Kim H. Veltman

"New Ways of Scholarly Work in a Networked World:" ***Image, Text, Sound and Technology: A strategic research support program,*** University of Calgary: University of Calgary Press, 2004. (In press)

Abstract

Serious visions of networked knowledge have existed since the 1930s. The advent of inter-net-worked new media points to a shift from Information and Communication Technologies (ICT) to Universal Convergence Technologies (UCT). This shift is transforming both knowledge organization and the ways of scholarly working although the consequences of the shift will take decades and even centuries. This paper focuses on three areas of change. First, there is a shift from fixed repositories (as isolated laboratories, libraries, museums and archives) to distributed resources. This is transforming the way we link to sources, and increasing enormously the amount of enduring knowledge that is accessible. At the same time it is bringing enormous amounts of new evidence through both sensors and novel forms of personal and collaborative knowledge.

Second, along with these developments in connectivity to enduring and new knowledge, there are enormous developments in the methods whereby we access new knowledge. This is pointing to virtual reference rooms that will transform the ways we deal with names, with classification systems and with interpretations of knowledge. Third, there is a trend towards collaborative research and creativity, which is affecting definitions of authors and communities. This is also leading to dynamic forms of knowledge. More significantly it is confronting us with challenges of encoding and sharing different ways of knowing.

---

1. Introduction

Scholarship requires access to sources and equipment. These were traditionally in memory institutions (libraries, museums and archives) and laboratories[1] The invention of microfilm (1900[2]), which led to the founding of University Microfilms (1938, now Proquest[3]), marked a first step away from direct dependence on physical presence at fixed repositories of memory institutions and labs. By 1945, Vannevar Bush had conceived of a Memex device that used microfiche and "in which an individual stores all his books, records and communications and which is mechanized so that it may be consulted with exceeding speed and flexibility." [4]

More than a decade earlier, Paul Otlet (1934),[5] who had worked with Lafontaine in establishing a universal bibliographical organization of intellectual work,[6] had envisaged a machine for reading:

> Soon television will be a problem that has essentially been resolved just as it is already from a scientific viewpoint: the image reproduces itself at a distance, wireless. One can imagine an electric telescope which allows one to read at home pre-ordered pages of the books exhibited in the "teleg" room of large libraries. This will be the telephoned book.[7]

This led to "a machine to think/imagine the world" (*la machine à penser le monde*, 1943)[8] as part of his larger vision of a Mundaneum for access to all human knowledge. This was a large, television-like sphere that allowed viewers to study information literally from all sides.

Neither of these machines were developed directly as Otlet or Bush envisaged. Nonetheless, Vannevar Bush's student, Claude Shannon, applied the principles of Boolean logic to define bits and bytes as they are now used in computers. Bush's essay also inspired Douglas Engelbart, Ted Nelson and most of the pioneers of what has since been called hypertext and hypermedia. By 1968 the Internet became a reality. In the 1990s, it began to spread far beyond narrow scientific and scholarly communities and continues to grow. There are now over 800 million internet users and this number is predicted to rise to 1100 million by the end of 2005.

There are also three further technological developments on the horizon. First, there is nanotechnology, whereby it is predicted that a single gramme of DNA will hold the equivalent of 1 trillion CD-ROMS. Second there is a new convergence whereby telephony, television and Internet will become a single set of technologies. Partly because of this, there is, third, a new trend whereby not only different media but also different senses are interchangeable. Hence a spoken message can be recorded as a written one and a published text can be reproduced in manuscript form or even orally. Internetworked media thus introduce a new kind of digital bridge whereby even illiterate persons can be included within the learning structure. These developments suggest that we are moving from Information and Communication Technologies towards Universal Convergence Technologies (from ICT to UCT). This essay explores emerging and possible consequences thereof for scholarly work with respect to distributed resources, virtual reference rooms and collaborative research and creativity.

2. Distributed Resources

2i) Enduring Knowledge

The past decade has seen enormous advances concerning access to enduring knowledge in memory institutions. For instance, Canada has been a pioneer in a Virtual Museum of Canada which brings together images of collections from 1035 institutions through CHIN.[9] Many countries now have national networks of libraries. The German network (DBV) now has over 40 million entries. Global efforts such as the Research Library Group provide access to 136,121,972[10] bibliographic records (as of 28 July 2004), which is roughly ten times the size of the British Library in 1950.

Increasingly these bibliographic records are being complemented by full text scans of the original books. High level scanning now ranges from 100-767 MB per page, which means that these facsimiles address the needs of all but the most subtle problems of connoisseurs. Through a method of the British Library called "Turning the Pages,"[11] the full texts of ten classics are now online. At the Bibliothèque Nationale de la France, the full text of over 75,000 books has been scanned in. In China, the complete classics (800 million characters) have been scanned in using Unicode and are available in 8 dialects. If the method of turning the pages becomes a general practice this could revolutionize the way scholars consult manuscripts, books and related materials.

An open question remains whether these materials will be accessible free of charge. Publishers especially in the periodicals field continue to assume that it is their prerogative to make considerable profits while providing minimal reimbursements to authors. This goes to extremes. For instance, Oxford University Press recently accepted for publication an article (based on a keynote at an international conference) then stopped publication of after the author refused to concede more than rights to a single edition of the work. Meanwhile, the academic community is exploring alternatives. In the United Kingdom, the JISC (Joint Information Systems Committee) has embarked on making major resources available at reasonable prices. For example over 250,000 articles of the Royal Society of Chemistry are being made available for £50.[12] In Germany, the Max Planck Gesellschaft has a vision of free access to all the e-science preprints from over 80 research institutes. This has led to a "Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities."[13] In the United States, the Open Archives Initiative has a similar rhetoric but less evidence of actual free content.[14] Internationally, and especially in Europe there is a growing sense that if public monies are used to create libraries and museums, then public monies such also be used to make these materials freely accessible to citizens.[15]

2ii) Sources

Traditionally the quality of a book was judged by the footnotes and bibliographies that linked its claims with books and sources elsewhere. This meant that a judgment about what was inside its cover depended on the precision with which it pointed beyond its covers. The good news was that this enabled readers to check a book's sources. The bad news was that readers inevitably had to enter upon elaborate hunting exercises in order to track down these often elusive sources.

A deeper implication of networked knowledge[16] is that these footnotes and references within its covers can, in turn, be linked to the original sources to which they refer (books, paintings etc.). In this context, the notion of pedigree applied to texts and especially to images becomes important. A printed book contains a list of illustrations with credits that tell us only who provided the image. For instance, an original painting, say 1) the Mona Lisa, may come via 2) an official ectachrome of the Réunion des Musées Nationaux (RMN), via 3) a photograph which leads to 4) a slide which leads to 5) a photograph that is used in a given printed book. But there is usually nothing in the printed book to say that the image is 5 generations distant from the original. Often the caption merely reads: Louvre.

By contrast, in an electronic world, one can potentially trace such a pedigree of an image via its various generations. It costs no more to link with the original source (Louvre, first generation) than to go via various intermediaries (more than first generation). Since each additional generation increases the potentials for distortions of the image, scholars might in future agree that a direct pedigree (only 1 generation) becomes a criterion for making serious claims about objects.[17] If so providing a direct link to the original version of the source becomes a new criterion for scholarly work.

2ii) New Evidence

The new media are producing enormous amounts of new data and new evidence especially in the scientific fields. It is said that there are 7 million new pages per day on the regular Internet.[18] Even this is only a small part of the big picture. The Radio telescope at Conception produces 224 gigabytes of new information per day. The NASA Earth Observation Satellite (EOS) produces 918 gigabytes of new information per day: i.e. 27.5 terabytes per month and 330 terabytes per year. By 2005, the new Large Hadron Collider at CERN[19] will generate the world's biggest challenge in distributed computing. "On average 20 particle collisions occur within an LHC detector every time the particle bunches in the beam cross at the centre of the apparatus. This happens 40 million times per second, or every 25 nanoseconds, giving a raw event rate of almost a thousand million per second. This corresponds to a (compressed) data rate of about 40 TeraBytes per second."[20] High Definition Television (HDTV) produces .23 gigabytes /second: i.e. 9.8 gigabytes per minute, 588 gigabytes per hour, 14.1 terabytes per day, or 98.7 terabytes per week and 4.1 petabytes per year. This is for one station. In a 500 channel HDTV universe, the figure would be just over 2 exabytes per year. If this still seems futuristic, in December 1999, Ian Foster et al. of the Argonne National Labs reported: "In the immediate future, we anticipate collecting data at the rate of terabytes per day from many classes of applications, including simulations running on teraflop-class computers and experimental data produced by increasingly more sensitive and accurate instruments."[21]

While there will be 1 billion (10^9) computers sometime in 2005, it is estimated that there may be a 1000 times more sensors than computers: i.e. 10^{12} or 1,000,000,000,000. Many of these sensors have a simple control function, but many are also used to provide us with enormous amounts of new data from the earth, from the oceans and even from outer space. The scope of such sensors is also increasing enormously. In oceanography there is

a project to insert sensors on the entire Juan de Fuca tectonic plate in the Pacific Ocean.[22] There are similar challenges in the humanities and social sciences as we gain access to hundreds of millions of titles and increasingly their full contents. Learning how to deal with these enormous new amounts of data and information recorded by sensors and computers will be a major stimulus in changing the ways of scholarly work.[23]

3. Virtual Reference Rooms

Reference rooms are effectively the search engines of collective memory institutions. Digital equivalents in the form of virtual reference rooms offer important new tools for new ways of scholarly work. Creating digital versions of existing reference works: dictionaries, encyclopaedias, bibliographies etc. is only a first step in this process, because this simply translates the limitations of linear print into digital form.

In physical libraries we carefully separate different media into different spaces: i.e. there are separate rooms for regular books, rare books, journals, magazines, maps, drawings in major libraries. In virtual libraries such as they are presented on the Internet these distinctions are often lacking. To be sure some advanced catalogues will add a symbol to indicate that the material is a manuscript or a map, but it is not possible to search by medium. In future being able to search using different levels of knowledge such that one could focus specifically on journal articles or monograph books concerning a topic would prove a very useful filtering strategy. Virtual reference rooms also potentially offer us new approaches to names, classifications and interpretations.

3i) Names

Names of persons, places and things inevitably have multiple spellings, versions and even alternate terms. With print media there was a necessity to identify an official version as an authority file and indeed many of the efforts of the 19th and 20th centuries focused on creating such standardized authority files, the assumption being that once these official names existed then everyone could use them and one could dismiss the rest as incorrect or wrong. A narrow application of this approach meant that one only found those names written precisely in the form found in the authority file and that variants were ignored. A slightly more generous application allowed a small number of variants through see also methods.

Digital media allow us to add all possible variants as alternative ways in reaching an agreed upon standard version. This enables us to search using any variant and still arrive at a standard version. In terms of titles this means I could type in a title in Hungarian and arrive at a standard Latin or English title. Potentially persons can type in a title in their own language and arrive at the standard version in the original language. In a world where there are over 6,500 languages, and where no one can hope to learn them all, this offers a fundamental breakthrough in ways of scholarly work.

3ii) Classifications

The past centuries have seen a few attempts at universal classification schemes: e.g. Göttingen, Library of Congress, Bliss, Ranganathan. In the visual arts the 20th century added Iconclass and the Art and Archeology Thesaurus (AAT). At the national, regional and local levels, libraries, museums and archives typically created their own classification systems. In the context of print and for purposes of organizing books on shelves, memory institutions were inevitably committed to using one system at a time. Meanwhile, scholars also frequently produced their own personal lists of terms. The challenges of relating terms in one system with those in another system inspired organizations such as the International Federation of Documentation (FID) and International Standards Organisation (ISO) committees 37 and 46.

Digital media potentially introduce two advances in this domain. First, with respect to well-established classifications one can create bridges between systems. Instead of needing to adopt only one system, one can search for a term using multiple systems. Second, in the case of emerging fields where categories are often loosely defined and prone to continual change, one can use dynamic systems, which reconfigure links on the fly.[24] These innovations point to first ways in which one can create links between enduring knowledge of memory institutions and personal and collaborative knowledge produced in the context of new media. Instead of pretending to create a new ontology that will replace all others, we can begin by interconnecting existing ontologies and linking them with emergent ones.

3iii) Interpretations

In Europe, the deeper consequences of printing after its introduction in the 1450s evolved slowly. It took nearly two centuries before a wide range of major texts had been published. It was not until the 1660s that the traditions of handwritten correspondence, from the world of letters, began to be published in the *Journal des Savans* and which led in the course of the next centuries into periodical literature as we know it today. In the course of the 19th century, scholars increasingly distinguished between original work as primary sources and interpretations thereof as secondary sources.[25] The 20th century pointed to a need for four further distinctions within the kinds of interpretations in secondary sources:

1) Internal Analyses. Especially in (English) literature studies there has been great emphasis through the new criticism, close reading, deconstructivists and the various post-schools (post-modern, post colonial, post-normal) etc. to focus on internal analyses of texts as if they existed in isolation.

2) External Analyses. Such careful analyses within a given text (or painting) deserve to be distinguished from comparative studies where such texts are compared with other texts: e.g. not just an analysis of what Hamlet says in Shakespeare's play, but rather comparing this with other stories of mad princes.

3) Conservation and Restoration. The 19th and 20th centuries saw the emergence of conservation and restoration as serious disciplines to the extent that there is a considerable amount of documentation and literature tracing how specific works (e.g. paintings, frescos and manuscripts) have been affected by physical interventions in the interests of conservation.

4) Reconstructions. The Renaissance brought the advent of new categories of studies whereby authors attempted both to document existing ruins and to reconstruct how these ruins might have appeared originally. From the 16th through the 19th century printing greatly spread these studies (e.g. Cock, Du Cerceau, Piranesi). These reconstructions also became more impressive as simple woodcuts gave way to engravings, etchings, daguerrotypes and other reproduction techniques. These reconstructions typically remained a subset of the printing world until the latter 20th century when the rise of Computer Aided Design (CAD) and related electronic tools transformed this into a new category of interpretation.[26]

Although scholars can readily recognize these categories as useful distinctions they simply have not yet made their way into our bibliographical and other search instruments. Once this has occurred we shall, in future, be better equipped to find the level of interpretation that interests us. This is another example of how changing ways of organizing knowledge will affect the ways we work with knowledge.

4. Collaborative Research and Creativity

From the time it began in the late 1960s the Internet has contributed to collaborative research in high-energy physics and astronomy. This is leading to new kinds of collaboratories whereby teams work together towards solutions. When the Hubble Telescope broke down, 10,000 scientists worked together to fix the problem. These networked teams are also combining to design the Eurofighter, new cars and many new products. In industry initiatives such as the Industry Foundation Classes (IFCs) and International Manufacturing Systems (IMS) are leading to collaborative methods in design. In addition to open source, there is discussion of open theory, and even open design. There is discussion of translating design principles of the Bauhaus school to an electronic context. In the context of the European E-Culture Net there is discussion of creating a new electronic virtual agora for collaborative research and creativity. Even so the implications for these developments for scholarship have yet to be fully appreciated

4i) Communities

In physics where the core international community is under 2,000 scientists, major articles now have hundreds of authors. In the humanities and social sciences, encyclopaedias (e.g. Britannica) and dictionaries (e.g. *Oxford* or the *Dictionary of Old English*), which also entail the efforts of hundreds of scholars, have been the exception rather than the rule. The tradition of the *Oxford English Dictionary*, whereby authors remain anonymous as in the Middle Ages is laudable in the eyes of some but poses problems with respect to those collecting the necessary quotas for tenure.

Anonymous contributions can make serious contributions as witnessed by the example of the Wikipedia, whereby many persons from all over the world give freely and anonymously of their time to contribute towards a whole that is much larger than the sum of its parts. There are contexts wherein this approach has major advantages. Long before the Internet authors discovered that it was wise to have a *nom de plume* when discussing delicate subjects, especially those with political consequences. Hence 18th century France developed a convention of discussing events in Persia when they were discussing France (cf. Voltaire's *Zadig*). In such cases anonymity is an effective and important tool against both censorship and the possible consequences of criticism in a public arena.

In other cases this approach has limitations. Is it not precisely because they are anonymous and do not always provide their sources, that the Wikipedia articles are more limited than they might be? As long as we are dealing with "facts" which are universally accepted such as the day when Charlemagne was crowned: i.e. Christmas day 800, then there is no problem. But the moment we enter themes where there are conflicting opinions as to precisely when a battle occurred, when a church was founded or when a work was written, then it becomes vital for the researcher to have clues that will help them trace the sources back to their origins in order to arrive at their own individual and independent interpretations of the evidence at hand. Such examples confirm that the ability to link back to original sources discussed earlier marks much more than a luxury item or handy extra bell and whistle. In cases where facts are disputed or interpretations are open to question it becomes essential for scholars to have methods to verify sources.

Paradoxically when we are dealing with impersonal "facts" such as where something is or when something happened, we need careful documentation about the person providing the information in order that we can track their sources back to their point of origin. By contrast when we are dealing with personal claims, we often need to create special environments to ensure that the privacy of the person is protected. If a discussion group on a sensitive subject is potentially open to everyone, then individuals will frequently be unable to speak freely.

4ii) Authors

Even before the Internet became something generally used there were persons who predicted that the book was dead,[27] and claimed that the author was dead.[28] Some enthusiasts of communities predict that learning in groups and communities will soon replace authorship altogether. Here the evidence of the past two millennia provides a sobering antidote. Most major works in the humanities and social sciences have been written from a distinct viewpoint, typically developed by a single individual after several decades of hard work. No trendy group discussion is likely to replace this demanding and long- term dedication anytime soon.

For both groups and individuals the rise of the Internet and subsequently the World Wide Web also introduced other problems. On the one hand, it initially required going from writing with a pen to typing with a keyboard. The promises of new systems, which will allow us to dictate our essays, have been around for decades but are still not yet part of

everyday practice. Moving from a simple writing software (e.g. Word, Word Perfect) to presence on the Internet remains much more complex than one would like and has generated many new jobs for a new generation.

Some experts claim that this is a time of transition and this may well be so. Yet it is noteworthy that this so-called transition entails contradictory trends that are affecting the output of scholarship. With the traditional printed book, the author wrote their manuscript by hand and then the publisher transformed this into a printed book. To achieve this publishers and printers had a number of professionals to do layout, typesetting, graphics etc. In the past fifty years new media have made popular the dictum that now everyone can be an author, and been very quiet about the corollary, that every author is now also expected to do graphics, think about layout, and most of the professional tasks of publishing. Some publishers unabashedly expect "camera ready texts" and still expect the author to abandon all their rights.

To be sure the Information Technology experts paint a very different picture. For decades they have promised us "authoring tools"[29] whereby the otherwise complex programming challenges of transforming a written text to something on the Web will be made so simple that anyone can do it. This has led to a whole new genre of How to do it books "for dummies." While such books are very successful it is noteworthy that more intelligent authors now tend to "outsource" the translation exercise from computer to the network to a new generation of web specialists. So we have a rhetoric that one person can do it all, while the practice shows that publishers are neglecting to update many traditional professional specialisms, while many (amateur) practitioners are emerging in their place. Among such practitioners are not for profit companies such as New Media Studio that claim to have solutions for schools, universities and museums.[30]

Thus far, many of the efforts concerning the Semantic Web continue the rhetoric of the past decades: promising us new authoring environments whereby even beginners can do everything.[31] Fortunately there are others such as Costa Oliviera et al.[32] who envisage new kids of authoring environments, which will effectively provide personalized versions of virtual reference rooms, using ontologies to provide authors access to thesauri, dictionaries and other reference materials as they compose their new articles and books. Many open questions remain. Will such tools become open source? Will they become another market for major publishers? Or, as in other domains, will there be freeware, off-the-shelf and highly expensive proprietary solutions?

5. Dynamic Knowledge

As Marshall McLuhan made us aware, print forced a static, linear presentation of knowledge. The great advantage of thereof was that it offered a fixed version, which could be shared. The most striking limitation thereof, was that it typically presented this time-bound version as if it were the only one. If, for instance, we print statistics about the number of persons on the Internet, we need to print a new version at least annually. The statistics of 100,000 users in 1989, changed to over a 1 million with the advent of HTML, to 50 million with the advent of Internet Explorer in 1995, to 200 million in 2000 and to

over 800 million today. In a networked world such statistics can become databases, which could potentially be updated automatically as statistics become updated. All this points to the possibility of creating new kinds of self-updating books in the future.

Internet statistics are a simple example. The same principle can be applied to lists of paintings by Rembrandt, to the known manuscripts by Leonardo da Vinci, to lists of books on a subject, to maps (which vary depending on political regime[33] as well as over time), and indeed to almost any topic. Whereas printed books and articles provided us with a single, static version of such knowledge, networked media can provide us with dynamic versions such that we can trace changes over time, space and other dimensions.

In the past there was a quest to create a single, static *catalogue raisonnée* of paintings, a single bibliography. The new media point to a more cumulative, dynamic approach whereby a catalogue or bibliography is integrated in a larger whole such that we can see its relation to earlier versions. Instead of working to create a product that will replace earlier efforts, the new challenge is to show where today's efforts fit into a larger pattern that has been emerging for centuries.

If print allowed scholars to record more effectively what they did, networked media potentially allow authors to see more effectively where and how what they are doing fits into a bigger picture by showing their work in the context of these earlier contributions.

This trend points to new ways of looking at the contributions of secondary literature as a whole. It is no secret that important works of art, literature and culture generally are open to different interpretations. On closer inspection, however, it soon becomes clear that these interpretations are much fewer in number than we might have expected. To take a concrete example. In 1434 or 1435 Leon Battista Alberti in his *On Painting* (De pictura) provided a first written decription of the so-called legitimate contruction (*costruzione legittima*) of linear perspective. There are well over 100 articles on the subject, but they effectively all subscribe to one of three interpretations.

In print culture one important dimension of scholarship consisted largely in making comprehensive bibliographies, typically in chronological order, of the various interpretations that had been made concerning a given work or topic. Occasionally, review articles tried to assess trends and patterns in research. In the context of new media there is likely to be a further challenge of subdividing such chronological lists into the arguments or claims that they support. Here again a contribution of the new media is in allowing us to look at existing interpretations from a more systematic point of view. New media thus provides much more than online access: it offers a more comprehensive view of the existing evidence. Hereby, we can discern more clearly patterns of interpretation, schools of thought, trends in scholarship.

6. Ways of Knowing

From the above, scholarship clearly entails a range of activities. At the simplest level it entails tracing and establishing the so-called facts: those elements of evidence, which should not be a matter of debate. For instance a given painting has a certain size. Whether

is a small Netherlandish painting of a few centimeters or a Tintoretto of many meters, these details are potentially objective. So too are technical questions of the chemical composition of the paints involved. More elusive are questions of colours especially since various resins covering the painting may have altered the painting's colour over time, sometimes even deliberately so.

Questions of identifying figures and symbols in a painting cover a whole range of certitude: from obvious attributions concerning figures such as Christ to difficult cases where a same figure may be open to multiple interpretations. The challenge of scholarship lies not only in making interpretations but also in indicating honestly the extent to which a claim or attribution is certain or uncertain.

Differences in interpretation and uncertainty are hardly a new theme in the worlds of learning, knowledge and belief. Experts in Talmudic interpretation have known this for over two millennia. In the Buddhist canon there is the famous *Parable of the Blind Men and the Elephant*,[34] wherein a series of blind men touch different parts of an elephant and reach very different conclusions about the animal at hand. Such differences also reflect, to a certain extent, different national traditions of knowledge:

> There is an old joke about an Englishman, a Frenchman, a German, and a Jew who were asked to write an essay on the elephant. The Englishman wrote on "The Elephant and the British Empire." The Frenchman wrote about "The Love Life of the Elephant." The pedantic German wrote a large treatise on "The Toenail of the Elephant", and the Jew wrote on "The Elephant and the Jewish Problem."[35]

It is too easy to dismiss such jokes as mildly racist, as not politically correct, while ignoring an insight that things said in jest frequently reflect deeper truths than banalities spoken in all seriousness. Even in very well established fields such as archaeology it is no secret that each school has its own interpretations: the French School at Athens comes to very different conclusions than the British, German or American Schools about the same monuments.

Print media were typically written from the viewpoint of a given school. Hence, in the past, a monograph by the French School at Athens typically considered and discussed in detail the findings of French archaeologists and frequently omitted to mention the findings of other schools. Strikingly enough the Internet in its early days has followed this tradition. The websites of the British School focus on British interpretations, the French School on French interpretations and so on. This is due not only to the archaeologists: it is due as much to the mindset of programmers of websites who start from a premise that there is a single way of knowing.

Ever since the work of Sapir[36] and Whorf,[37] there has been a growing recognition that language is something more profound than saying the same thing with different sounds. If Feuerbach could proclaim that man is what he/she eats, the descendants of Sapir and Whorf would insist that man is what language he/she speaks. Thinkers such as Philippe Quéau have drawn attention to different connotations of knowledge in different languages. In English, thanks to the Anglo-Saxon links between knowing and doing (English *can*, Scottish *ken*, German *können*), there is a link between knowledge and

power. In French, savoir (cf. Italian sapienza) is linked with taste (Italian sapore, English, sap, savour). In Hebrew to now someone is to have carnal knowledge of them. In Sanskrit the root term for knowledge is linked with parturition, less on the act and more on the consequences. The anecdotes above suggested that every culture would want to know different things about a topic such as the elephant. The insights of the past century into the nature of language suggest something more fundamental: there are certain things which each language and culture can know about a given topic because of their linguistic diversity and others which it will tend to ignore.

In the decades, or even centuries to come, there will be a growing challenge to build these dimensions into our knowledge bases, and into our practices of work as scholars. A future equivalent of Google should provide us with both a) objective facts about elephants, their usual size, weight etc. which are independent of language and culture and b) more subjective dimensions of elephants which vary in terms of language and culture. We clearly do not want to impose only an English or French or German view on the world but we do need to develop tools which help citizens in each of these traditions to understand why and in what ways their colleagues look at, analyse and express knowledge so differently. By 2006 when it is predicted that there will be more Internet users in China than in the United States this challenge will acquire a new dimension.

Scholars such as Havelock, Innis and McLuhan made us very aware that the alphabet was a two-sided invention. The good news was that it brought literacy and many new potentials for sharing knowledge. The bad news was that it created a literacy divide, whereby scholarship was for the literati and illiterate persons were effectively excluded from these new advantages.

In March 2004, India began selling the Simputer (Simple Computer), the first computer specifically designed to address the needs of illiterate persons. This device costs less than $200. Within a decade the cost of such a device could fall to the range of $10-$25 and such a device could be linked with Noah Samarra's (Worldspace) vision of satellite controlled radios reaching four billion persons. This means that for the first time in history inter-networked devices could potentially reach all members of the world population. If so, those who assume that the new media are simply repeating the earlier patterns of the literary divide, and continue to speak of a digital divide, would do well to begin thinking of networked computers in terms of a digital bridge.

In retrospect we know that the ways of scholarly work were changed forever in the 19th century with the rise of sociology, anthropology, ethnology, ethnography, when scholars looked beyond the evidence of objects and books and began to study so called primitive persons who were not only illiterate with respect to European languages but often had no codified language of their own. The 20th century gradually revealed to these scholars that so-called natives or aboriginals had profound knowledge systems even if they existed only in oral form. What will happen if new tools allow these persons not only to have a "voice" as the polite saying goes, but also to make their voices part of the cumulative memory of civilization? While it is clearly premature to know precisely how this will change our approaches to scholarly work, it is safe to predict that the full implications of the new technologies have only just begun. The Internet is truly for everyone.

7. Conclusions

Marshall McLuhan became famous for drawing attention to the effects of media: how the choice of a given medium as a form has an influence on the content it contains and conveys. He focused on how the shifts a) from oral communication to literacy and b) from written literacy to printing led to an increasing emphasis on static, linear communication.[38] McLuhan made a number of intuitive, visionary assessments, probes as he would have called them, about the implications of the new electric media (especially radio and television). Since then many have been tempted to use these insights of the 1950s and 1960s and apply them to the emerging Inter-net-worked world of today.

Our essay suggests that the Internet should simply be seen as yet another medium that will replace an earlier form, in the way that printing replaced manuscripts. One of the distinguishing characteristics of the new digital form is an ability to translate any medium into any other medium: e.g. handwritten to printed, printed to oral etc. A corollary is that each of the human senses is potentially interchangeable such that sensory transduction is possible: i.e. we can feel what we see, or see what we hear. A deeper implication of this revolution is a new kind of digital bridge whereby even illiterate persons can be included within the knowledge loop of collective memory institutions.

The new internet-worked media are affecting our ways of working on at least three major fronts. A first is in terms of new approaches to distributed resources. Here they are transforming the way we organize and gain access to enduring knowledge. They are pointing to new links with sources. In terms of new knowledge they are providing us with enormous amounts of new evidence through sensors. A second front is in terms of virtual reference rooms. Here they are pointing to new approaches to alternative names, comparative classifications and systematic approaches to different interpretations. Third, the new media are pointing to new forms of collaborative research and creativity.

With analog print, the products of scholarly working were inevitably in linear and static form. Digital media enable us to develop dynamic forms of knowledge that allow us to show where we fit into earlier contributions rather than simply trying to replace them. This includes potentials for comparative languages, chronologies, geographies and even interpretations. These add new cumulative dimensions to the processes of collective memory institutions.

This new awareness of past knowledge and new tools for dealing with emerging and new knowledge is bringing new awareness that knowledge is more than something dynamic which changes over time. Different languages and cultures are actually different ways of knowing. The deeper insight offered by the new media is that there will not be one new way of working but many. The challenge of technologists will be in ensuring that the many ways can communicate, without reducing them to a monolithic system. The challenge of scholars will be to reveal in new ways the unity of diversities that is culture.

Notes

[1] There are of course many different kinds of scholarly activities. In the humanities and social sciences one's research was very much defined by access to major libraries such as London, Paris or Rome, although this varied depending one whether one's topic was a narrow edition of a specific text or a wide ranging historical questions. By contrast, in science, one traditionally needed to be at an institution that had access to high level equipment. In High Energy Physics, for example, this meant one need to be at one of a handful of major labs such as CERN (Centre Européenne de Recherche Nucléaire), Fermi Lab, DESY (Deutsches Elektronen Synchroton), and FOM (Instituut voor de Fondamenteele Onderzoek van Materie). One of the first effects of the Internet was to link these hitherto isolated labs in a single global network whereby physicists communicate with one another and share data and insights. This is leading to new kinds of collaboratories whereby scientists can operate multi million-dollar electron microscopes and other equipment from remote sites. As a result scientists can work from home as well as at major centres.

[2] Sarah Allen's Weblog, "Historical perspective on links and trails," December 28, 2003. See: http://www.ultrasaurus.com/sarahblog/archives/000097.html
Frederick Seitz, "The Cosmic Inventor, Reginald Aubrey Fessenden (1866 - 1932)," *Transactions of the American Philosophical Society Held at Philadelphia For Promoting Useful Knowledge*, Volume 89, Pt. 6, 1999. See: http://www.radiocom.net/Fessenden/.
Cf. Welcome to the Reginald A. Fessenden - ARS 'W1FRV' (first radio voice) Marshfield, Massachusetts Where the world's first radio voice broadcast took place...

[3] Proquest. See: http://www.fact-index.com/p/pr/proquest.html

[4] Image of the Memex Machine. See: www.infonet.co.jp/.../ bookshelf/Memex.html

[5] Paul Otlet. See: http://www.sims.berkeley.edu/~buckland/otlet.html. Ron Day, "Paul Otlet's Book and The Writing of Social Space," *Journal of the American Society for Information Science*, 48(4) (1997), pp.310-317. It was reprinted in: *Historical Studies in Information `Science*, edited by Trudi Bellardo Hahn and Michael Buckland. (*Information Today*), New York: Wiley, 1998, pp. 42-50.
See: http://www.lisp.wayne.edu/~ai2398/newpage4.htm
Cf. Mundanaeum Historie. 2. Une machine à faire la Paix.
See: www.mundaneum.be/content/ mundaneum/histoire3.html
Cf.: http://www.lis.uiuc.edu/~wrayward/rayward.html

[6] Mundaeneum. See: http://www.mundaneum.be/

[7] Translated by the author from the original French cited on the Mundaneum website at http://www.mundaneum.be/content/mundaneum/histoire9.html:

> Bientôt la télévision sera un problème essentiellement résolu, comme il l'est déjà scientifiquement, l'image se reproduit à distance, sans fil. On peut imaginer le télescope électrique, permettant de lire de chez soi des livres exposés dans la salle "teleg" des grandes bibliothèques, aux pages demandées d'avance. Ce sera le livre téléphoné".

[8] The sketch dated 1948.08.15 from the Otlet Archive at the Mundaneum in Mons is being published by Charles van den Heuvel, "Visualising Organisations and Organising Visualisations. Otlet's Sketches for a Synthesis of the Sciences in a European Context

and the WWW," Science in Europe – Europe in Science: 1500-2000- Maastricht, 4-6-Nov. 2004.

[9] See: http://www.virtualmuseum.ca/English/index_flash.html

[10] RLG. See: http://www.rlg.org/en/page.php?Page_ID=174

[11] British Library Turning the Pages.
See: http://www.bl.uk/collections/treasures/digitisation3.html. For an open source version, Cf. Yi-Chun-Chu, Ian H. Witten, Richard Lobb, David Bainbridge, "How to turn the page,"See: http://www.nzdl.org/html/open_the_book/p186-chu.pdf
Cf. Ian H Witten. See: http://www.cs.waikato.ac.nz/~ihw/

[12] Rebecca Carver, "Historic chemistry archive goes online for £50," *The Guardian*, London, Thursday July 29, 2004.
See: http://education.guardian.co.uk/higher/sciences/story/0,12243,1271107,00.html

[13] Max Planck. See: http://www.mpiwg-berlin.mpg.de/OPEN_ACCESS_EN.HTM

[14] Open Archives Initiative See: http://www.openarchives.org/. Cf. also the Coalition for Networked Information which used to have more optimistic visions.
See: http://www.cni.org/.

[15] Cf. the author's:"E-Culture, Art History and Museum Studies," *Changing the way research is done. The user perspective on e-infrastructure.* Brussels: European Commission, December 2003, pp. 66-67. Available on-line at: ftp://ftp.cordis.lu/pub/ist/docs/rn/veltman.pdf.

[16] Early hypertext experiments focused on electronic versions of footnotes and references within the book.

[17] To be sure persons will remain free to choose images with as many generations from the original as they wish, but then not everyone needs to be taken seriously.

[18] Cited by Jeffrey Harrow, RCFoC, 24 July 2000. See the Cyveillance Whitepaper called Sizing the Internet.
See: http://www.cyveillance.com/web/corporate/white_papers.htm.

[19] Jean Jacques Dactwyler, "Die globale Rechenmaschine," *Tages Anzeiger*, 18 Januar 2002, p. 38.

[20] Kors Bos, François Etienne, Enrique Fernandez, Fabrizio Gagliardi, Hans Hoffmann, Mirco Mazzucato, Robin Middleton, Les Robertson, Federico Ruggieri, Gyorgy Vesztergombi, *Outline of a Proposal for a Data-intensive Computational GRID Testbed*. Proposal for an outline of a project to be submitted to the EU and the NSF, Version 1.4, 20 March 2000.

> Most of these collisions are uninteresting *background* events such as collisions with residual gas molecules in the vacuum of the beam pipe. The process of selecting the interesting events for recording and further analysis is called *triggering*

[21] Ian Foster, Joseph Insley, Gregor von Laszewski, Carl Kesselman, and Marcus Thiebaux, "Distance Visualization: Data Exploration on the Grid," *Computer*, Vol. 32, No. 12, December 1999, pp. 36-43.
Cf. http://csdl.computer.org/comp/mags/co/1999/12/rz036abs.htm
(Formerly available at: http://computer.org/computer/articles/visualization_1299_1.htm)
A tutorial at INET 2001 (Stockholm) on Grids and Grid Technologies organized by Ian Foster (Argonne) and presented by Carl Kesselman gave a survey of these developments. See: http://www.mcs.anl.gov/~foster. There is now a Global Grid Forum. See: www.gridforum.org

There is also a Grid Physics Network. See: http://www.griphyn.org/

[22] NEPTUNE Project. See: www.neptune.washington.edu. This is but one of a number of global projects that include te Victoria Experimental Network under the Sea (VENUS); the Monterey Accelerated Research System (MARS); the Ocean Observatories Initiative (OOI), the Global Ocean Observing System (GOOS); the Integrated Ocean Observing System (IOOS) and the Ocean Research Interactive Observatory Networks (ORION).

[23] For one recent assessment see: *Changing the way research is done. The user perspective on e-infrastructure*, ed. Antonella Karson, François Grey, Mario Campolargo, Marie-Gabrielle Dejardin, Brussels: European Commission (DG INFSO).

[24] E.g. Aquabrowser, Brain Software, Exystence database for complexity studies.

[25] For a modern account see, for instance, http://library.albany.edu/usered/basics/primary.html.
In the sciences distinctions between primary, secondary and tertiary are different. See, for example, http://grinnell.unh.edu/primarylit.html

[26] Although methodological rigour still needs to be developed with respect to this category it is a separate area of study

[27] Frank Ogden, The Last Book You'll Ever Read, Toronto**:** Walter & Ross, 1993.

[28] E.g. Roland Barthes, "The Author is Dead," Image, Music, Text, ed. And trans. Stephen Heath, New York: Hill, 1977, Cf. K. Schmoz, D. Reitter, The Author Concept, University of Potsdam, 2001. See: http://www.reitter-it-media.de/kurse/other/absence_of_the_author_2001-slides.pdf

[29] Cf. for instance a site on Computer based teaching and Learning Tools. Multimedia Authoring tools. See: http://lorien.ncl.ac.uk/ming/resources/cal/mmedia.htm. Critics have rightly noted that such authoring tools are only aimed at persons who have no physical problems. To address this shortcoming the World Wide Web Consortium has initiated an important Authoring Tool Accessibility Guidelines Working Group. See: http://www.w3.org/WAI/AU/. As new tools are developed the promise of creating new tools is made anew with more bells ad whistles. As an exmaple, Columbia University now speaks of Authoring Environments for Mobile Augmented and Virtual Reality. See: http://www1.cs.columbia.edu/graphics/projects/mars/Authoring.html

[30] New Media Studio works with the National Science Digital Library.
See: http://www.newmediastudio.org/DataDiscovery/Authorsub.html

[31] For a review and critique of these efforts see the author's "Towards a Semantic Web for Culture," Journal of Digital Information, Volume 4, Issue 4, Article No. 255, 2004-03-15. Special issue on New Applications of Knowledge Organization Systems. See: http://jodi.ecs.soton.ac.uk/Articles/v04/i04/Veltman/ (87 pp.).

[32] Edgard Costa Oliveira, Frank van Harmelen, Mamede Lima Marques, "A Framework for Ontology Based Learning Environments," See: http://www.cs.vu.nl/~edgard/Pub/PosterISWC04Oli.pdf. This poster reports on work in progress on a significant PhD dissertation.
Important pioneers in this field include Paul de Bra's work on hypermedia. See: http://wwwis.win.tue.nl/~debra/; Thea van der Geest, Mike Sharples: The New Writing Environment. Writers at Work in a World of Technology. London: Springer Verlag, 1996. See also the website: http://www.rwth-aachen.de/technische-redaktion/summerschool/participants/theavandergeest.htm. Also of interest is work on Adaptive Multimedia. See: http://adiret.cs.uni-magdeburg.de/~nuernb/amr2003/

[33] E.g.Kashmir's maps of itself may well differ from India or Pakistan's maps of the "same" territory.
[34] The Parable of the Blind Men and the Elephant: the original version from the Buddhist canon. See:http://www.kheper.net/topics/blind_men_and_elephant/Buddhist.html
[35] Elhanan Adler, "The Digital Library and the Jews," *Meyer and Rosaline Feinstein Lecture, 2003*, National Foundation for Jewish Culture, July 28, 2004. See: http://www.jewishculture.org/jewish_scholarship/jewish_scholarship_feinstein_adler.html
[36] Cf. Edward Sapir, *Language*. New York: Harcourt, Brace & Co. 1921; and Ibid., "The Status of Linguistics as a Science." *Language* 5, 1929:209. For a discussion cf. the Sapir-Whorf Hypothesis." See: http://www.angelfire.com/journal/worldtour99/sapirwhorf.html
[37] Benjamin Whorf, 1940. "Science and Linguistics." reprinted in *Language, Thought & Reality*. Cambridge, MA: MIT Press, 1956.
[38] Giesecke explored further these implications of printing and discovered that printing, which began in Korea, was used in China to create standardized versions of their laws so that they could rule more efficiently. Germany used printing to share knowledge and to spread the word about Protestantism. So Gutenberg's innovation lay not so much in the technology as in the purposes for which he employed the technology.
Michael Giesecke, Der Buchdruck in der frühen Neuzeit - Eine historische Fallstudie über die Durchsetzung neuer Informations- und Kommunikationstechnologien, Frankfurt/Main: Suhrkamp, 1991, 2. Aufl. 1994; Michael Giesecke: *"Von den Mythen der Buchkultur zu den Visionen der Informationsgesellschaft". Trendforschungen zur kulturellen Medienökologie. Mit einer CD-ROM mit dem Volltext des Buches sowie weiteren Aufsätzen und Materialien.* Suhrkamp Verlag: Frankfurt am Main, 2002.. Since then Matellart has examined how the rise of global communications systems has transformed claims about the universality of knowledge. Armand Matellart, *The Invention of Communication*, Minneapolis: University of Minnesota Press, 1996; Armand Matellart, *Networking the World, 1794-2000*, Minneapolis: University of Minnesota Press, 2000.