause the car to change colour is preferred over one that does (all other aspects of the two models being equal).

Circumscription is a powerful and versatile logical tool, but has to be applied with care to get the intended effect. Hanks and McDermott's (1987) "Yale Shooting" example domain infamously exposed a loophole in an early attempt to solve the Frame Problem using circumscription.) Murray Shanahan first applied circumscription to solve the Frame Problem in the context of the Event Calculus (see Shanahan (1997) for a comprehensive discussion of this), and we have since used it in relation to several enhancements and extensions to the core theory. For example, Miller and Shanahan (1996) how the Event Calculus can be combined with standard mathematical modelling using sets of differential equations, to represent dynamic systems displaying both discrete and continuous change. This work illustrates that when physicists, engineers

and the like reason about such systems, in terms of transitions from one set of differential equations to an- other, the default reasoning they often apply in identifying "discontinuities" in otherwise smoothly changing parameters can be expressed succinctly in terms of circumscription.

2.3. Knowledge producing and conditional actions

Much of the computationally-oriented work in AI to date on automated planning and scheduling has concentrated on the special case where everything about the environment is known beforehand. But this is often not the case in everyday life. For example, if I wish to go to the town centre, I might not know if there is a bus at a convenient time or not. So I have to express a plan for this goal as follows: go to the bus stop – have a look to see if there is a number 10 bus waiting there – if there is, get on it – if there isn't, walk. The action "Have a look to see if there is a number 10 bus" is a knowledge producing action in that it changes my knowledge of the world rather than the world itself. "If there is (a bus), get on it" is a conditional action contingent on my future knowledge.

Leora Morgenstern (at the IBM T.J. Watson Research Center in New York) and I are currently working on extending the Event Calculus, and an associated circumscriptive solution to the Frame Problem, to deal with knowledge producing and conditional actions. Preliminary investigations indicate that a promising direction is to borrow some ideas from modal logic. In particular, we are experimenting with expanding the framework's ontology to include the notion of a possible world. Each model of the theory will include a number of possible worlds other than the "real world", each with its own time line, and an accessibility relation indicating which of these possible worlds are compatible with the agent's current state of knowledge. The agent's growth in knowledge over time, as more and more knowledge producing actions are executed, is simulated by making more and more possible worlds inaccessible over time. (In other words, alternative possibilities are progressively eliminated as the agent's ignorance diminishes.) This work is largely inspired by related work by Scherl and Levesque (1993) using the Situation Calculus, but with some important and interesting differences that arise because of the different ontologies and capabilities of the two calculi.

3. Modular-E and the Ramification Problem

The Event Calculus is a specialised theory written in standard, general purpose logic. A complementary methodology, now widely used in the area of reasoning about actions, is to develop and analyse highly specialised logics called action languages (Gelfond and Lifschitz, 1998). Action languages retain the fundamental logical notions of models and entailment, but have a restricted syntax that allows the particular concepts and principles under investigation to be hard-wired into the semantics (i.e. the underlying definition of what a model is), rather than represented as axioms. Such an approach can be a useful extra step when conducting an in depth investigation of a particular issue, because it allows representational details that might otherwise obscure the central research questions to be abstracted away. The resultant logics can also be very useful for direct use as specifications for computer programs.

Tony Kakas (at the University of Cyprus), Loizos Michael (at Harvard) and I have recently developed a particular action language called Modular-E (ME) to help us study two problems intimately related to the Frame Problem (Kakas *et al.*, 2005). The first of

these is the Ramification Problem. This is the problem of adequately representing and inferring information about the knock-on and indirect effects (ramifications) that might accompany the direct effects of an action or event. Indirect effects are commonplace in the everyday world. For instance, in our "car" example, turning the key will cause the engine to run, which (if the car has no oil) will in turn cause it to overheat, which in turn will cause smoke to come out of the bonnet, etc. The Ramification Problem is related to the Frame Problem in that an over-zealous solution to the latter precludes an adequate solution to the former. For example, if our solution to the Frame Problem has allowed us to conclude (by default) that turning the key doesn't cause smoke to come out of the bonnet (because it is not an explicitly stated effect of that action), we cannot then admit this as a consequence of our knowledge of indirect effects.

Statements called propositions are ME's equivalent to Event Calculus axioms. Here are three c-propositions ("c" for "causes"), two h-propositions ("h" for "holds") and one o-proposition ("o" for "occurs") to describe some facts about the car domain:

$$\{\text{TurnsKey, Petrol, BatteryCharged}\}\ \text{causes EngineRunning} \tag{4}$$

$$\{\text{EngineRunning}, \neg\ \text{Oil}\}\ \text{causes EngineTooHot} \tag{5}$$

$$\text{EngineTooHot causes Smoke} \tag{6}$$

$$(\text{Petrol} \wedge \neg\ \text{EngineRunning})\ \text{holds} - \text{at 2} \tag{7}$$

$$(\neg\ \text{EngineTooHot} \wedge \neg\ \text{Smoke})\ \text{holds} - \text{at 3} \tag{8}$$

$$\text{TurnKey occurs} - \text{at 5.} \tag{9}$$

Because of ME's hard-wired principles of fluent persistence (part of its solution to the Frame Problem), we can infer from the above that at all times before the TurnKey action at 5 there is petrol in the car, the engine is neither running nor too hot, and there is no smoke (i.e. these will be characteristics of all models of this domain). But the propositions say nothing about whether the car battery is in fact charged or if there is oil in the car, and so we obtain four models from this domain description. In one of these models the battery is charged but there is no oil in the car, and ME's solution to the Ramification Problem ensures that in this same model there is smoke coming out of the bonnet at, for example, time 6.

ME has an exceptionally robust and complete solution to the Ramification Problem. It can be used, for example, to describe domains in which some actions or effects are non-deterministic, in which simultaneously executed actions have competing causal effects, in which particular outcomes are contingent on a particular micro-ordering of indirect effects, or in which "looping" indirect effects give rise to rapid oscillations in fluent values (e.g. to model the mechanism of an electric buzzer). Technically it achieves this flexibility by viewing all effects as micro-processes which operate in an infinitesimal time span just "long enough" to explore and resolve all potential causal conflicts and outcomes.

Work with action languages can feed back into work with mainstream general purpose logics. Jeremy Forth (at Imperial College London) and I are currently

investigating how an earlier incarnation of ME's solution to the Ramification Problem, expressed in the Language E (Kakas and Miller, 1997a, b), can be translated into Event Calculus terms. One benefit of this is that the Event Calculus will then be able to take advantage of some of the computational techniques that have been developed for E and its successor ME.

4. The Qualification Problem, elaboration tolerance and free will

The second problem that motivated the development of ME is known as the Qualification Problem. This has two major aspects, the endogenous problem and the exogenous problem, as described below.

4.1. The Endogenous Qualification Problem

Many "causal laws" come with qualifications. For example, the assertion that turning the ignition key causes a car's engine to start, as represented in equations (2) and (4) above, is qualified by the provisos that there is petrol in the car and that the battery is charged. In the case of this particular causal law, these two conditions are explicitly included within the rule itself, but it is not always desirable to express qualifications in this way. One important reason for this is that we would like our theories to be elaboration tolerant – when new information becomes available we would like to be able to add this as a new proposition or axiom without having to revisit or revise the (possibly many hundreds of) existing rules already in the representation. As an example, consider the "car" domain augmented with two more ways to start the engine, but also supposing we did not at first know about the necessity of petrol. We can record information about starting the engine with three c-propositions:

$$\{\text{TurnKey}, \text{BatteryCharged}\} \text{ causes EngineRunning} \tag{10}$$

$$\text{JumpStart causes EngineRunning} \tag{11}$$

$$\text{PushStart causes EngineRunning}. \tag{12}$$

However, now we are told that it is not possible to start the engine without petrol. We could sift through all our causal rules, and add petrol as a qualification to each of (10), (11) and (12), but ME offers a better alternative. Instead we can add this new information as a p-proposition ("p" for "prevents"):

$$\neg \text{ Petrol prevents EngineRunning}. \tag{13}$$

ME will automatically reconcile this new statement with the existing theory by inferring extra qualifications where necessary. In a similar manner, it is also able to reconcile conflicting causal rules as a form of non-determinism, and can also accommodate arbitrary logical constraints between fluents, expressed as fluent formulae within a-propositions ("a" for "always"). For example, we may also wish to record that broken cars' engines never run:

$$\text{always}(\text{Broken} \rightarrow \neg \text{ EngineRunning}). \tag{14}$$

The Endogenous Qualification Problem is the problem of assimilating and combining qualifications and logical constraints in this way, and ME offers a particularly

complete solution, because we have shown that the logic is able to reconcile any arbitrary combination of c-, p-, o-and a-propositions, provided only that the collection of fluent formulae contained in the a-propositions are not logically inconsistent. In turn, this has enabled us to articulate a notion of free will – in ME the implicit agent can attempt any action in any circumstance, and the expression of this will not lead to an inconsistency. (This is not the case in a number of other formalisms, and the lack of this property can cause technical problems, for example, when reasoning about plans.)

4.2. The Exogenous Qualification Problem

The Exogenous Qualification Problem relates to the fact that in complex practical situations, actions' direct and indirect effects will occasionally fail due to factors (qualifications) completely outside the scope of the associated Event Calculus or ME theory. This will in turn result in incoming observations contradicting the predictions given by the theory. The Exogenous Qualification Problem is the problem of reconciling such observations with the theory and appropriately assimilating them, adjusting future predictions as necessary. For instance, suppose that we add an extra h-proposition to the "car" example (4)-(9) to the effect that there is oil in the engine and the battery is charged:

$$(\text{Oil} \wedge \text{BatteryCharged}) \text{ holds} - \text{at } 2. \tag{15}$$

Propositions (4)-(9) together with (15) now eliminate all models of the domain except one in which the car is smoothly running at time 6. In other words, this domain entails the h-proposition "EngineRunning holds-at 6", so that this is a "prediction" of the theory. But now suppose that some un- usual circumstance has caused the TurnKey action to fail, at least in respect of starting the engine[1], so that we observe that the engine is not running at time 6. We therefore now wish to add " ¬ EngineRunning holds-at 6", but the problem is that this will eliminate the one remaining model, leaving an inconsistent theory.

Once again, default reasoning techniques come to the rescue. The trick is to make use of a default reasoning facility already available in ME via its inclusion of n-propositions ("n" for "normally"). An n-proposition is used to declare that a particular fluent has a particular truth value in normal circumstances. For example, to state that cars are not normally broken we could write:

$$\text{normally} \neg \text{Broken}. \tag{16}$$

This acts as an instruction to ME to prefer models in which the car is not broken to those in which it is (except in those cases where a specific cause for the car to be broken is also present in the model). To solve the exogenous qualification problem we can mechanically transform all the c-propositions in the theory as follows. For c-proposition (x) we create the fluent NormalExo(x) to symbolise all the otherwise unrepresented "normal" conditions that the rule is contingent on, and then use an n-proposition to also state that those conditions are indeed normal. For example, (4) is transformed into:

$$\{\text{TurnKey}, \text{Petrol}, \text{BatteryCharged}, \text{NormalExo}(4)\} \text{ causes EngineRunning} \tag{4a}$$

$$\text{normally NormalExo}. \tag{4b}$$

The transformed theory consisting of (4a), (4b), (5)-(9) and (15) has two models, but (4b) ensures that we prefer the model in which NormalExo(4) is true, allowing us to conclude "EngineRunning holds-at 6" as before. However, as soon as we add the observation " ¬ EngineRunning holds-at 6" to the theory we eliminate this model, leaving us with just the previously less preferred model in which the car engine is not running.

Default reasoning techniques have to be applied with care to avoid un- intended results (as famously illustrated with the "Yale Shooting Problem" mentioned in Section 2.2), and it might be expected that the various default mechanisms built into ME to solve the Frame, Ramification and Qualification Problems would interfere with each other. But this turns out not to be the case, largely because of the elaboration tolerance and "free will" properties that characterise ME, as mentioned in Section 4.1. The development of ME has once again highlighted the intimate and interesting relationships between these three fundamental problems of logic-based AI.

5. Concluding remarks

The intended focus of this article is on fundamental and technical research is- sues rather than applications. However, the Event Calculus and its action language counterparts in various forms have been applied in a number of areas of practical information technology, such as scheduling systems, databases, accident report processing and automated legal reasoning (see the introductory section of Miller and Shanahan (2002) for a comprehensive set of references). Knowledge-based systems such as these have a potential advantage over traditional computer applications because of the flexibility and transparency of their computation (e.g. allowing reasoning both from cause to effect and vice versa, or enabling output to be accompanied by an explanation of the underlying information processing). For example, Alessandra Russo (at Imperial College London), myself and others have employed an Event Calculus implementation for use in software system specification (Russo *et al.*, 2002), developing a tool able to indicate reasons why a software prototype does not meet particular requirements in its specification. The next stage of this work will be to take the more complete argumentation-based computational procedures for E and ME (Kakas *et al.*, 1999, 2001) developed in collaboration with Tony Kakas and Francesca Toni (also at Imperial) and apply them in the same context, to give us an even more versatile tool for software requirements engineering.

More generally, research in Reasoning about Actions is proving useful in a number of cutting-edge applications. To give just two examples, Sheila McIlraith (at Toronto) and her research team have explored how the Situation Calculus can be integrated with the DAML markup language to provide semantic web services (see, e.g. McIlraith *et al.*, 2001), and Chitta Baral (at Arizona State University) and his team have applied the action language A to research in molecular biology, producing a system able to reason about interaction networks in cells, in particular to advance knowledge of the causes of certain types of cancer (Tran *et al.*, 2005).

These immediate benefits in contemporary computer applications and applied research are all evidence for the fundamental and over-arching reason why the AI sub-field of Knowledge Representation, exemplified by the research described in this article, should be of interest to computer scientists, information scientists, information managers and those working in related fields. This is that formal and detailed models

of how information is organised and used inside our heads must be relevant to decisions about how we organise and use information more generally.

6. Further reading

I hope that this article has prompted at least a few readers to find out more about AI, logic-based Knowledge Representation, and perhaps even the Event Calculus and ME. Russell and Norvig's (1995) book *Artificial Intelligence, A Modern Approach* is a particularly good general introductory text on AI. For a more in-depth but still introductory treatment of logic and Knowledge Representation, see Genesereth and Nilsson's (1987) *Logical Foundations of Artificial Intelligence*. Still more specifically, Davis's (1990) *Representations of Commonsense Knowledge* gives an excellent overview of techniques and topics in Commonsense Reasoning. It is always worth revisiting the influential and inspiring papers of McCarthy, and a fine selection of these, spanning three decades and including the three specific McCarthy papers cited in this article, is gathered together in Lifschitz (1990). Another more contemporary collection of influential research papers is contained in Hobbs and Moore (1985), including Pat Hayes's classic *The Second Naive Physics Manifesto*. Finally, Shanahan's *Solving the Frame Problem: A Mathematical Investigation of the Common Sense Law of Inertia* gives a detailed account of the development of the Event Calculus.

There is, of course, also a vigorous debate at a more philosophical level about the futility or utility of AI, and about the worth of computational and logic-based models of human reasoning. Famous doubters include John Searle, Noam Chomsky and, more recently, the physicist Penrose (1989) in his book *The Emperor's New Mind*. A good summary of these philosophical standpoints, and an arguably effective demolition of them from a strong Darwinian perspective, is given in Dennett's(1995) *Darwin's Dangerous Idea*.

Note

1. McCarthy's example is that somebody has stuck a potato in the exhaust pipe. It is unlikely that anyone creating an ME "car" theory would include a "PotatoInExhaustPipe" fluent, let alone include the negation of this as a qualification in proposition (4).

References

Davis, E. (1990), *Representations of Commonsense Knowledge*, Morgan Kaufmann, San Francisco, CA.

Dennett, D. (1995), *Darwin's Dangerous Idea*, Allen Lane, Penguin Press, Harmondsworth.

Gelfond, M. and Lifschitz, V. (1998), "Action languages", *Artificial Intelligence*, Vol. 2 Nos 3-4, pp. 193-210, available at: www.ep.liu.se/ej/etai/1998/007/ (accessed 22 November 2005).

Genesereth, M. and Nilsson, N. (1987), *Logical Foundations of Artificial Intelligence*, Morgan Kaufmann, Palo Alto, CA.

Hanks, S. and McDermott, D. (1987), "Nonmonotonic logic and temporal projection", *Artificial Intelligence*, Vol. 33 No. 3, pp. 379-412.

Hobbs, J. and Moore, R. (Eds) (1985), *Formal Theories of the Commonsense World*, Ablex Publishing, Norwood, NJ.

Kakas, A. and Miller, R. (1997a), "A simple declarative language for describing narratives with actions", *Journal of Logic Programming*, Vol. 31 Nos 1-3, special issue on Reasoning about Action and Change, pp. 157-200.

Kakas, A. and Miller, R. (1997b), "Reasoning about actions, narratives and ramifications", *Journal of Electronic Transactions on Artificial Intelligence*, Vol. 1 No. 4, pp. 39-72, available at: www.ep.liu.se/ej/etai/1997/003/ (accessed 23 November 2005).

Kakas, A., Michael, L. and Miller, R. (2005), "Modular-E: an elaboration tolerant approach to the ramification and qualification problems", in Baral, C., Greco, G., Leone, N. and Terracina, G. (Eds), *Proceedings of the 8th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR'05), Diamante, Italy, September 5-8, 2005*, *Lecture Notes in Artificial Intelligence*, Vol. 3663 No. 1, Springer-Verlag, Berlin, pp. 211-26.

Kakas, A., Miller, R. and Toni, F. (1999), "An argumentation framework for reasoning about actions and change", in Gelfond, M., Leone, N. and Pfeifer, G. (Eds), *Proceedings of the 5th International Conference on Logic Programming and Nonmonotonic Reasoning (LP-NMR'99), El Paso, Texas, December 2-4, 1999*, *Lecture Notes in Artificial Intelligence*, Vol. 1730 No. 1, Springer-Verlag, Berlin, pp. 78-91.

Kakas, A., Miller, R. and Toni, F. (2001), "E-RES – reasoning about actions, events and observations", in Eiter, T., Truszczynski, M. and Faber, W. (Eds), *Proceedings of the 6th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR'2001), Vienna, Austria, September 17-19, 2001*, *Lecture Notes in Artificial Intelligence*, Vol. 2173 No. 1, Springer-Verlag, Berlin, pp. 254-66.

Kowalski, R. and Sergot, M. (1986), "A logic-based calculus of events", *New Generation Computing*, Vol. 4 No. 1, pp. 67-95.

Lifschitz, V. (Ed.) (1990), *Formalizing Common Sense – Papers by John McCarthy*, Ablex Publishing, Norwood, NJ.

McCarthy, J. (1959), "Programs with common sense", *Proceedings of the Teddington Conference on the Mechanization of Thought Processes*, Her Majesty's Stationery Office, London, pp. 75-91.

McCarthy, J. (1980), "Circumscription: a form of non-monotonic reasoning", *Artificial Intelligence*, Vol. 13 Nos 1-2, pp. 27-39.

McCarthy, J. and Hayes, P. (1969), "Some philosophical problems from the standpoint of artificial intelligence", *Machine Intelligence*, Vol. 4 No. 1, pp. 463-502.

McIlraith, S., Son, T. and Zeng, H. (2001), "Semantic web services", *IEEE Intelligent Systems*, Vol. 16 Nos 2, special issue on the Semantic Web, pp. 46-53.

Miller, R. and Shanahan, M. (1996), "Reasoning about discontinuities in the event calculus", in Doyle, J., Aiello, L. and Shapiro, S. (Eds), *Proceedings of the 5th International Conference on Principles of Knowledge Representation and Reasoning (KR'96), Cambridge, Massachusetts, USA, November 5-8*, Morgan Kaufmann Publishers, San Francisco, CA, pp. 63-74.

Miller, R. and Shanahan, M. (2002), "Some alternative formulations of the event calculus, in computational logic: logic programming and beyond", in Kakas, A. and Kowalski, R. (Eds), *Essays in Honour of Robert Kowalski Part 2*, *Lecture Notes in Artificial Intelligence*, Vol. 2408 No. 1, Springer-Verlag, Berlin, pp. 452-90.

Penrose, R. (1989), *The Emperor's New Mind: Concerning Computers, Minds and the Laws of Physics*, Oxford University Press, Oxford.

Russell, S. and Norvig, P. (1995), *Artificial Intelligence: A Modern Approach*, Prentice-Hall, Englewood Cliffs, NJ.

Russo, A., Miller, R., Nuseibeh, B. and Kramer, J. (2002), "An abductive approach for analysing event-based requirements specifications: best application paper award", in Stuckey, P. (Ed.), *Proceedings of the 2002 International Conference on Logic Programming (ICLP'02), Copenhagen, Denmark, 29 July-1 August, 2002*, *Lecture Notes in Computer Science*, Vol. 1364 No. 1, Springer-Verlag, Berlin, Berlin, pp. 22-37.

Scherl, R. and Levesque, H. (1993), "The frame problem and knowledge-producing actions", in Fikes, R. and Lehnert, W. (Eds), *Proceedings of the 11th National Conference on Artificial Intelligence (AAAI'93), Washington DC, USA, August 18-20, 1993*, AAAI Press, Menlo Park, CA, pp. 689-95.

Shanahan, M. (1997), *Solving the Frame Problem: A Mathematical Investigation of the Common Sense Law of Inertia*, MIT Press, Cambridge, MA.

Tran, N., Baral, C. and Shankland, C. (2005), "Issues in reasoning about interaction networks in cells: necessity of event ordering knowledge", *Proceedings of the 20th National Conference on Artificial Intelligence (AAAI'05), Pittsburgh, Pennsylvania, USA, July 9-13, 2005*, AAAI Press, Menlo Park, CA, pp. 676-81.

Corresponding author
Rob Miller can be contacted at: rsm@ucl.ac.uk