

# Information Science in Transition

Michael Day<sup>1</sup>

UKOLN, University of Bath, Bath BA2 7AY, United Kingdom  
[m.day@ukoln.ac.uk](mailto:m.day@ukoln.ac.uk)

Review of: *Information science in transition*, edited by Alan Gilchrist, London: Facet Publishing, 2009, 401 pages, hardback, ISBN 978-1-85604-693-0, £49.95.

Until it joined with the Library Association in 2002 to form the Chartered Institute of Library and Information Professionals (CILIP), the Institute of Information Scientists was a professional organisation for those primarily working in scientific and technical information work. The chapters in this volume were first published in 2008 as a special issue of the *Journal of Information Science* to commemorate the founding of the institute in 1958. In accordance with this, many of the chapters provide a retrospective - sometimes even anecdotal - overview of developments in information science in the UK since the 1950s. While the approach of the volume is thematic, a major focus is on key initiatives and individuals, the latter including such luminaries as Jason Farradane, Cyril Cleverden and Karen Spärk Jones.

Following a guest editorial by Brian Vickery, there are sixteen chapters in the book. While each chapter stands alone, conceptually the volume moves - with some exceptions - from largely retrospective reviews of past progress in information science by scholars of the older generation to overviews of current trends and technologies by their younger colleagues. Vickery's editorial tries to place information science in its historical context, explaining how the advent of digital computers and the Internet has transformed the discipline dramatically while simultaneously making its future more uncertain. This is also a view articulated by several of the volume contributors.

The opening chapter is an attempt by Jack Meadows to discern the main research themes in UK information science over the past 50 years. A survey of the *Journal of Information Science* and other journals showed that the predominant theme was information retrieval, but that there was also important research being undertaken into information seeking, communication and bibliometrics. The chapter also tries to delineate some of the factors affecting information science research in the UK, for example noting the negative consequences of the demise of the old British Library Research and Development Department in the 1990s [1]. He concludes, however, on a positive note, pointing out that "activities that were relatively marginal decades ago - such as automated information retrieval - are now at the heart of major growth industries" (p. 17). He also notes that the widening interest in information science concepts has brought in researchers from other disciplines - which is probably one of the key lessons of the whole book. In the second chapter, David Bawden (City

---

<sup>1</sup> Draft book review submitted to *Ariadne* issue 63 (April 2010)

University) again uses the *Journal of Information Science* as a means of exploring the development of the information science discipline itself, focusing on the underlying philosophical bases of the subject proposed by scholars like Bertie Brookes and Jason Farradane.

The third chapter is by Stella Dextre Clarke. This is a retrospective of fifty years of knowledge organisation work in the information science domain that takes a partly anecdotal approach, attempting to illustrate "how it felt to work in those times" (p. 45). Perhaps the best aspect of this is that it enables Dextre Clarke to give the reader a feel for what information retrieval could be like in the card-based pre-computer age. The chapter opens with a brief overview of the state of subject classification in the late 1950s, noting the continued practical predominance of enumerative schemes like the Dewey Decimal Classification while the theoreticians S. R. Ranganathan and Henry E. Bliss were still working away developing their (then) revolutionary ideas of 'faceted classification.' The focus then changes to the development of thesauri, noting the importance of Jean Aitchison's pioneering work on thesaurus construction. Dextre Clarke then provides a very brief overview of the role of controlled vocabularies in the early information retrieval tests conducted as part of the Aslib-Cranfield Research Project, a topic covered in more detail in the following chapter. Finally, moving to the present day, Dextre Clarke notes the continued importance of controlled vocabularies in the form of taxonomies and provides some pointers for a future Semantic Web.

Stephen Robertson (Microsoft Research Laboratory, Cambridge) provides the following chapter on the history of evaluation in information retrieval. He starts out with an outline of the Cranfield experiments led by Cyril Cleverdon in the 1960s and follows this with a brief overview of US developments in computer-based retrieval as represented by Gerard Salton's SMART Information Retrieval System and the National Library of Medicine's MEDLARS bibliographic retrieval service. The focus then shifts back to the UK with the attempts by Karen Spärk Jones and others to define how to plan and execute information retrieval experiments, including the development of standard test collections. The proposal for an 'ideal' test collection was not taken forward at the time, but its role is now fulfilled in part by the annual Text REtrieval Conference (TREC), which has been running since 1992 [2]. Robinson provides an outline of the TREC process and methodology and its evolution into the Web era.

Chapter 5 turns to the all important subject of user studies. Tom Wilson starts his review with J. D. Bernal's study of the use of scientific literature that he presented at the Royal Society Scientific Information Conference in 1948. Follow-up research was relatively slow in getting underway, but by the 1970s, there had been detailed studies of scientists' literature searching practices by John Martyn of the Aslib Research Group and of social scientists' information requirements by a research group based at the University of Bath Library [3]. Wilson identifies the growth period for user studies as the 1980s, by which time "information use and users had become a curriculum topic in the schools of librarianship in the UK" (p. 99). Despite this growth, which Wilson tracks through Web of Science data, the chapter concludes with a warning about the dangers of this branch of information science becoming too removed from actual practice.

Blaise Cronin (Indiana University) then contributes a typically elegant chapter entitled "The sociological turn in information science." Partly anecdotal in nature, the

chapter explores the interdisciplinary nature of information science by tracing the influence of the wider social sciences. Cronin's starting point is that 'the social' has long been part of the information science discipline, whether implicitly or explicitly. Using the metaphor of a 'turn' - e.g., as pioneered by the "linguistic turn" in philosophy - Cronin outlines some of the major contemporary influences on the social sciences, including the range of approaches characterised as critical theory. He concludes that the information science discipline "has long been mindful of, and indeed receptive to, sociological thinking" (p. 122).

The following two chapters focus on particular application domains. The first is a historical introduction to the discipline now known as chemoinformatics, written by Peter Willett of the University of Sheffield. This chapter outlines the information-rich retrieval challenges of chemistry, which has long had specific requirements, typified by the development of services like the American Chemical Society's CAS (Chemical Abstracts Service) registry or the Beilstein database. More recently, the advent of high-throughput experimentation and combinatorial chemistry has vastly increased the size and diversity of the data that needs to be interrogated by computational tools. Willett sees the emergence of chemoinformatics as being "driven in large part by the scaling-up of ... techniques, which had traditionally been aimed at just a few tens of molecules, to the very large datasets that characterize modern pharmaceutical research" (p. 155). Chemoinformatics itself is viewed by Willett as a specialised form of data mining, "involving the analysis of chemical and biological information to support the discovery of new bioactive molecules" (p. 133). The following chapter, by Peter Bath (also of the University of Sheffield) covers health informatics. As with chemistry, healthcare planning and clinical practice is increasingly becoming dependent on the widespread capture and availability of vast amounts of data, "collected, stored, analysed, transferred, and accessed on a daily basis" (p. 170). The chapter outlines the particular challenges of health informatics, including the significant problem of integrating legacy data and ethical issues like security, privacy and confidentiality.

In chapter 9, Elisabeth Davenport of Napier University explores the connections between social informatics research in the USA - as practiced by Rob Kling and others [4] and the tradition of sociotechnical studies undertaken in the UK, described as a "fusion of sociology and computing in ICT research" (p. 201). The chapter mainly focuses on work undertaken by three UK research groups, Enid Mumford and colleagues at the University of Manchester, the Science Studies Unit at the University of Edinburgh, and the 'critical informatics' approach of the London School of Economics.

The final chapters take a broadly thematic approach to particular information science challenges. Peter Enser (University of Brighton) contributes a chapter on visual information retrieval, both semantic and content-based, noting how the field has been entirely transformed by greatly increased availability of images (both still and moving) on the Internet. Elizabeth Orna and Barry Mahon then present chapters respectively on the development of information policies at national and institutional level and on professional development in the information sector. Mahon's chapter is described as a 'personal view' and is written in an informal style at variance with other parts of this volume. That said, he does usefully define metadata as "a default term for

what information scientists would recognize as the whole range of activities and processes designed to make information retrievable" (p. 295).

The next three chapters cover some topics of active current interest. First, Charles Oppenheim of Loughborough University provides an overview of electronic scholarly publishing and open access. After a brief explanation of some of the factors affecting scholarly publishing in the scientific, technical and medical (STM) domain, Oppenheim turns to address the open access (OA) agenda in more detail. The chapter outlines the two main forms of OA that have been adopted so far, often characterised as the 'green' and 'gold' approaches, respectively depositing copies of papers in repositories and publishing in OA journals. The business models of both are compared. That for the green approach is said to be "simply that the body maintaining the repository pays for the ingest of materials, addition of metadata and other technical and administrative requirements" (p. 303). The business models for gold OA journals shift the main costs of publication from subscribers to authors - or in many cases their employers or research funders. Oppenheim points out that if gold OA was ever adopted on a wide scale, HE institutions would probably have to carry a bigger proportion of costs in order to maintain the market (p. 306). Other topics covered in this chapter include the implications of OA publishing on journal quality (including peer review), the potential increase of citation counts of OA materials, copyright, and the attitudes of publishers, HE institutions and funding bodies.

The following chapter covers the extremely topical subjects of social software and social networking tools. In this, Wendy Warr introduces a number of Web 2.0 technologies, and indicates how some of them are used in publishing and industrial contexts. The chapter provides very brief introductions to: wikis, blogs and RSS feeds, social networking tools (e.g., Facebook, LinkedIn, Flickr, YouTube), social bookmarking sites (e.g., delicious, Connotea, CiteULike), and virtual worlds (e.g., Second Life). Warr notes that the "world of Web 2.0 is fast changing," which doubtless explains why the chapter does not mention the social networking tool de nos jours: Twitter. The chapter ends with an overview of how these tools are being used by information professionals, publishers, and the pharmaceutical and chemical industries, e.g. for developing collaborative workspaces.

Jack Meadows's opening chapter identified bibliometrics as one of the key topics of the past 50 years of information science. In chapter 15, Mike Thelwall of the University of Wolverhampton provides an overview of recent developments in that field. Thelwall dates the emergence of bibliometrics as a scientific field to the development of the Institute for Scientific Information (ISI) citation databases in the 1960s. Their existence encouraged the development of both evaluative bibliometrics based (in the main) on the analysis of citations and relational bibliometrics interested in exploring the structure and nature of scientific disciplines. Thelwall argues that the biggest recent change in bibliometrics "stems from the availability of new significant sources of information about scholarly communication, such as patents, web pages and digital library usage statistics" (pp. 351-352). The chapter then introduces how bibliometrics has been used recently to develop tools for the evaluation of individual researchers (the h-index) and for supporting national research evaluation exercises. In the final section, Thelwall provides a brief introduction to the domain that he has done much to shape himself, i.e. Webometrics [5]. The methods used for this include the

analysis of hyperlinks - e.g. for generating web impact factors - and web citations in journal or conference papers.

The final chapter in the volume is a short piece by Eugene Garfield jauntily entitled: "How I learned to love the Brits." This anecdotal piece focuses in part on his attendance at the Dorking Conference on Classification in 1957 and the many friendships that he had forged then and since with his UK colleagues.

Looking at the whole volume, a few key themes stand out. Firstly, some of the authors emphasise the important balance that needs to be maintained between theory and practice. Wilson's comments on the potentially overly-theoretical nature of user studies have already been mentioned. In a similar vein, Robertson notes that the laboratory nature of the TREC experiments means that they are essentially an abstraction, while "certain aspects of the real world are highly resistant to abstraction (p. 86). Secondly, it is clear that certain aspects of information science research have been completely transformed by the growing availability of vast amounts of networked information. This applies in particular to information retrieval research, where the TREC experiments have needed to demonstrate scalability as well as precision and recall, as well as to more specialised areas like image retrieval or subject classification. In other areas, it has led to the development of completely new areas of research, e.g. Webometrics. Thirdly, Cronin's chapter, together with the chapters on chemoinformatics and health informatics, demonstrates that information science is highly interdisciplinary in nature. Cronin has, however, noticed that citation relationships are often unidirectional. With reference to the sociology of science, he notes that this domain's "intellectual vanguard ... are cited routinely in the leading information science journals," but that "regrettably, there is little citation in the reverse direction" (p. 119).

Perhaps the most important general insight from this volume is the importance of linking information science research to its wider contexts. Given the sociological turn in information science described in Cronin's chapter, it would seem useful to apply some of these techniques to topics covered in other chapters. For example, in his discussion of institutional OA policies and mandates, Oppenheim comments that "surveys and anecdotal evidence indicate that the real issue is not that scholars are not convinced by OA, but rather they do not have the time to convert their materials into a format suitable for an IR [institutional repository], and the matter is simply not a high enough priority for them" (p. 314). Perhaps this area would represent a profitable one for sociologists of science to explore in more detail.

This volume would be of interest to anyone interested in the history of information science in the UK and for those wanting to get an overview of current trends. The volume is well produced and the layout very clear. I noticed the occasional typographic error, chiefly in personal names, e.g. Cramer for Cranmer (p. 98), Jacob for Jacobs (p. 321). One minor annoyance was that Mahon's conception of 'Europe' seems to conflate the geographical expression with the political entity of the European Union (and its predecessor organisations). The main remaining question that needs to be asked is whether papers already published in an academic journal really need to be published again in book form? After all, many potential readers may already have access to this content via their institutional subscriptions to e-journals. The editor addresses this challenge in his preface, commenting that publishing in book form

enables the content to be distributed to a wider readership. On balance, this assessment is probably correct.

## References

1. Goulding, A. "Searching for a research agenda for the library and information science community." *Journal of Library and Information Science* 39(3), September 2007, 123-125.
2. Text REtrieval Conference (TREC): <http://trec.nist.gov/>
3. Line, M. B. "The information uses and needs of social scientists: an overview of INFROSS." *Aslib Proceedings* 23(8), 1971, 412-434.
4. Van House, Nancy A. "Science and technology studies and information studies." *Annual Review of Information Science and Technology* 38, 2004, 3-86.
5. Thelwall, M., Vaughan, L., and Björneborn, L. "Webometrics." *Annual Review of Information Science and Technology* 39, 2005, 81-135.