The information paradox

Alan GILCHRIST

Cura Consortium and Metataxis Ltd.
alangilchrist77@gmail.com

Resumen

Reflexiones personales sobre la situación actual y perspectivas futuras de la ciencia de la información, cuya práctica está muy extendida pero cuya teoría no se ha consolidado todavía. Tras varias décadas de fuerte desarrollo de las prácticas de gestión de la información, fundamentalmente como soporte de las disciplinas técnicas y científicas, los avances en las áreas tradicionales de actividad práctica se han ralentizado a medida que se generalizaba la aplicación de la tecnología de la información. Mientras tanto, no se ha materializado la aparición de una verdadera teoría de la ciencia de la información.

Palabras clave: Ciencia de la información. Teoría. Práctica. Tecnología de la información. Evolución. Prospectiva.

Abstract

Personal thoughts on the varied practice and still unformed theory of information science. After some four decades of strong development of information handling practice, largely in support of the disciplines of science and technology, advances in the traditional areas of practice have slowed down as information technology has been more widely applied. Meanwhile, a true theory of information science has failed to materialize.

Keywords: Information Science. Theory. Practice. Information technology. Evolution. Prospective.

1. Some etymology

Everyone "owns" information but few understand it. Economists rank it with energy, materials and capital as the basic building blocks, yet 'information handling' is diffuse, often unconnected and poorly managed. Consequently, the word information has become almost meaningless on its own and often misleading when combined with another. Shannon's 'Information theory' is actually the theory of the transmission of signals, independent of their possible meaning; 'Information retrieval', coined by Calvin Mooers was judged by Robert Fairthorne to be 'Reference retrieval' (and even today Google retrieves websites); the 'Information revolution' is actually the 'Communications revolution', and so on.

More recently, the word 'Knowledge' has crept in, making further assaults on people's understanding. It is difficult though, or even impossible, in considering the terms Knowledge Management and Knowledge Organization to know how one can 'manage' or 'organize' knowledge. Those who are active in these two areas know what they mean by the two terms, but in themselves the terms remain opaque to others.

It should be remembered that the term 'Information scientist' was first coined in the 1950s by Jason Farradane and colleagues to describe qualified scientists who would search the litera-

ture on behalf of other scientists. For them, the Institute of Information Scientists was founded, and to be a member one needed to have a degree in science or engineering, a second language and five year's experience of 'information work'. Not long after, Farradane set up the first course in 'Information science', teaching at a London Polytechnic that later became The City University. Doubtless, this elevation in terms was to satisfy academic respectability.

Similarly, the first (slim and small format) journal of the Institute was called *The Bulletin of The Institute of Information Scientists* later becoming *The Information Scientist*, which later still changed into the current *Journal of Information Science.* The word information, with little variation between European languages (Spanish *información*, French *information*, German *Informationen*, Italian *informazioni*) was easily combined with the word science to give '*ciencias de la información*' in Spanish; '*sciences de l'information*' in French (note the plurals); '*Informationswissenshaft*' in German (where the word *Wissenschaft* does not translate literally into science, but comes from the German word for 'knowledge'); and '*scienza dell'informazione*' in Italian.

The term 'information science' was also adopted in the U.S.A., where the journal *American Documentation* changed its title to *Journal of the*

American Institute of Information Science, later adding *'& Technology'.* So 'information science' appeared to become an international discipline, though it can be noted that cross-national journal citations are less common than those in the exact sciences.

2. Rise and decline of Information Science

2.1. A Golden Age?

World War 2 accelerated the advance of science and technology, and in particular the development of the digital computer. One of the first bibliographic tasks was to analyse the large number of technical reports won from the Germans and as an aid in this task H. P. Luhn of the IBM Corporation came up with the idea of a KWIC Index (Key Word in Context) in order to sort titles into manageable lists. The inventor and science administrator Vannevar Bush, in an influential paper published in 1945 had already predicted that "Wholly new forms of encyclopaedias will appear, ready made with a mesh of associative trails running through them ready to be dropped into the Memex and then applied". (Memex stood for MEMory EXtender).

Five years later, Calvin Mooers coined the expression 'Information retrieval' and developments accelerated fast as the computer grew more powerful allowing online searching. Early successes included the database of the National Library of Medicine in the early 1970s and the Lockheed Dialog system.

In the early days of searching, trained experts had to construct complex formulae using the Boolean operators of AND, NOT, and OR, careful to nest components where necessary. The formulae were then fed into a mainframe computer and the searches were run overnight and any errors had to be corrected and re-run the following night. Then, Stephen Robertson and Karen Sparck Jones of the U.K. University of Cambridge developed Boolean probabilistic searching and Gerard Salton of Cornell University followed this up with the more mathematically complex Vector Space Modelling.

Helen Brownson of the U.S National Science Foundation is credited with being the first person to use the word 'thesaurus' in 1957 in relation to information retrieval, and there followed the compilation of a large number of bulky and complex thesauri produced by, among others, the Engineers Joint Council with *TEST (Thesaurus of Engineering and Scientific Terms, 1967)* containing 17,800 descriptors, followed by thesauri from the larger international agencies such as UNESCO, OECD and the World Bank.

In the U.K., lagging behind the U.S. in computer technology and use there was a greater emphasis on classification and not only were contributions made to the Universal Decimal Classification, but the CRG (Classification Research Group) founded in 1952, developed the Bliss Classification with funding from NATO (extraordinary in retrospect). The CRG organized a Conference on 'Classification Research for Information Retrieval' which brought together such experts as Brian Vickery, Robert Fairthorne and Jack Mills. Also present were Eugene Garfield, the 'inventor' of the *Science Citation Index,* and the Indian mathematician and librarian S. R. Ranganathan, the originator of faceted classification and the Colon Classification, who had a huge influence on British classificationists. Another attendee was Cyril Cleverdon who instituted the influential Cranfield experiment providing comparative evaluations of leading retrieval techniques. This was succeeded by the TREC (Text Retrieval Conference), funded by the National Science Foundation, through which researchers and vendors could evaluate retrieval packages and techniques.

Moving from batch processing to online retrieval opened the way for information scientists to act on behalf of users, with iterative interrogation techniques. Eventually, the end users were provided with terminals as well, giving access to the central computer, but often with limited success. One U.K. government department provided proforma access with provisions for either 'simple search' or 'advanced search' (i. e. supporting concealed Boolean operators). A survey then showed that simple search was preferred in a ratio of some 95% to 5%.

In 1948 the Royal Society of London held a seminal Conference on Scientific Information in which the internationally renowned scientist J. D. Bernal was prominent. Bernal also attended another important conference in Washington in 1958 jointly organized by the National Science Foundation, the National Academy of Sciences and the American Documentation Institute (later ASIS&T).

It was Bernal who was instrumental in setting up a Research Department in Aslib, the Association of Special Libraries and Information Bureaux that had been set up in 1924 when the Library Association failed to take an interest in the nascent industrial libraries. This gave Aslib a new lease of life as it managed to attract government research funding through the Department of Scientific and Industrial Research along with such organizations as the Building Research Station and the National Physical Laboratory. Indeed, the first Director of the Research De-

partment was C.W. Hanson who had been Information Officer of the Scientific Instrument Research Association. He was succeeded by Brian Vickery who later became Professor at University College London. Other members of the Department included Stephen Robertson who later became Professor at The City University in London and Blaise Cronin, now Rudy Professor Emeritus of Information Science at Indiana University. As mentioned above, the expansion of Aslib was followed by the establishment of the Institute of Information Scientists for individual membership (unlike Aslib which had corporate membership) and with a strong interest in education.

2.2. A slow decline

With the advent of distributed processing and the personal computer, the world of work changed radically. First, the typing pool disappeared and then, when the World Wide Web appeared, many special libraries and information departments were either closed or drastically reduced as senior managers thought that money could be saved: surely, they thought, individuals could now get the information they needed from the Web.

Disintermediation set in and the fewer information scientists and librarians that were left were used as sources of help as a last resort. Then it was realised that the Web could not provide all that was needed, so Intranets were set up by IT Departments with minimal support grudgingly received from the LIS staff. Many organizations even had separate units designing and managing their outward facing Internet website and inward-facing intranet.

Every individual with a personal computer now became an amateur author, graphics artist, publisher, librarian and information scientist in addition to the core professional job. This, of course, was good news for the IT industry which was now able to sell thousands of PCs instead of a few mainframes, and a cycle was established whereby new software was produced, requiring more powerful hardware which attracted new software which worked better on updated hardware and so on.

Along with these developments there were in-house rivalries for budget where the Human Resource Department (previously called the Personnel Department) 'hijacked' the emerging area of knowledge management; and the post of CIO (Chief Information Officer) was introduced, being actually the head of ICT operations. Both the CIO and the Director of HR (often as members of the Board) are big spenders on neces-sary infrastructure while the LIS facility, though spending far less, found it difficult to justify its budgets in monetary terms.

One of the results of this is that while senior management continues to state confidently that good information is essential to the success of the business, surveys continually show that the 'knowledge workers' (another silly term merely meaning that they don't work, for example, on an assembly line) are strongly dissatisfied with facilities supporting 'enterprise search'.

The decline has been reflected in the U.K. by reorientations within academia and a related fragmentation of professional associations. Some of the traditional Schools of Library and Information Science have changed their descriptive titles and moved into Faculties of Computing or Business. Some have dropped the word 'science' after the word 'Information' in favour of the word 'studies', while others have opted for the term 'Information Management'. A recent informal survey concluded that the traditionally core subjects of classification and retrieval languages had either disappeared or were given brief attention, University College London being one of very few that still give these subjects prominence in its curriculum.

Following a decline in membership numbers, the Institute of Information Scientists merged with the Library Association in 2002 to form CILIP, the Chartered Institute of Librarians and Information Professionals (a strange title that seems to suggest that librarians are not information professionals while failing to define the scope of the latter term).

When the two bodies merged there were some 28,000 members but that number has since fallen to around 13,000. Within CILIP there are no less than 25 Special Interest Groups, many of whose interests were undivided in the defunct Institute. Possibly the most successful, and the only one allowed to have members outside the membership of the parent body is UKeIG (U.K. Electronic Information Group) formerly UKOLUG (the U.K. Online Users Group). Other Special Interest Groups are devoted to Aerospace and Defence Libraries; Government Information; and Commercial, Legal and Scientific Information. Outside CILIP, the British Computer Society hosts the Information Retrieval Specialist Group, the Business Information Systems Specialist Group and a third titled Knowledge, Information, Data and Metadata Management (1).

Other independent associations include the-International Society for Knowledge Organization (UK Chapter), the Information and Records Management Society, the British and Irish Asso-

ciation of Law Librarians, the Archives and Records Association, and the Society of College, National and University Libraries.

A small, but potentially significant addition to the list is provided by the use of the word 'Informatics' (broadly speaking a combination of information science and information technology) in specific areas, such as Cheminformatics, Health Informatics and Social Informatics and there are now university courses available in all three.

It is quite possible that there is a similar picture of fragmentation in other countries with the concepts of library, documentation and information in separate camps.

As a minor and perhaps overstated indication of the partial eclipse of the information profession it is perhaps worth noting the recent publication of a book in 2010 by a prominent publisher advising graduate students on how to find the right resources. The principal author is a senior lecturer in political studies and his co-author a 'subject adviser' in that person's university library. There is a glossary included in the book with an entry for the term 'Classification' (but none for thesaurus or taxonomy) which includes the advice

> Contrary to common belief, these classification tools [typologies, concepts, theories] are not for confusing the reader: the idea is to bring order to the complexity which is social life.

An entry for the term 'Keyword' reads

> A keyword sums up what a document is about and is used when searching for documents in bibliographic databases and subject gateways.

There is a feeling that the shadow of Google is behind these definitions.

3. Is a revival of Information Science possible?

The decline was noted by Meadows (2009) in a volume of essays under the title *Information Science in Transition* when he said

> […] the information science activities developed over the past 50 years have triumphed, but information science as a separate entity may be on the wane. If so, its final epitaph may well be that of Sir Christopher Wren (in the St Paul's Cathedral that he had designed) - If you want a monument, look around you.

But perhaps the epitaph to information science is presented too soon. Robinson (2014) in a review of a book by Stock and Stock (2013) remarks that several more books on information science have recently appeared (Davis and Shaw, 2011; Bawden and Robinson, 2012,

Iberkwe-SanJuan, 2012) (2), which might indicate some small flames springing up from the embers. It is interesting to note that these four books emanate from Germany, the U.S.A., the U.K., and France respectively.

So far in this paper the use of the term information science has inferred its practice rather than any possible theory. What seems clear as we move forward into an increasingly technological twenty-first century is that the practice is becoming even wider than that described by Vickery in his seminal book of 1973 in which he said "The field of study is so wide and varied".

This presumably means that any theory must become more complex, unless we can accept that we may be talking of interconnected information sciences in the plural and a consequent plurality of theories, perhaps with some overarching connector. There have been many attempts to find a place for information science in 'the greater scheme of things'. For example, Brookes (1980) in a paper invoking a theory of Karl Popper's, claimed that information science was one of the social sciences, if not the most important one, while Garcia Marco (2013) convincingly describes information science(s) as an inter-discipline within an 'ecology' of sciences.

Ecology seems to be a useful concept here when one considers such sub-disciplines as biochemistry, socio-economics and psycholinguistics. Informology may be an ugly term, but it will do for the moment in suggesting the possibility of such areas of study as socio-informology, informo-economics and psycho-informology. These might usefully be linked and each might support various theories each of which might have healthily conflicting viewpoints leading to cross-fertilization, as happens between the various schools of economic theory.

This concept of ecology may become even more convincing as the practice of information work develops within the realities of globalization and the accelerating development of the 'information sciences' in such countries as Korea, China, India, and Brazil.

The Figures that follow are over-simplified, but intended to suggest in broad terms the scope and practice of the information science(s).

Fig. 1 is a simple statement of what must surely be the 'backbone' of the information sciences. An author or authors working together issue a message or messages designed specifically for an intended recipient or recipients, or broadcast to unknown recipients.
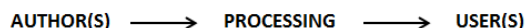
AUTHOR(S) ⟶ PROCESSING ⟶ USER(S)

Fig. 1. The 'backbone' of information sciences

In Fig. 2 the message is either transmitted orally, or via some recording medium, largely through primary and secondary publishing or direct onto the Internet. This involves considerations of Open Access Publishing, copyright, commercial aspects and the form of publication.
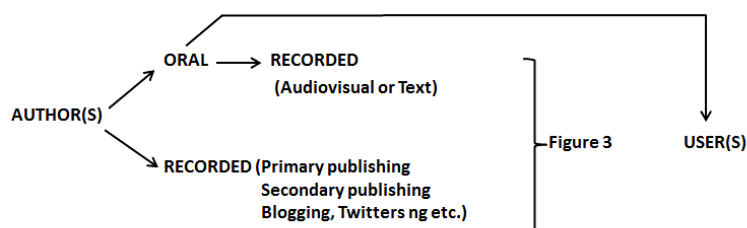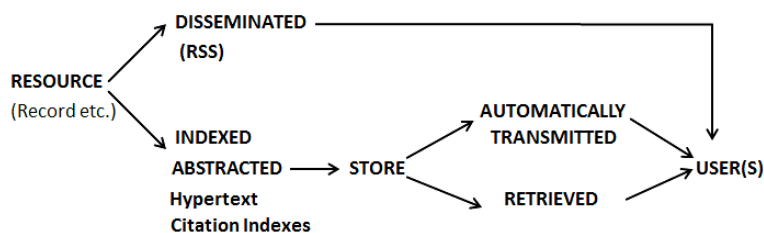

Fig. 2.

Fig. 3. Information 'processing'

At this stage the resource (text, visual, spoken) is either disseminated directly or is processed, perhaps by indexing and abstracting before being placed in storage (i.e. a memory bank) from where it may, again, be transmitted automatically (as with RSS feeds) or retrieved by physical or online search.

The 'processing' in Fig. 1 is effected by 'information systems', some of which are shown in Figure 4, together with broader aspects of the systems and supporting disciplines, such as Information Architecture. One of the more important features of this Figure is to show how important it will be, for example, that Knowledge Organization and Knowledge Representation should work together; and for the managers of 'memory stores' to learn more from each other in the use of techniques than they have in the past. It would also be sensible if the two areas of Information Management and Knowledge Man-

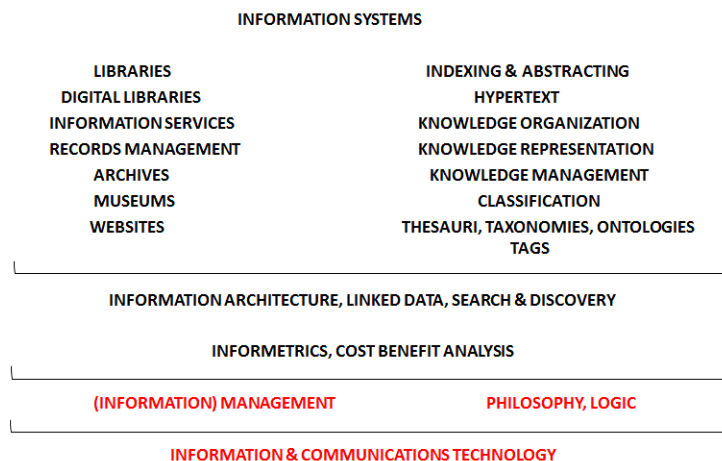agement were to merge wherever that seemed profitable.

Figure 4. Information systems

Finally, there is the consideration of the important aspects of the immediate use and the application of the information extracted, assimilated and correlated by individuals and groups of users working in collaboration.

So far in this paper the emphasis has been, by implication, on the practice of information science, but what of a possible theory? Brookes, in claiming that information science was a social science in the paper cited above, viewed information science in the singular, but it now seems clear that it is a plural discipline. Just as there is no overarching theory of the social sciences, so there is no likelihood that there can be one for the information sciences. This is not to say that theory is not applicable to the study of the information sciences or that such study might be unprofitable. Research and the construction of hypotheses will always be useful.

It is becoming increasingly difficult to predict what the world might be like in the next ten or twenty years as information technology becomes both more powerful and more pervasive. Bostrom (2014) has reviewed in some detail the paths, dangers and strategies posed by the emergence of 'superintelligence' based possibly on genetic engineering and certainly on advanced artificial intelligence.

If traditional information science (whatever that is) does not take a wider and more considered view of its scope and potential before and possibly within this revolution, then its component parts will be assimilated by others and the fragmentation will continue. As Tom Wilson (2010) put it in his speech accepting an Honorary Doctorate at the University of Murcia in 2010:

Let us not restrict ourselves to grubbing around in the garden patch of a limited, little information science, restricted to the relationship between information and machine. Instead, let us expand, reach out, embrace and explore the wider world of information, to develop a vision of information science as a central synthesising discipline in understanding not simply information, but the world we live in. Because the world we live in is surely a world of information.

Notes

(1) I was once talking to the Dean of a Faculty of Computing, Mathematics and Information Science about the merger of the LA and the IIS and he expressed his opinion, while admitting its impossibility, that a better solution would have been to split the BCS in two, "sending" the BCS technicians to the then Institution of Electrical Engineers and merging the rest with the IIS.

(2) None of these books were available to me at the time of writing.

References

Bawden, D.; Robinson, Lyn (2012). Introduction to information science. London: Facet Publishing.

Bostrom, Nick (2014). Superintelligence: Paths, Dangers, Strategies. Oxford: Oxford University Press.

Brookes, B. C. (1980). The foundations of information science. Part 1 – Philosophical aspects. // Journal of Information Science. 2:3/4, 125-134.

Davis, C.H.; Shaw, D. (2011). Introduction to information science and technology. Medford NJ: Information Today.

Garcia Marco, J. F. (2013). The place of information science(s) in the ecology of sciences. Unpublished. Paper presented at Ibersid 2013.

Ibekwe-SanJuan, F. (2012). La science de l'information, origins, theories and paradigms. Paris: Lavoisier.

Meadows, Jack (2009). Fifty years of UK research in information science. // Gilchrist, Alan, Editor. Information science in transition. London: Facet Publishing.

Robinson, Lyn (2014). Review of Stock, W.G. And Stock, M. Handbook of Information Science. // Journal of Information Science. 40:2 (2014) 264-265.

Stock, W. G.; Stock, M. (2013). Handbook of information science. Berlin: De Gruyter.

Vickery, B. C. (1973). Information systems. London: Butterworth.

Wilson, T. D. (2010). Information and information science: An address on the occasion of receiving the award of Doctor Honoris Causa, at the University of Murcia, September 2010. Information Research. 15:4, paper 439. http://informationr.net/ir/15-4/paper439.html.