Editorial

MICHAEL SEADLE
Dean, Faculty of Arts, and Director, Berlin School of Library and Information Science, Germany

World Digital Libraries 6(2): v-vi

Knowledge Management plays a significant role in this issue and increasingly in the world of library and information science, yet the definitions of Knowledge Management are quite various, as Chakraborty and Laloo point out in their table on p. 78-79. Most of the definitions have something to do with capturing knowledge or with capturing information to transform it into knowledge. Information is relatively concrete. It includes all forms of text, images, numerical data, and can be sorted, structured, reprocessed and repurposed. Information comes in various types of containers including paper and computer-based storage, and it comes in formats that are generally well established. Knowledge is a far more slippery term that represents a level of abstraction above and beyond the information base on which it builds. Knowledge is, in a sense, how people understand information after they have interacted with it, but the moment that knowledge becomes concrete in written or spoken or visual or numeric form, it also crosses a boundary that transforms it into information that machines or paper or other media can store. Knowledge degrades into information with every keystroke of every author.

The term Knowledge Management suggests a misleadingly static process in which humans or machines put everything that has been thought or known into convenient categories and classifications for easy retrieval, and it misses the essential dynamic by which knowledge creation occurs. Knowledge Management originated in the business world, and one of the early goals of relational databases was to make it possible for anyone in an organization to query the full range of stored information (or at least the range for which they were authorized) in order to learn something new. That something new was knowledge creation, which might be stored as information (knowledge), or might merely remain in the mind of the person who executed the query in the form of a not quite fully formulated idea.

In the scholarly world knowledge creation is the goal, and every scholar who builds on existing works in the form of stored information has probably interacted with some form of Knowledge Management system, even if it is only the online catalogue of the local library or the table of contents of a book. Knowledge Management systems are so ubiquitous that

it is easy to forgot how ancient they are, and equally easy to forget how meaningless the term can become if not used precisely in a way that distinguishes it from information management. Retrieval is only one aspect of Knowledge Management, and generally the least transformative. The key question is how people will behave with the information once they have it.

In the 21st century the number of transformative tools for Knowledge Management has grown considerably. The most discussed of these cluster under the heading of the semantic web, which builds linkages that enable the dynamic recombination of information sources. A related aspect of the semantic web are the triple-store databases with inference engines of varying levels of sophistication. A well-designed and well-functioning triple store resembles the old expert systems of the 1980s in modern form. Developments in machine intelligence rarely come into discussions of Knowledge Management in the library and information science world, but the ability of machines to transform information into something like knowledge has grown rapidly. In Google's experiments with self-driving cars, for example, computer systems are processing new information (for example, speed, distance between cars, and road conditions) in conjunction with stored information about the car's mass, inertia, breaking capabilities, turning radius, and tipping point. The result is a set of decisions, in effect the knowledge, about how to drive that particular car under those particular traffic and road conditions.

Another transformative process that manages knowledge, but is rarely discussed in that context, is today sometimes called "distant reading" and more traditionally falls under "text mining". It is essentially an old process that linguists and literary scholars used as soon as they could get access to computers and digital texts. Early results were used to look at word counts and to make arguments about putative authorship. The basic tools such as Regular Expressions in languages like C, Perl, or Python have not fundamentally changed, but the questions have evolved as the quantity of stored and readily available digital information has grown. Historians can, for example, now write programs to answer questions that they want to ask of archival information in digital form. At one time the only choice was to go to an archive and read the texts page by page – a slow and laborious process that limited data collection. In so far as digital content exists, and especially if it is internet-accessible, a program can do the reading. The most serious difficulty is neither retrieval nor programming, but the scholar's ability to formulate questions that the program can reasonably answer with the available data. In other words, this process depends fundamentally on whether scholars can define precisely what they want to know about the information.

Ultimately Knowledge Management is not information storage via a database or a classification system. Managing knowledge comes essentially through the intellectual process of formulating a useful and answerable question that the stored information can answer.