**Title**
Knowledge organization and the technology of intellectual work



**ISBN**
9783956500404

**Author**
Buckland, MK

**Publication Date**
2014

Peer reviewed

Manuscript. Published in *Knowledge Organization in the 21st Century: between Historical Patterns and Future Prospects. Proceedings of the thirteen internation ISKO Conference 19-22 May 2014, Kraków, Poland.* Ed. by W. Babik. Würzburg: Ergon-Verlag, 2014. Pp 14-21. Published version may differ slightly.

**Michael K. Buckland**

# Knowledge Organization and the Technology of Intellectual Work

**Abstract**

Since ancient times intellectual work has required tools for writing, documents for reading, and bibliographies for finding, not to mention more specialized techniques and technologies. Direct personal discussion is often impractical and we depend on documents instead. Document technology evolved through writing, printing, telecommunications, copying, and computing and facilitated an 'information flood' which motivated important knowledge organization initiatives, especially in the nineteenth century (library science, bibliography, documentation). Electronics and the Internet amplified these trends. As an example we consider an initiative to provide shared access to the working notes of editors preparing scholarly editions of historically important texts. For the future, we can project trends leading to ubiquitous recording, pervasive representations, simultaneous interaction regardless of geography, and powerful analysis and visualization of the records resulting from that ubiquitous recording. This evolving situation has implications for publishing, archival practice, and knowledge organization. The passing of time is of special interest in knowledge organization because knowing is cultural, living, and always changing. Technique and technology are also cultural ("material culture") but fixed and inanimate, as can be seen in the obsolescence of subject headings, which remain as inscribed while culture moves on. The tension between the benefits of technology and the limitations imposed by fixity in a changing world provide a central tension in knowledge organization over time.

This twenty-fifth anniversary of ISKO is an opportunity to recognize past achievements and to think about the future. Here I will consider some aspects of technology and of time and, for this, it is important to remember the ambiguity of the word *knowledge*: There is knowledge in the head and there is recorded knowledge. I will use *knowing* and *documents* respectively.

## The history of document technology long term

All communities, all societies, all collaborations arise through interaction and communication among the members. Prehistoric societies used speech, dance, and gesture to communicate and art to record. A performance of speech, dance, or gesture is ephemeral. It is extremely limited in time and highly localized in space. If you are not present there and then, you do not see or hear it. Evolving technologies steadily reduced these limitations. The usual distinction between prehistoric and historic times is the use of writing and it is convenient to start with writing and four other lines of technical development: printing, telecommunications, document copying, and computing.

***Writing.*** By recording, writing provides an alternative to speech, whether by making it permanent or by making it unnecessary. Writing can 're-mediate' speech into an enduring form and, being portable, it can resist the effects of distance as well as time. The statement or image remains and so can be seen at any future time. But writing is not limited to the recording of speech and gesture. It can simply be an original inscription, often to record that something has happened (history) or than something should happen (agenda).

In all cases the effect is to establish a trace, an evidence of something that can thereafter be perceived by others for the first time (communication) or serve as a reminder for oneself or for others (memorandum). In this way the written record endures, thereby conquering the passage of time, at least for as long as the record remains legible.

But writing does more than prevail against the tyranny of the passing of time. By being portable it can also reduce the effect of distance. A single written record can, in principle, be read by anyone anywhere, even though it can only be in only one place at any given time. Writing, then, diminishes the effect of time and so provide an alternative to human memory, an 'external memory'. Much has been written on the consequences of the invention of writing in providing an enduring stable form of evidence, thereby facilitating communication, control, and commerce. It is hard to imagine life without writing.

***Printing.*** Printing provides an extreme multiplication of writing and so multiplies the effects of writing with two consequences. First, printing makes copies which can be in as many different places as copies have been made. The more widely that copies are scattered the more convenient for geographically dispersed individuals. This matters because convenience of access is a powerful determinant of use. Second, since any individual record is vulnerable to alteration, including falsification and destruction, there is safety in numbers. The more copies that are made and the more widely they are distributed, the more likely it is that at least one authentic copy will survive.

Printing enabled churches to standardize their liturgies more effectively. It also facilitated the Reformation and the rise of modern science. Much has been written on the impact of printing.

***Telecommunication.*** Until the nineteenth century, telecommunication was a person on foot, horse, or ship bringing good news or bad. The rise of transmission technologies, notably railways, telegraph, telephone, radio, and now the internet, had the effect of reducing the effects of distance and diminishing the delays associated with travel. Telecommunications, like printing facilitated coordination, control, and commerce. There is a large literature on the effects of telecommunication.

***Copying.*** Copying has thee meanings: Transcribing, duplicating, and facsimile. Transcribing is a special case of writing. The numerous forms of duplicating (e.g. mimeographing) involve the creation of a new original and so are more properly treated as a form of printing. In the eighteenth century handwritten documents were copied by "letter press": a thin damp sheet of paper was pressed against the original so that some of the ink of the original would transfer into the moist sheet. During the nineteenth century photography was used occasionally to copy documents, but techniques for generating rapid, reliable, economical facsimile copies of documents is a twentieth-century development with three really important techniques: Photostat, microfilm, and electrostatic (xerography). Compared with writing, printing, and telecommunications, there has been relatively little historical or social commentary on the impact of modern copying technology.

*Photostat*, direct-projection photography on to sensitized paper without an intermediate negative image, was pioneered around 1900 by René Graffin in Paris to facilitate his editing of early Christian writings in Syriac. The resulting image was negative (white writing in a black ground). The left-to-right reversal was corrected by using a 45° mirror. The speed, accuracy, and efficiency of photostats for copying text and images compared with manual transcription or typewritten copies were quickly recognized and the photostat process was widely adopted (Hawken 1960; United States 1912).

*Microfilm* was famously used by Dagron in the siege of Paris in 1870, but widespread use was delayed until the 1930s when compact precision cameras, standard film speeds, and 35 mm safety film were available. Banks, newspapers, libraries, and other organizations adopted microfilm and its variants on a large scale.

*Electrostatic copying*, also known as xerography, became available during the 1960s and remains the technology of choice for copying and for printing digital documents.

The photographic copying of documents developed jointly with forensic copying, making clear copies of documents that were damaged, illegible, or suspect. Lodewyk Bendikson, a Dutch surgeon who became a librarian in California, demonstrated how photographic techniques could make legible copies of documents with ink spills, erasures, burned, obliterated by censors, or falsified (Buckland 2012). From a technical perspective document facsimile copying is not separate from image enhancement.

***Digital computing*** can be seen as a sequence of just two kinds of primitive operation: (i) Deriving new records from old, which could be subsumed under writing; and (ii) Sorting existing records, as in information retrieval and visualization (Buckland & Plaunt 1994; Plaunt 1997).

**The challenge for knowledge organization**

These developments in document technology reinforce each other to increase enormously the number of documents. People speak of an "information society," but all communities, including prehistoric societies, depend on information and communication. The real change is the increase in documents so we should speak of an emerging "document society." In the nineteenth century people worried about an "information flood." In the twentieth century it became the "information explosion" and, now, "big data."

Many records are temporary, but some are expected to have continuing value. We cannot remember everything, so we need an artificial, external memory in the form of documents. But both individuals and institutions need to identify, locate, and retrieve specific records, which required another line of technical development with many names, including bibliography, documentation, information science, or knowledge organization.

Libraries were in the care of scholars whose familiarity with the collection enabled them to recommend the most suitable documents for any reader's need, but Martin Schrettinger and others recognized that the memories of scholarly librarians are lost when they forget, move away, or die. Schrettinger (1808) asserted the need for formal systems for finding and retrieving documents and he coined the term '*Bibliothek-Wissenschaft*' ("library science") for it. The techniques of modern librarianship were well-developed by Dewey and others by the late nineteenth century when there was an increase in interest in bibliography and documentation.

**The technology of intellectual work**

Authors generate documents. Librarians collect them. Bibliographers index them. Authors use the indexes to find documents and then write new ones, which librarians collect and bibliographers index. This dynamic of complementary roles continues like a double helix.

The documentalists of the early twentieth century, notably Paul Otlet, Wilhelm Ostwald, and Otto Frank, tried to take a more integrative approach. Otlet, especially in his later years was interested in sophisticated work stations and communication networks (Otlet 1934). Wilhelm Ostwald established an institute in Munich named The Bridge (*Die Brücke*). The manifesto of the The Bridge was entitled *Technik der geistigen Arbeit* (Bührer & Saager. 1911). Ostwald wanted to extract facts and concepts from books and periodicals in order to recall units of recorded knowledge ("monographs") more effectively. New concepts and facts would be added and each element updated as needed. By 1911 he and his colleagues were writing lyrically about this "World Brain" (*Das Gehirn der Welt*). H. G. Wells promoted the same idea until he realized that available technology was insufficient. Ostwald wanted more. Just as individual letters of moveable type could be rearranged to form new words, rearranging concepts could form new knowledge, he hoped (Ostwald 1911).

Ostwald and Otlet represented a modernist view based on systems, logic, standards, machinery, efficiency, and progress (Buckland 2007b). Sadly, existing technology was inadequate. It was also a naïve, utopian view based on a limited understanding of knowing. Even scientific facts cannot be properly understood out of context. As Ludwik Fleck (1935) explained, there were always three elements: the person, the record, and the cultural context of the person. This has large consequences for anyone imagining that a technological system of knowledge organization and retrieval could be sufficient.

**The case of documentary editors**

Scholarly annotated editions of historically significant texts constitute an important foundation for learning and research in the humanities. Documentary editing requires a major investment of highly specialized expertise. Editors and their assistants spend much time researching people, places, events, and themes connected with their selected papers in one way or another. The results of these time-consuming investigations are typically kept in the project offices and may result in a few lines of explanatory note in the eventual published volume. The reality is that most of what editors learn is not included in the published volumes, is not shared with other researchers, and is probably discarded when grants for publication expire.

This situation is more regrettable when editorial work is duplicated in other editorial projects with overlapping scope. For example, Emma Goldman and Margaret Sanger knew each other and were active in some of the same circles, so the editors of the Goldman papers and the editors of the Sanger papers, located over 3,000 kilometers apart, often research the same individuals and topics, not to mention overlap with editorial work at other documentary editing projects, or other scholars, such as historians, archivists, and curators of special collections.

There is currently considerable work on software for the production of scholarly editions, especially of literary texts, and on techniques for adding annotations. Here, however, we are interested in the research resources created and used during the editing process rather than with the production of the eventual edition itself: Working notes, collected data, lists, references, clippings, photocopies, etc. There are also specialized locally-developed products such as itineraries, chronologies, legislative histories, and lists of names, places and organizations. Only some of this material appears in the very concise footnotes, endnotes, commentary, and appendices in the eventual published print or online editions.

Aspects of documentary editing work could be improved with Web technology: Shared access within an editorial project; shared access between related editorial projects to enhance collaboration; public access to editorial research findings; and more interoperability between editorial projects and other kinds of scholarly enterprise.

A project is based at the University of California, Berkeley, is addressing these issues in collaboration with three editing projects: The Emma Goldman Papers Project (Berkeley); The Margaret Sanger Papers (New York University), and The Elizabeth Cady Stanton and Susan B. Anthony Papers Project (Rutgers). Each project agreed to place working notes, mainly simple flat text files and scanned images, on a shared website provided for them (http://editorsnotes.org/).

Such a website facilitates the management of a multi-person editorial project because it becomes much easier for a senior editor to see, monitor, and add comments and instruction to the work of assistants. A shared website enables projects with overlapping scope to see each other's work. In August 2012 password control to the site was quietly removed making the site openly available to both humans and webcrawlers. A month later, after web search engines, including Google, Bing, and Baidu (China), had indexed the contents, the resources on the site were being visited by scholars from around the world. Other editorial projects could adopt the same or a similar approach to "open note book science" in the humanities.

The importance of this change in work practices is modest in the context of a single project. The real impact of a new "editor=s notes on the Web" genre would come from its adoption as a standard best practice for

editorial projects generally, internationally. With time, the accumulation of indexed, accessible editors= notes could have a substantial impact and bring an increased visibility and recognition that the projects need and deserve.

**The future of document technology long-term**

If we extrapolate our concise history of document technology into the future what do we find? I suggest that the logic of each is as follows:

*Writing*, a means for the recording of speech, is moving steadily towards *the recording of everything*.

*Printing*, the multiplication of texts, is evolving into *the representation of anything*, especially now with 3D printers.

*Telecommunications*, in effect the transport of documents, becomes, with sustained improvement, *pervasive simultaneous interaction*.

*Copying*, because it depends for versatility on the use of image analysis and enhancement, leads towards the *analysis of documents and data*.

*Computing*, in additional to deriving new documents and enabling other technologies, supports ever *more complex visualizations*.

A simple extrapolation of past trends, then, leads to a document society increasingly characterized by ubiquitous recording, pervasive representations, simultaneous interaction regardless of geographical distance, and powerful analysis and visualization of the records resulting from that ubiquitous recording. New technology allows disparate genres to be woven together much more completely than before and, of course, more complex challenges and opportunities for knowledge organization.

**Knowledge organization, culture, and time**

It is generally accepted that knowledge organization is formed within cultural contexts (Bowker & Star 2000). Since culture constantly changes, time is important in understanding knowledge organization. This can be seen in the obsolescence of subject descriptions (Buckland 2007a; 2011).

By definition, the descriptions created by indexers are for future use. This requires the indexer to think about likely needs and to describe (name) in a forward-looking way. It is not simply a matter of what the document is about, but of how it might be viewed in an imagined future (Greisdorf & O'Connor 2003). Meanings, however, are established by usage and so always draw on the past. Subject description, therefore, is also backward-looking We create descriptions by drawing on the past, but work with an eye to the future. This Janus-like stance might seem difficult enough in a stable world, but reality is made much worse by time, by technology, and by social change.

The act of recording the topical description of a document or of specifying a relationship between named topics, is necessarily performed at some point in time and inscribed into the knowledge organization apparatus: bibliographies, catalogs, and indexes. As time passes that act recedes from the present into the past. During the same flow of time the prior discourse, upon which the choice of name was based, has evolved and changed, so description can be expected to change too. Also, as the future becomes the present, new futures continue to be foreseen, and the forward-looking perspective of the indexer increasingly relates to new and different future discourses. New names arise, especially for new topics, through figurative use of language, especially through metaphor (Norgard 2002).

Information systems need stable records and an assigned description, once inscribed, is fixed. So, with the passing of time, its relationship with both the then-past discourses and also the then-future expected needs both drift away from an advancing present. Assigned names are, therefore, inherently obsolescent with respect to both the past and the future. Discourses and the indexer flow forward with time, but the assigned names have been inscribed for, and fixed in, a receding past. A static subject indexing vocabulary ceases to be effective.

There are consequences for description from affective changes. Even when the denotation is stable, the connotation or attitudes to the connotation may change. Always, some linguistic expressions are socially unacceptable. What is deemed acceptable or unacceptable not only differs from one cultural group to another, but changes over time. (Berman 1971; Marshall 1977).

**Technology, culture, and time**

Knowledge organization depends heavily on technology. Documents are physical objects on paper, film, magnetic disks, or other physical media. Efficiency depends on tasks being routinized, reduced to clerical procedures, or delegated to machines. Knowing is cultural, living, and changing. Technology, too, is cultural. Objects and systems reflect the context of their origin and are, in academic terminology, "material culture," but they are not, like knowing persons, living and changing. In knowledge organization there is a tension between the unstable character of knowing and the machinic tendencies essential for cost-effective performance.

**Summary**

The history of information technology shows continuing increases in the number of documents. The so-called "information society" is really a "document society." Knowledge organization exists to advance learning and understanding by helping to deal with the flood of documents. Intellectual work requires technology which needs to be formal and machinic to be efficient. But what people know, what they would like to know, and what others have learned and written about, resist mechanical treatment. If it were otherwise, education and knowledge management could be reduced to data processing.

Indexes and categorization systems use language and, for that reason, are in conflict with the need to have stable, unambiguous marks to enable knowledge organization systems to perform efficiently. Knowing is cultural, alive and changing. Technology is cultural too, but inanimate and stable. This tension will remain important in the relationship between knowledge organization and the technology of intellectual work.

**Acknowledgments**

The support of the A. W. Mellon Foundation is gratefully acknowledged.

**References**

Berman, S. (1971). *Prejudices and antipathies: A tract on the LC Subject Heads concerning people*. Scarecrow, Metuchen, NJ.

Bowker, G. & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press.

Buckland, M. K. (2007a). Naming in the library: Marks, meaning, and machines. In *Nominalization, nomination and naming in texts*. C. Todenhagen & W. Thiele, eds. Tübingen: Stauffenburg. Pp. 249-260.

Buckland, M. K. (2007b). On the cultural and intellectual context of European Documentation in the early twentieth century. In *European Modernism and the Information Society: informing the present, understanding the past*, ed. W. B. Rayward. Aldershot, UK: Ashgate. Pp. 44-57.

Buckland, M. K. (2011). Obsolescence in subject description. *Journal of Documentation* 68 (2): 154-161.

Buckland, M. K. (2012). Lodewyk Bendikson and photographic techniques in documentation, 1910 – 1943. In *International Perspectives on the History of Information Science and Technology Worldwide, Baltimore, 2012*. Ed. by T. Carbo and T. B. Hahn. Medford, NJ: Information Today. Pp. 99-106. Preprint at http://people.ischool.berkeley.edu/~buckland/bendikson2012.pdf

Buckland, M. K. & C. Plaunt. (1994). On the construction of selection systems. *Library Hi Tech* 48: 15-28.

Bührer, K. W. & Saager, A. (1911). *Die Organisierung der geistigen Arbeit durch die Brücke.* Ansbach: Seybold.

Fleck, L. (1935/1979). *The genesis and development of a scientific fact*. Chicago: University of Chicago Press.

Greisdorf, H. & O'Connor, B. (2003). Nodes of topicality: Modelling user notions of *on topic* documents. *Journal of the American Society for Information Science and Technology*, 54: 1296-1304.

Hawken, W. R. (1960). *Full-size photocopying*. New Brunswick: Graduate Library School, Rutgers—the State University, 1960. Hathi: http://hdl.handle.net/2027/mdp.39015023479549

Marshall, J., comp. (1977). *On equal terms: A thesaurus for non-sexist indexing and cataloging*. New York: Neal-Schuman.

Norgard, B. A. (2002). *Linguistic expressions and indexing information resources*. Ph.D dissertation in Library and Information Studies, University of California, Berkeley.

Ostwald, W. (1912). Das Gehirn der Welt. *Nord & Süd* 36, (1): 63-66

Otlet, P. *Traité de documentation: Le livre sur le livre*. Brussels: Editiones Mundaneum, 1934.

Plaunt, C. J. (1997). *A functional model of information retrieval systems*. Ph.D dissertation in Library and Information Studies, University of California, Berkeley.

United States. *Report of the Commission on Economy and Efficiency*. Washington: Govt. Printing Office, 1912. Hathi: http://catalog.hathitrust.org/Record/001044389.