Kim H. Veltman

New Media and Transformations in Knowledge

Opening keynote: “Metadata und die Transformation des Wissens,”
Published as New Media and Transformations in Knowledge I, II *in:*
Euphorie Digital? Aspekte der Wissensvermittlung in Kunst, Kultur und Technologie, Heinz Nixdorf Museums Forum, Paderborn, September 1998, Bielefeld: Transcript Verlag, 2000, pp. 35-62, 131-166

1. Introduction

As media change so also do our concepts of what constitutes knowledge. This, in a sentence, is a fundamental insight that has emerged from research over the past sixty years.[1] In the field of classics, Eric Havelock,[2] showed that introducing a written alphabet, shifting from an oral towards a written tradition, was much more than bringing in a new medium for recording knowledge. When claims are oral they vary from person to person. Once claims are written down, a single version of a claim can be shared by a community, which is then potentially open to public scrutiny, and verification.[3] The introduction of a written alphabet thus transformed the Greek concept of truth (*episteme*) and their concepts of knowledge itself. In the field of English Literature, Marshall McLuhan,[4] influenced also by historians of technology such as Harold Innis,[5] went much further to show that this applied to all major shifts in media. He drew attention, for example, to the ways in which the shift from handwritten manuscripts to printed books at the time of Gutenberg had both positive and negative consequences on our world-view.[6] In addition, he explored how the introduction of radio and television further changed our definitions of knowledge. These insights he distilled in his now famous phrase: “the medium is the message.”

Pioneers in technology, such as Vannevar Bush,[7] Douglas Engelbart,[8] and visionaries such as Ted Nelson,[9] have claimed from the outset that new media such as computers and networks also have implications for our approaches to knowledge. Members of academia and scholars have become increasingly interested in such claims, leading to a spectrum of conclusions. At one extreme, individuals such as Derrick de Kerckhove,[10] follow the technologists in assuming that the overall results will invariably be positive. This group emphasizes the potentials of collective intelligence. This view is sometimes shared by thinkers such as Pierre Lévy[11] who also warn of the dangers of a second flood, whereby we shall be overwhelmed by the new materials made available by the web.

Meanwhile, others have explored more nuanced assessments. Michael Giesecke,[12] for instance, in his standard history of printing (focussed on Germany), has examined in considerable detail the epistemological implications of printing in the fifteenth and sixteenth centuries and outlined why the advent of computers invites comparison with Gutenberg's revolution in printing. Armand Mattelart,[13] in his fundamental studies, has pointed out that the rise of networked computers needs to be seen as another step towards global communications. He has also shown masterfully that earlier steps in this process, such as the introduction of the telegraph, telephone, radio and television, were each accompanied by more global approaches to knowledge, particularly in the realm of the social sciences.

The present author has explored some implications of computers for museums,[14] libraries,[15] education[16] and knowledge in general.[17] In the context of museums seven elements were outlined: scale, context, variants, parallels, history, theory and practice; abstract and concrete; static and dynamic. Two basic aspects of these problems were also considered. First, computers entail much more than the introduction of yet another medium. In the past, each new innovation sought to replace former solutions: papyrus was a replacement for cuneiform tablets; manuscripts set out to replace papyrus and printing set out to replace manuscripts. Each new output form required its own new input method. Computers introduce fundamentally new dimensions in this evolution by introducing methods of translating any input method into any output method. Hence, an input in the form of an oral voice command can be output as a voice command (as in a tape recording), but can equally readily be printed, could also be rendered in manuscript form or potentially even in cuneiform. Evolution is embracing not replacing. Second, networked computers introduce a new cumulative dimension to knowledge. In the past, collections of cuneiform tablets, papyri, manuscripts and books were stored in libraries, but the amount of accessible knowledge was effectively limited to the size of the largest library. Hence knowledge was collected in many parts but remained limited by the size of its largest part. In a world of networked computing the amount of accessible knowledge is potentially equal to the sum of all its distributed parts.

In deference to the mediaeval tradition, we shall begin by expressing some doubts (*dubitationes*), concerning the effectiveness of present day computers. In a fascinating recent article, Classen assessed some major trends in new media[18]. He claimed that while technology was expanding exponentially, the usefulness[19] of that technology was

1000 bytes	=	1 kilobyte
1000 kilobytes	=	1 megabyte
1000 megabytes	=	1 gigabyte
1000 gigabytes	=	1 terabyte
1000 terabytes	=	1 petabyte
1000 petabytes	=	1 exobyte

Figure 1. Basic terms of size in electronic storage.

expanding logarithmically and that these different curves tended to balance each other out to produce a linear increase of usefulness with time. He concluded i) that society was keeping up with this exponential growth in technology, ii) that in order to have substantial improvements especially in education “fortunes have to be spent on R&D to get there,” and finally iii) that “we in industrial electronics research can still continue in our work, while society eagerly adopts all our results.”[20]

Dr. Classen’s review of technological progress and trends is brilliant, and we would fully accept his second and third conclusions. In terms of his first conclusion, however, we would offer an alternative explanation. He claims that the (useful) applications of computers have not kept up with the exponential expansion of technology due to inherent limits imposed by a law of usefulness. We would suggest a simpler reason: because the technology has not yet been applied. In technical terms, engineers and scientists have focussed on ISO layers 1-6[21] and have effectively ignored layer 7: applications.

Some examples will serve to make this point. Technologists have produced storage devices, which can deal with exobytes at a time (figure 1). Yet all that is available to ordinary users is a few gigabytes at a time. If I were only interested in word processing this would be more than sufficient. As a scholar, however, I have a modest collection of 15,000 slides, 150 microfilms, a few thousand books and seven meters of photocopies. For the purposes of this discussion we shall focus only on the slides. If I wished to make my 15,000 slides available on line, even at a minimal level of 1 MB per slide, that would be 15 gigabytes. Following the standards being used at the National Gallery in Washington of using 30 megabytes per image, that figure would rise to 450 gigabytes. Accordingly, a colleague in Rome, who has a collection of 100,000 slides, would need either need 100 gigabytes for a low-resolution version or 4 terabytes for a more detailed version.

In Europe museums tend to scan at 50MB/image which would raise those figures to 5 terabyes, while research institutions such as the Uffizi are scanning images at 1.4 gigabytes per square meter. At this resolution 100,000 images would make 1400 terabytes or 1.4 petabytes. There are no machines available at a reasonable price in the range of 450 gigabytes to 1.4 petabytes. The net result of this math exercise is thus very simple. As a user I cannot even begin to use the technology so it might as well not exist. There is no mysterious law of usefulness holding me back, simply lack of access to the technology. If users had access to exobytes of material, then the usefulness of these storage devices would probably go up much more than logarithmically. It might well go up exponentially and open up undreamed of markets for technology.

Two more considerations will suffice for this brief excursus on usefulness. Faced with the limitations of storage space at present, I am forced as a user to employ a number of technologies: microfilm readers, slide projectors, video players (sometimes in NTSC, sometimes in PAL), televisions, telephones, and the usual new technologies of fax machines, computers and the Internet. All the equipment exists. It is almost impossible to find all of it together working in a same place, and even if it does, it is well nigh impossible to translate materials available in one medium, to those in another medium. We are told of course that many committees around the world are busily working on the standards (e.g. JPEG, JHEG, MPEG) to make such translations among media simple and nearly automatic. In the meantime, however, all the hype in the world about interoperability, does not help me one iota in my everyday efforts as a scholar and teacher. The net result again is that many of these fancy devices are almost completely useless, because they do not address my needs. The non-compatibility of an American, a European and a Japanese device may solve someone's notion of positioning their country's technology, but it does not help users at all. Hence most of us end up not buying the latest device. And once again, if we knew that they solved our needs, their usefulness and their use might well rise exponentially.

Finally, it is worthwhile to consider the example of bandwidth. Technologists have recently demonstrated the first transmission at a rate of a terabyte per second. A few weeks ago at the Internet Summit a very senior figure working with the U.S. military reported that they are presently working with 20-40 gigabits a second and that they are confident they can reach terabyte speeds for daily operations within two years. Meanwhile, attempts by G7 pilot project five to develop demonstration centres to make the best products of cultural heritage accessible on an ATM network (a mere 622 MB/second) have been unsuccessful. A small number of persons in large cities now have access to ADSL (1.5 MB/sec), while others have access to cable modems (.5 MB/second). Even optimistic salesmen specializing in hype are not talking about having access to ATM speeds directly into the home anywhere in the foreseeable future. Hence, most persons are limited to connectivity rates of .028 or .056 MB/second, (in theory, while the throughput is usually much lower still), which is a very long way from the 1,000,000,000 MB (i.e. terabyte) that is technically possible today.

With bandwidth as with so many other aspects of technology, the simple reality is that use in real applications by actual users has not nearly kept pace with developments in technology. If no one has access to and chances to use the technology, if there are no examples to demonstrate what the technology can do, then it is hardly surprising that so-called usefulness of the technology lags behind. We would conclude therefore that there is no need to assert logarithmic laws of usefulness. If technology is truly made available, its use will explode. The Internet is a superb example. The basic technology was there in the 1960's. It was used for over two decades by a very select group. Since the advent of the World Wide Web, when it was made available to users in general, it has expanded much more each year than it did in the first twenty years of its existence.

So what would happen if all the technological advances in storage capacity, processing power, bandwidth were available for use with complete interoperability? What would change? There would be major developments in over thirty application fields (Appendix 1). Rather than attempt to examine these systematically, however, this paper will focus instead on some of the larger trends implicit in these changes. I shall assert that computers are much more than containers for recording knowledge, which can then be searched more systematically. They introduce at least seven innovations, which are transforming our concepts of knowledge. First, they offer new methods for looking at processes, how things are done, which also helps in understanding why they are done in such ways. Second, and more fundamentally, they offer tools for creating numerous views of the same facts, methods for studying knowledge at different levels of abstraction. Connected with this is a third innovation: they allow us to examine the same object or process in terms of different kinds of reality. Fourth, computers introduce more systematic means of dealing with scale, which some would associate with the field of complex systems. Fifth, they imply a fundamental shift in our methods for dealing with age-old problems of relating universals and particulars. Sixth, they transform our potential access to data through the use of meta-data. Seventh and finally, computers introduce new methods for mediated learning and knowledge through agents. This paper explores both the positive consequences of these innovations and examines some of the myriad challenges and dangers posed thereby.

2. Processes

Media also affect the kinds of questions one asks and the kinds of answers one gives to them. The oral culture of the Greeks favoured the use of What? and Why? questions. The advent of printing in the Renaissance saw the rise of How? questions. As storage devices, computers are most obviously suited to answering questions concerning biography (Who?), subjects (What?), places (Where?) and chronology (When?). But they are also transforming our understanding of processes (How?) and hence our comprehension of relations between theory and practice. In the past decades there has, for instance, been a great rise in workflow software, which attempts to break down all the tasks into clearly defined steps and thus to rationalize the steps required for the completion of a task. This atomization of tasks was time consuming, expensive and not infrequently very artificial in that it often presented isolated steps without due regard to context.

Companies such as Boeing have introduced augmented reality techniques to help understand repair processes. A worker fixing a jet engine sees superimposed on a section of the engine, the steps required to repair it. Companies such as Lockheed are going further: reconstructing an entire workspace of a ship's deck and using avatars to explain the operating procedures. This contextualization in virtual space allows users to follow all the steps in the work process.[22]

More recently companies such as Xerox[23] have very consciously developed related strategies whereby they study what is done in a firm in order to understand what can be done. In the case of her Majesty's Stationary Office, for example, they used VRML models to reconstruct all the workspaces and trace the activities on the work- floor. As a result one can examine a given activity from a variety of different viewpoints: a manager, a regular employee or an apprentice. One can also relate activities at one site with those at a number of other sites in order to reach a more global view of a firm's activities. Simulation of precisely how things are done provides insights into why they are done in that way.

In the eighteenth century, Diderot and D'Alembert attempted to record all the professions in their vast encyclopaedia. This monumental effort was mainly limited to lists of what was used with very brief descriptions of the processes. The new computer technologies introduce the possibility of a new kind of encyclopaedia, which would not only record how things were done, but could also show how different cultures perform the same tasks in slightly or even quite different ways. Hence, one could show, for instance, how a Japanese engineer's approach is different from that of a German or an American engineer. Instead of just speaking about quality one could actually demonstrate how it is carried out.

Computers were initially static objects in isolation. The rise of networks transformed their connectivity among these terminals into a World Wide Web. More recently there have been trends towards mobile or nomadic computing. The old notion of computers as large, bulky objects dominating our desks is being replaced by a whole range of new devices: laptop computers, palmtop and even wearable computers.[24] This is leading to a new vision called ubiquitous computing, whereby any object can effectively be linked to the network. In the past each computer required its own Internet Protocol (IP) address. In future, we are told, this could be extended to all the devices that surround us: persons, offices, cars, trains, planes, telephones, refrigerators and even light bulbs.

Assuming that a person wishes to be reached, the network will be able to determine whether they are at home, in their office, or elsewhere and route the call accordingly. If the person is in a meeting the system will be able to adjust its signal from an obtrusive ring to a simple written message on one's portable screen, with an option to have a flashing light in urgent cases. More elaborate scenarios will adjust automatically room temperatures, lighting and other features of the environment to the personal preferences of the individual. Taken to its logical conclusions this has considerable social consequences,[25] for it means that traditionally passive environments will be reactive to users' needs and tastes, removing numerous menial tasks from everyday life and thus leaving individuals with more time and energy for intellectual pursuits or pure diversion.

At the international level one of the working groups of the International Standards Organization (ISO/IEC JTC1/WG4) is devoted to Document Description and Processing Languages, SGML Standards Committee. At the level of G8, a consortium spearheaded by Siemens is working on a Global Engineering Network (GEN).[26] Autodesk is leading a consortium of companies to produce Industry Foundation Classes, which will effectively integrate standards for building parts such as doors and windows. Hence, when someone wishes to add a window into a design for a skyscraper, the system will "know" what kind of window is required. In future, it will be desireable to add to these foundation classes both cultural and historical dimensions such that the system can recognize the differences between a Florentine door and a Sienese door of the 1470's or some other period.

The Solution Exchange Standard Consortium (SEL) consists of 60 hardware, software, and commercial companies, which are working to create an industry specific SGML markup language for technical support information among vendors, system integration and corporate helpdesks. Meanwhile, the Pinnacles Group, a consortium which includes Intel, National Semiconductor, Philips, Texas Instruments and Hitachi, is creating an industry specific SGML markup language for semiconductors. In the United States, as part of the National Information Infrastructure (NII)[27] for Industry with Special Attention to Manufacturing, there is a Multidisciplinary Analysis and Design Industrial Consortium (MADIC), which includes NASA, Georgia Tech, Rice, NPAC and is working on an Affordable Systems Optimization Process (ASOP). Meanwhile, companies such as General Electric are developing a Manufacturing Technology Library, with a Computer Aids to Manufacturing Network (ARPA/CAMnet).[28] ESI Technologies is developing Enterprise Management Information Systems (EMIS).[29] In the automotive industry the recent merger of Daimler-Benz and Chrysler point to a new globalization. A new Automotive Network eXchange (ANX)[30]means that even competitors are sharing ideas, a process which will, no doubt, be speeded by the newly announced automotive consortium at MIT. A preliminary attempt to classify the roles of different interaction devices for different tasks has recently been made by Dr. Flaig.[31]

As Mylopoulos et al.[32] have noted, in the database world, this tendency to reduce reality to activities and data goes back at least to the Structured Analysis and Design Technique (SADT). It is intriguing to note that the quest for such an approach has a considerable history. In the United States, where behaviorism became a major branch of psychology, Charles S. Pierce claimed that: "The only function of thinking is to produce acting habits."[33] Such ideas have been taken up particularly in Scandinavia. For instance,
Sarvimäki (1988),[34] claimed that there is a continuous interaction between knowledge and action; that knowledge is created through and in action. These ideas have more recently been developed by Hjørland (1997).[35] Some would see this as part of a larger trend to emphasize subjective dimensions of reality in terms either of purpose (Hjelmslev)[36] or interest (Habermas).[37] Meanwhile, Albrechtsen and Jacob (1998),[38] have attempted to analyse work from a descriptive rather than a normative point of view. Building on the ideas of Davenport,[39] Star[40] and Law,[41] they have outlined an activity theory in terms of four types of work, namely, industrialized bureaucratically regulated work, learning network organization, craft type of individualised work and semi-independent market-driven result units.

If activities are seen as one aspect of the human condition such an activities based approach makes perfect sense. If, however, such activities are deemed to be the sole area to be studied, then one encounters the same problems familiar with a number of Marxist theoreticians. While claiming that reality must be reduced to the visible dimensions of practical, physical activities, they wish, at the same time, to create a conceptual, theoretical framework which goes beyond those very limits on which they insist.

Software Defined	User Defined
Base Program	Meta Program
Base Interface	Meta Interface

Figure 2. Separation of basic software from user defined modalities through meta-object protocols in programming.

3. Views and Levels of Abstraction

One of the fundamental changes brought about by computers is increasingly to separate our basic knowledge from views of that knowledge. Computer scientists refer to such views as contextualization, and describe them variously in terms of modules, scopes and scope rules.[42] The importance of these views has increased with the shift towards conceptual modelling.[43] In the case of earlier media such as cuneiform, manuscripts and books, content was irrevocably linked with a given form. Changing the form or layout required producing a new edition of the content. In electronic media this link between form and content no longer holds. Databases, for instance, separate the content of fields from views of that content. Once the content has been input, it can be queried and displayed in many ways without altering the content each time. This same principle applies to Markup Languages for use on the Internet. Hence, in the case of Standard Generalized Markup Language (SGML) and Extensible Markup Language (XML), the rules for content and rules for display are separate. Similarly in the case of programming, the use of meta-object protocols is leading to a new kind of open implementation whereby software defined aspects are separated from user defined aspects (figure 2). An emerging vision of network computers, foresees a day when all software will be available on line, and users will need only to state their goals to find themselves with the personally adapted tools. Linked with this vision are trends towards reusable code.[44]

Related to the development of these different views of reality, is the advent of spreadsheets and data-mining techniques, whereby one can look at the basic facts in a database from a series of views at different levels of abstraction. Once a bibliography exists as a database, it is easy to produce graphs relating publications to time, by subject, by city, country or by continent. In the past any one of these tasks would have comprised a separate study. Now they are merely a different "view."

One of the serious problems in the new electronic methods is that those designing the systems are frequently unfamiliar with the complexities of historical knowledge. An excellent case in point is the entity-relationship model, developed by Chen,[45] which is the basis of most relational databases and object-oriented approaches. On the surface it is very simple. It assumes that every entity has various relationships in terms of attributes. Accordingly a person has attributes such as date of birth, date of death and profession. In the case of modern individuals this is often sufficient. In historical cases, however, the situation may be much more complex. For instance, there are at least five different theories about the year in which the painter Titian died, so we need not only these varying dates but also the sources of these claims. Although entity-relationship models do not cope with this, other systems with conceptual modelling do. We need new attention to the often, implicit presuppositions[46] underlying software and databases and we need to bring professionals in the world of knowledge organisation up to date concerning the developments in databases.

4. Scale

These developments in views and different levels of abstraction are also transforming notions of scale. Traditionally every scale required a separate study and even a generation ago posed serious methodological problems.[47] The introduction of pyramidal tiling[48] means that one can now move seamlessly from a satellite image of the earth (at a scale of 1:10,000,000) to a life-size view of an object and then through a spectrum of microscopic ranges. These innovations are as interesting for the reconstruction of real environments such as shopping malls and tourist sites as they are for the creation of virtual spaces such as Alpha-World[49]. Conceptually it means that many more things can be related. Systematic scales are a powerful tool for contextualization of objects.

These innovations in co-ordinating different scales are particularly evident in fields such as medicine. In Antiquity, Galen's description of medicine was limited mainly to words. These verbal descriptions of organs were in general terms such that there was no clear distinction between a generic description of a heart and the unique characteristics of an individual heart. Indeed the approach was so generic that the organs of other animals such as a cow were believed to be interchangeable with those of an individual.

During the Renaissance, Leonardo added illustrations as part of his descriptive method. Adding visual images to the repertoire of description meant that one could show the same organ from a number of different viewpoints and potentially show the difference between a typical sample and an individual one. However, the limitations of printing at the time made infeasible any attempts to record all the complexities of individual differences.

Today, medicine is evolving on at least five different levels. The GALEN project is analysing the basic anatomical parts (heart, lung, liver etc.) and systematically studying their functions and inter-relationships at a conceptual level. The Visible Human project is photographing the entire human body in terms of thin slices, which are being used to create Computer Aided Design (CAD) drawings at new levels of realism. In Germany, the Medically Augmented Immersive Environment (MAIE), developed by the Gesellschaft für Mathematik und Datenverarbeitung (GMD) and three Berlin hospitals, dedicated to radiology (Virchow), pathology (Charité) and surgery (RRK) respectively, are developing models for showing structural relations among body parts in real time. This system includes haptic simulation based on reconstructed tomographic scans. Other projects are examining the human body at the molecular and atomic level (figure 3). At present these projects are evolving in tandem without explicit attempts to co-relate them. A next step will lie in integrating all this material such that one can move at will from a macroscopic view of the body to a study of any microscopic part at any desired scale.

Conceptual	GALEN[50]
Physical	Visible Human[51]
Structural	OP 2000 Medically Augmented Immersive Environment (MAIE)[52]
Molecular	Bio-Chemical
Atomic	Human Genome[53]

Figure 3. Different levels of scale in the study of contemporary medicine.

In the past, anatomical textbooks typically provided doctors with a general model of the body and idealized views of the various organs. The Virtual Human is providing very detailed information concerning individuals (three to date), which can then serve as the basis for a new level of realism in making models. These models can then be confronted with x-rays, ultra-sound and other medical imaging techniques, which record the particular characteristics of individual patients.

Elsewhere, in the Medical Subject Headings (MeSH) project, a semantic net includes five relationship classes: identity, physical, spatial, conceptual and functional, with tree category groupings for anatomic spaces, anatomic structural systems, anatomic substances and diseases.[54] Potentially such projects could lead to a systematic linking of our general knowledge about universals and our specialized knowledge about particulars (see section 7 below).

A somewhat different approach is being taken in the case of the human genome project. Individual examples are studied and on the basis of these a “typical model” is produced, which is then used as a set of reference points in examining other individual examples. Those deviating from this typical model by a considerable amount are deemed defective or aberrant, requiring modification and improvement. A danger in this approach is that if the parameters of the normal are too narrowly defined, it could lead to a new a version of eugenics seriously decreasing the bio-diversity of the human race.[55] If we are not careful we shall succumb to believing that complexity can be resolved through the regularities of universal generalizations rather than in the enormously varying details of individuals. Needed is a more inductive approach, whereby our models are built up from the evidence of all the variations.

5. Kinds of Reality

Another important way in which computers are changing our approach to knowledge relates to new combinations of reality. In the 1960’s the earliest attempts at virtual reality created a) digital copies of physical spaces, b) simplified digital subsets of a more complex physical world or c) digital visualizations of imaginary spaces. These alternatives tended to compete with one another. In the latter 90’s there has been a new trend to integrate different versions of reality to produce both augmented reality and augmented virtuality. As a result one can, for instance, begin with the walls of a room, superimposed on which are the positions of electrical wires, pipes and other fixtures.

Reality (Nature, Man Made World)	
Virtual Reality	Sutherland, Furness
Augmented Reality	Feiner, Stricker
Augmented Virtuality	Mankoff[56]
Double Augmented Reality[57]	Gelernter, Ishii

Figure 4. Basic classes of simulated reality and their proponents.

Such combinations have enormous implications for training and repair work of all kinds. Recently, for instance, a Harvard medical doctor superimposed an image of an internal tumour onto the head of a patient and used this as an orientation method for the operation. (This method is strikingly similar to the supposedly science fiction operation of the protagonist's daughter in the movie *Lost in Space*). As noted elsewhere, this basic method of superimposition can also be very fruitful in dealing with alternative reconstructions of an ancient ruin or different interpretations of a painting's spatial layout. Other alternatives include augmented virtuality, in which a virtual image is augmented and double augmented reality in which a real object such a refrigerator has superimposed on it a virtual list which is then imbued with further functions.[58] (cf. figure 4).

Other techniques are also contributing to this increasing interplay between reality and various constructed forms thereof. In the past, for instance, Computer Aided Design (CAD) and video were fundamentally separate media. Recently Bell Labs have introduced the principle of pop-up video, which permits one to move seamlessly between a three-dimensional CAD version of a scene and the two-dimensional video recording of an identical or at least equivalent scene.[59] Meanwhile, films such as *Forrest Gump* integrate segments of "real" historical video seamlessly within a purely fictional story. This has led some sceptics to speak of the death of photographic veracity,[60] which may well prove to be an overreaction. Major bodies such as the Joint Picture Expert Group (JPEG) are working on a whole new framework for deciding the veracity of images, which will help to resolve many of these fears.

On the positive side, these developments in interplay among different kinds of reality introduce immense possibilities for the re-contextualization of knowledge. As noted earlier, while viewing images of a museum one will be able to move seamlessly to CAD reconstructions of the rooms and to videos explaining particular details. One will be able to move from a digital photograph of a painting, through images of various layers of the painting to CAD reconstructions of the painted space as well as x-rays and electron-microscope images of its micro-structures. One will be able to study parallels, and many aspects of the history of the painting. A new integration of static and dynamic records will emerge.

6. Complex Systems

The systematic mastery of scale in the past decades has lent enormous power to the zoom metaphor, to such an extent that one could speak of Hollywoodization in a new sense.

Reality is seen as a film. The amount of detail, the granularity, depends on one's scale. As one goes further one sees larger patterns, as one comes closer one notices new details. Proponents of complex systems such as Yaneer Bar-Yam,[61] believe that this zoom metaphor can serve as a tool for explaining nearly all problems as one moves from atomic to molecular, cellular, human and societal levels. Precisely how one moves from physical to conceptual levels is, however, not explained in this approach.

Complex systems entail an interdisciplinary approach to knowledge, which builds on work in artificial neural networks to explain why the whole is more than the sum of its parts. The director of the New England Center for Complex Systems (NECSI) believes that this approach can explain human civilization:

> One system particularly important for the field of complex systems is human civilization the history of social and economic structures and the emergence of an interconnected global civilization. Applying principles of complex systems to enable us to gain an understanding of its past and future course is ultimately an important objective of this field. We can anticipate a time when the implications of economic and social developments for human beings and civilization will become an important application of the study of complex systems.[62]

Underlying this approach is an assumption that the history of civilization can effectively be reduced to a history of different control systems, all of them hierarchically structured. This may well provide a key to understanding the history of many military, political and business structures, but can hardly account for the most important cultural expressions. If anything the reverse could well be argued. Greece was more creative than other cultures at the time because it imposed less hierarchical structures on artists. Totalitarian regimes, by contrast, typically tolerate considerably less creativity, because most of these expressions are invariably seen as beyond the parameters of their narrow norms. Hence, complex systems with their intriguing concepts of emergence, may well offer new insights into the history of governments, corporations, and other bureaucracies. They do not address a fundamental aspect of creativity, which has to do with the emergence of new individuals and particulars, non-controlled elements of freedom, rather than products of a rule based system.

7. Individuals and Particulars

As was already suggested above, one of the central questions is how we define knowledge. Does knowledge lie in the details of particulars or in the universals based on those details? The debate is as old as knowledge itself. In Antiquity, Plato argued for universals: Aristotle insisted on particulars. In the Middle Ages, the debate continued mainly in the context of logic and philosophy. While this debate often seemed as if it were a question of either/or, the rise of early modern science made it clear that the situation is more complex. One needs particular facts. But in isolation these are merely raw data. Lists of information are one step better. Yet scientific knowledge is concerned with laws, which are effectively summaries of those facts. So one needs both the particulars as a starting point in order to arrive at more generalized universals, which can then explain the particulars in question.

Each change in media has affected this changing relationship between particulars and universals. In pre-literate societies, the central memory unit was limited to the brain of an individual and oral communication was limited to the speed with which one individual could speak to another. The introduction of various written media such as cuneiform, parchment, and manuscripts meant that lists of observations were increasingly accessible. Printing helped to standardize this process and introduced the possibility of much more systematic lists. The number of particular observations on which universal claims and laws could be established thus grew accordingly. While there were clearly other factors such as the increased accuracy of instruments, printing made Tycho Brahe's observations more accessible than those made at the court of Alphonse the Wise and played their role in making Kepler's new planetary laws more inclusive and universal.

The existence of regular printed tables greatly increased the scope of materials, which could readily be consulted. It still depended entirely on the memory and integrating power of the individual human brain in order to recognize patterns in the data and to reach new levels of synthesis. Once these tables are available on networked computers, the memory capacities are expanded to the size of the computer. The computer can also be programmed to search both for consistencies and anomalies. So a number of the pattern discoveries, which depended solely on human perception, can now be automated and the human dimension can be focussed on discerning particularly subtle patterns and raising further questions.

In the context of universities, the arts and sciences have traditionally been part of a single faculty. This has led quite naturally to many comparisons between the arts and the sciences, and even references to the art of science or the science of art in order to emphasize their interdependence. It is important to remember, however, that art and science differ fundamentally in terms of their approach to universals and particulars. Scientists gather and study particulars in order to discern some underlying universal and eternal pattern. Artists gather and study examples in order to create a particular object, which is unique, although it may be universal in its appeal. Scientists are forever revising their model of the universe. Each new discovery leads them to discard some piece or even large sections of their previous attempt. Notwithstanding Newton's phrase that he was standing on the shoulders of giants, science is ultimately not cumulative in the sense of keeping everything of value from an earlier age. Computers, which are only concerned with showing us the latest version of our text or programme, are a direct reflection of this scientific tradition.[63]

In this sense, art and culture are fundamentally different in their premises. Precisely because they emphasize the uniqueness of each object, each new discovery poses no threat to the value of what came before. Most would agree, for instance, that the Greeks introduced elements not present in Egyptian sculpture, just as Bernini introduced elements not present in Michelangelo, and he, in turn, introduced elements not present in the work of Donatello. Yet it would be simplistic to deduce from this that Bernini is better than Michelangelo or he in turn better than Donatello. If later were always better it would be sufficient to know the latest artists' work in the way that scientists feel they only need to know the latest findings of science. The person who knows about the Egyptians, Greeks, Donatello, Michelangelo and Bernini is much richer than one who knows only the latest phase. Art and culture are cumulative. The greatest scientist succeeds in reducing the enormity of particular instances to the fewest number of laws which to the best of their knowledge are unchanging. The most cultured individual succeeds in bringing to light the greatest number of unique examples of expression as proof of creative richness of the human condition. These differing goals of art and science pose their own challenges for our changing understanding of knowledge.

Before the advent of printing, an enterprising traveller might have recorded their impressions of a painting, sculpture or other work of art, which they encountered in the form of a verbal description or at best with a fleeting sketch. In very rare cases they might have attempted a copy. The first centuries after Gutenberg saw no fundamental changes to this procedure. In the nineteenth century, lithographs of art gradually became popular. In the late nineteenth century, black and white photographs made their debut.[64] In the latter part of the twentieth century colour images gradually became popular.

Even so it is striking to what extent the horizons of authors writing on the history of their subject remained limited to the city where they happened to be living. It has often been noted, for example, that Vasari's *Lives of the Artists*, focussed much more on Florence than other Italian cities such as Rome, Bologna, Milan or Urbino. At the turn of the century, art historians writing in Vienna tended to cite examples found in the Kunsthistorisches Museum, just as others since living in Paris, London or New York have tended to focus on the great museum that was nearest to home. The limitations of printing images meant that they could give only a few key masterpieces by way of example. From all this emerged a number of fascinating glimpses into the history of art, which were effectively summaries of the dominant taste in the main halls of the great galleries. It did not reflect the up to 95% of collections that is typically in storage. Nor did it provide a serious glimpse of art outside the major centres.

A generation ago scholars such as Chastel[65] pointed to the importance of studying the smaller cities and towns in the periphery of such great cities: to look not only at Milan but also at Pavia, Crema, Cremona, Brescia and Bergamo. Even so, in the case of Italy, for instance, our picture is still influenced by Vasari's emphases from over four centuries ago. Everyone knows Florence and Rome. But who is aware of the frescoes at Bominaco or Subiaco, of the monasteries at Grottaferata and Padulo, or the architecture of Gerace, Urbania or Asolo? The art in these smaller centres does not replace, nor does it even pretend to compete, with the greatest masterpieces which have usually made their way to the world's chief galleries. What they do, however, is to provide us with a much richer and more complex picture of the variations in expression on a given theme. In the case of Piero della Francesca, for example, who was active for much of his life in San Sepolcro, Arezzo and Urbino, we discover that these masterpieces actually originated in smaller centres although they are now associated with great cities (London, Paris, Florence). In other cases we discover that the smaller centres do not simply copy the great masterpieces. They adapt familiar themes and subjects to their own tastes. The narrative sequences at San Gimignano, Montefalco, Atri add dimensions not found even in Florence or Rome.

To be sure some of this richness has been conveyed by the medium of printing, through local guidebooks and tourist brochures. However, in these the works of art are typically shown in isolation without any reference to more famous parallels in the centres. Computers will fundamentally change our approach to this tradition. First they will make all these disparate materials accessible. Hence a search for themes such as *Virgin and Child* will not only bring up the usual examples by Botticelli or Raphael but also those in museums such as L'Aquila, Padua, and Volterra (each of which were centres in a previous era). Databases will allow us to study developments in terms of chronology as well as by region and by individual artist. Filtering techniques will allow us to study the interplay of centre and periphery in new ways.

More importantly, we shall be able to trace much more fully the cumulative dimensions of culture, retaining the uniqueness of each particular object. In the past, each of the earlier media precluded serious reproductions of the original objects. As noted above, colour printing has only been introduced gradually over the past half-century. Even then, a single colour image of a temple or church, can hardly do justice to all its complexities. The advent of virtual and augmented reality, and the possibility of stereo-lithographic printing, means that a whole new set of tools for understanding culture is emerging. They will not replace the value and sometimes the absolute necessity of studying some of the originals in situ, but if we always had to visit everything, which we wished to study in its original place, the scope of our study would be very limited indeed.

Earlier media typically meant that one emphasized one example often forgetting that it represented a much larger phenomenon. The Coliseum in Rome is an excellent case in point. History books typically focus on this amphitheatre and tell us nothing of the great number of amphitheatres spread throughout the Roman empire. Networked computers can make us aware of all known examples from Arles and Nîmes in France to El-Djem in Tunisia and Pula in Croatia. This new encyclopaedic approach means that we shall have a much better understanding of how a given structure spreads throughout a culture to form a significant element in our cultural heritage such as the Greek temple, the Romanesque and Gothic Church, and the Renaissance villa. It means that we shall also have a new repertoire of examples for showing even as these styles spread, each new execution of the principle introduces local uniqueness. Hence the cathedrals at St. Denis, Chartres, Notre Dame, Cologne, Magdeburg, Bamberg, Ulm and Burgos are all Gothic, and yet none is a simple copy of the other.

A generation ago when Marshall McLuhan coined the phrase "the global village", some assumed that the new technologies would invariably take us in the direction of a world where every place was more or less the same: where Hiltons and McDonalds would spread throughout an increasingly homogenized planet. This danger is very real. But as critical observers such as Barber have noted,[66] the new technologies have been accompanied by a parallel trend in the direction of regionalism and new local awareness.

The same technologies, that are posing the possibility of global corporations, are introducing tremendous new efforts in the realms of citizen participation groups and of local democracy. Networked computers may link us together with persons all over the world as if we were in a global village but this does not necessarily mean that every village has to look the same. Indeed, the more the mass-media try to convince us that we are all inhabitants of a single interdependent ecosystem, the more individuals are likely to articulate how and even why their particular village is different from others. In this context, the new access to individuals and particulars introduced by networked computers, becomes much more than an interesting technological advance. It provides a key to maintaining the cultural equivalent of bio-diversity, which is essential for our well being and development in the long run.

In themselves the particulars are, of course, only lists and as such merely represent data or, at best, information. Hence they should be seen as starting points rather than as results per se. Their vital importance lies in vastly increasing the sample, the available sources upon which we attempt to draw conclusions. The person who has access to only one book in art history will necessarily have a much narrower view than someone who is able to work with the resources of a Vatican or a British Library. In the past, scholars have often spent much more time searching for a document than actually reading it. In future, computers will greatly lighten the burden of finding. Hence, scholarship will focus increasingly on determining the veracity of sources, weighing their significance, interpreting and contextualizing sources, and learning to abstract from the myriad details which they offer, some larger patterns of understanding. Access to new amounts of particulars will lead to a whole new series of universal abstractions.

Implicit in the above discussion are larger issues of knowledge organization that go far beyond the scope of this paper. We noted that while the arts and science typically share the same faculty and are in many ways interdependent, there are two fundamental ways in which they differ. First, the sciences examine individual facts and particulars in order to arrive at new universal summaries of knowledge. The arts, by contrast, are concerned with creating particulars, which are unique in themselves. They may be influenced by or even inspired by other particular works, but they are not necessarily universal abstractions in the way that the sciences are. Second, and partly as a result thereof, the sciences are not cumulative in the same way that the arts and culture are. In the sciences only the latest law, theory, postulate etc. is what counts. In the arts, by contrast, the advent of Picasso does not make Rubens or Leonardo obsolete, any more than they made Giotto or Phidias obsolete. The arts and culture are defined by the cumulative sum of our collective heritage, all the particulars collected together, whereas the sciences are concerned only with the universals abstracted from the myriad particulars they examine.[67]

It follows that, while both the arts and sciences have a history, these histories ultimately need to be told in very different ways. In the arts, that history is about how we learned to collect and remember more and more of our past. Some scholars have claimed, for instance, that we know a lot more about the Greeks than Aristotle himself. In the sciences, by contrast, that history is at once about how scientists developed ever better instruments with which to make measurable that which is not apparent to the naked eye, and how they used the results of their observations to construct ever more generalized, universal, and at the same time, testable theories. To put it simply, we need very different kinds of histories to reflect these two fundamentally different approaches to universals and particulars, which underlie fundamental differences between the arts and sciences. With the advent of networked computers the whole of history needs to be rewritten: at least twice, a process that will continue in future.

8. Now and Eternity

Not unrelated to the debates concerning particulars and universals are those connected with the (static) fine arts versus (semi-dynamic) arts such as sculpture and architecture[68] and (dynamic) performance arts such as dance, theatre, and music. Earlier media such as manuscripts or print were at best limited to static media. They could not hope to reproduce the complexities of dynamic performance arts. Even the introduction of video offered only a partial solution to this challenge, insomuch that it reduced the three-dimensional field to a particular point of view reduced to a two-dimensional surface. Hence, if a video captured a frontal view of actors or dancers their backs were necessarily occluded. These limitations of recording media have led perforce to a greater emphasis on the history of fine arts such as painting than on the semi-dynamic arts such as sculpture and architecture or the dynamic arts such as dance and theatre.[69]

These limitations have had both an interesting and distorting effect on our histories of culture. It has meant, for instance, that we traditionally knew a lot more about the history of static art than dynamic art: a lot more about painting than about dance, theatre or music. It has meant that certain cultures such as the Hebrew tradition, which emphasize the now of dynamic dance and music over the eternal static forms of sculpture and painting were under-represented in traditional histories of culture. Conversely, it has meant that the recent additions of film, television, video and computers have focussed new attention on the dynamic arts, to the extent of undermining our appreciation of the enduring forms. Our visions of eternal art are being replaced by a new focus on the now.

From a more global context these limitations have also had a more general, subtle, impact on our views of world culture. Those strands, which focussed on the static, fine arts were considered the cornerstones of world cultural development. Since this was more so in the West (Europe, the Mediterranean and more in recently North America), sections of Asia Minor (Iran, Iraq, Turkey), and certain parts of the Far East (China, Japan and India),[70] these dominated our histories of art. Countries with strong traditions of dance, theatre and other types of performance (including puppet theatre, shadow theatre and mime) such as Malaysia, Java and Indonesia were typically dismissed as being uncultured. The reality of course was quite different. What typically occurred is that these cultures took narratives from static art forms such as literature and translated them into dynamic forms. Hence, the stories of an Indian epic, the *Ramayana*, made their way through Southeast Asia in the form of theatre, shadow puppet plays, dances and the like. Scholars such as Mair[71] have rightly drawn attention to the importance of these performance arts (figure 5).

Narrative Performance	Country
Etoki	Japan
Par	India
Parda Da	Iran
Pien Wen	China
Waysang Beber	Malaysia

Figure 5. Examples of narrative based performance art in various countries.

Ultimately, however, the challenge goes far beyond simple dichotomies of taste, namely, whether one prefers the static, eternal arts of painting to the dynamic, now, arts of dance and music. A more fundamental challenge will lie in re-writing the whole of our history of art and culture to reflect how these seeming oppositions have in fact been complementary to one another. In the West, for instance, we know that much Renaissance and Baroque art was based directly on Ancient mythology either directly via books such as Ovid's *Metamorphoses,* or indirectly via Mediaeval commentaries on these myths. We need a new kind of hyper-linking to connect all these sources with the products, which they inspired. Such hyperlinks will be even more useful in the East where a same mythical story may well be translated into half a dozen art forms ranging from static (scupture and painting) to dynamic (dance, mime, shadow theatre, puppet theatre, theatre). From all this there could emerge new criteria for what constitutes a seminal work: for it will become clear that a few texts have inspired works over the whole gamut of cultural expression. The true key to eternal works lies in those which affect everything from now to eternity.

9. Meta-Data

How is the enormity of this challenge to be dealt with in practice? It is generally assumed that meta-data offers a solution. The meta concept is not new. It played a central role in the meta-physics of Aristotle. In the past years with the rise of networked computing, meta has increasingly become a buzzword. There is much discussion of meta-data, meta-databases, and meta-data dictionaries. There is a Metadata Coalition,[72] a Metadata Council[73] and even a Metadata Review.[74] Some now speak of meta-meta data in ways reminiscent of those who spoke of the meaning of meaning a generation ago.

The shift in attention from data to meta-data[75] and meta-meta-data is part of a more fundamental shift in the locus of learning in our society. In Antiquity, academies were the centres of learning and repositories of human knowledge. In the Latin West, monasteries became the new centres of learning and remained so until the twelfth century, when this locus began to shift towards universities. From the mid-sixteenth to the mid-nineteenth centuries universities believed they had a near monopoly on learning and knowledge. Then came changes. First, there was a gradual shift of technical subjects to polytechnics. New links between professional schools (e.g. law, business) and universities introduced more short-term training goals while also giving universities a new lease on life.

The twentieth century brought corporate universities of which there are now over 1,200. It also brought national research centres (NRC, CNR, GMD), military research laboratories (Lawrence Livermore, Los Alamos, Argonne, Rome), specialized institutes (such as Max Planck and Fraunhofer in Germany) and research institutes funded by large corporations (AT&T, General Motors, IBM, Hitachi, Nortel). Initially the universities saw themselves as doing basic research. They defined and identified the problems the practical consequences of which would then be pursued by business and industry. In the past decades all this has changed. The research staffs of the largest corporations far outnumber those of the greatest universities. AT&T's Lucent Technologies has 24,000 in its Bell Laboratories alone and some 137,000 in all its branches. Hitachi has over 34,000, i.e. more researchers than the number of students at many universities. Nortel has over 17,000 researchers. The cumulative information produced by all these new institutions means that traditional attempts to gather (a copy of) all known knowledge and information in a single location is no longer feasible. On the other hand a completely distributed framework is also no longer feasible. A new framework is needed and meta-data seems to be a new holy grail. To gain some understanding of this topic and the scope of the international efforts already underway will require a detour that entails near lists of information. Those too impatient with details are invited to skip the next twelve pages at which point we shall return to the larger framework and questions.

It is generally accepted that meta-data is data about data,[76] or key information about larger bodies of information. Even so discussions of meta-data are frequently confusing for several reasons. First, they often do not define the scope of information being considered. In Internet circles, for instance, many authors assume that meta-data refers strictly to Internet documents, while others use it more generally to include the efforts of publishers and librarians. Secondly, distinctions need to be made concerning the level of detail entailed by the meta-data. Internet users, for instance, are often concerned only with the most basic information about a given site. In extreme cases, they believe that this can be covered through Generic Top Level Domain Names (GTLD), while publishers are convinced that some kind of unique identifying number will be sufficient for these purposes (see figure 6). Present day search engines such as Altavista, and Lycos also use a minimal approach to these problems, relying only on a title and a simple tag with a few keywords serving as the metadata.

Others, particularly those in libraries, feel that summary descriptions, full library catalogue descriptions or methods for full text descriptions are required. Meanwhile some are convinced that while full text analysis or at least proper cataloguing methods are very much desireable, it is not feasible that the enormity of materials available on the web can be subjected to rigorous methods requiring considerable professional training. For these the Dublin Core is seen as a pragmatic compromise (figure 7).

Basic Description		
Internet/Computer		
	Generic Top Level Domain Names	GTLD[77]
	Hypertext Transfer Protocol	http
	Multipurpose Internet Mail Exchange	MIME
	Uniform Resource Name	URN
	Uniform Resource Locator	URL[78]
Int. Standards Org.		ISO
	International Standard Book Numbering, ISO 2108:1992	ISBN[79]
	International Standard Music Number, ISO 10957:1993	ISMN[80]
	International Standard Technical Report Number	ISRN[81]
	Formal Public Identifiers	FPI[82]
Nat.Info.Standards Off.		NISO
	Serial Item and Contribution Identifier	SICI
	International Standard Serials Number	ISSN[83]
Publishers		
	Conf. Int. des Sociétés d'Auteurs et Compositeurs	CISAC[84]
	Common Information System	CIS
	International Standard Works Code	ISWC
	Works Information Database	WID
	Global and Interested Parties Database	GIPD
	International Standard Audiovisual Number	ISAN[85]
	International Federation of the Phonogram Industry	IFPI
	International Standard Recording Code	ISRC[86]
	Cf. Other Standard Identifier	OSI[87]
	Universal Product Code	UPC
	International Standard Music Number	ISMN
	International Article Number	IAN
	Serial Item and Contribution Identifier	SICI
	Elsevier Publisher Item Identifier	PII[88]
	Corp.for Nat.Res.Initiatives: Digital Object Identifier	DOI[89]
Libraries		
	Persistent Uniform Resource Locator	PURL[90]
	Handles	
Universities		
	Uniform Object Identifier	UOI[91]
	Object ID	OID

Figure 6. Major trends in meta-data with respect to basic identification.

Summary Description		
Internet/ W3		
	Hyper Text Markup Language: Header	HTML Header
	META Tag[92]	
	Hyper Text Markup Language Appendage	HTML Appendage
	Resource Description Format	RDF
	Extensible Markup Language	XML
	Protocol for Internet Content Selection	PICS
	Uniform Resource Identifier	URI
	Uniform Resource Characteristics	URC
	Universally Unique Identifiers[93]	UUID
	Globally Unique Identifiers	GUID
	Whois++ Templates	
	Internet Anonymous FTP Archives Templates	IAFA[94]
	Linux Software Map Templates	LSM
	Harvest Information Discovery and Access System	
	Summary Object Interchange Format	SOIF[95]
	Netscape Meta Content Framework	MCF[96]
	Microsoft Web Collections[97]	
Libraries		
	International Federation of Library Associations	IFLA [98]
	International Standard Bibliographic Description	ISBD[99]
	Electronic Records	ISBD (ER)
	Dublin Core	DC
	Resource Org.and Discovery in Subject Based Services	ROADS
	Social Science Information Gateway	SOSIG
	Medical Information Gateway	OMNI[100]
	Art, Design, Architecture, Media	ADAM
Full Description	**(Library Catalogue Record)**	
Libraries		
	Z.39.50	
	Machine Readable Record[101]: many national variants	MARC[102]
	Other Catalogue formats summarized in Eversberg [103]	e.g. PICA, MAB
Full Text		
Libraries and Museums		
	Standard Generalized Markup Language	SGML[104]
	Library of Congress Encoding Archival Description	LC EAD[105]
	Text Encoding Initiative	TEI
	Consortium for Interchange of Museum Information	CIMI

Figure 7. Major trends in meta-data with respect to more complete description.

As can be inferred from the lists above, there are a great number of initiatives with common goals, often working in isolation, sometimes even ignorant of the others' existence. Nonetheless, a number of organizations are working at integrated solutions for meta-data. What follows is by no means comprehensive. Gilliland-Swetland, for instance, has recently identified five different kinds of metadata: administrative, descriptive, preservation, technical and use.[106] We shall begin by examining four crucial players. While presented separately, it is important to recognize that there are increasing synergies between/among these players and their solutions, which are to a certain extent competing with one another.

i) Internet Engineering Task Force (IETF)

The IETF, which is directly linked with the Internet Society, is active on a great number of fronts. At present, sites on the World Wide Web typically have a Uniform Resource Locator (URL). These suffer from at least two basic problems: i) they often change location and ii) there may be several mirror sites for the same material. The IETF has been working on a more comprehensive approach:

> Resources are named by a URN (Uniform Resource Name), and are retrieved by means of a URL (Uniform Resource Locator). Describing the resource for purposes of discovery, as well as making the binding between a resource's name and its location(s) is the role of the URC (Uniform Resource Characteristic).
> The purpose or function of a URC is to provide a vehicle or structure for the representation of URIs [Uniform Resource Indicators] and their associated meta-information.[107]

The precise meaning of these terms is not as clear as one might wish. Weider,[108] for instance calls Universal Resource Names (URNs)[109] the equivalent of an ISBD number for electronic resources, whereas Iannella calls them a naming method. As for Universal Resource Characteristics (URC), Iannella calls them meta-data, whereas Ron Daniels[110]gives them quite a different take. Similarly, the exact nature and function of the Uniform Resource Indicators (URI) has been the subject of considerable debate and at a meeting in Stockholm (September 1997), the IETF URI committee was officially disbanded. Subsequently, the W3 Consortium has taken up the problem (see below). Meanwhile, URNs still need to be mapped back to a series of disparate URLs. To this end the IETF is exploring at least four methods of URN to URL Mapping (Resource Discovery) and URC[111] using http:

i) Domain Name Server (dns)[112]
ii) x-Domain Name Server 2 (x-dns-2) with trivial URC syntax[113]
iii) SGML designed to interoperate with the trivial URC scenario[114]
iv) Path, same as 2 above except that it is hierarchically arranged.[115]

A fifth method, Handle, is being explored by ARPA. Ultimately the technical details of these competing schemes is less important than the result that they promise: a framework which will allow various sources to interoperate. It is noteworthy that institutions around the world are working on these challenges. The Distributed Technology Centre (DSTC) in Brisbane has a Basic URN Service for the Internet (BURNS) project,[116] and The URN Interoperability Project[117] (TURNIP), while Earth Observation at the Joint Research Centre (JRC) has an URN Resolver Experiment[118] as part of its European Wide Service Exchange (EWSE) initiative. Meanwhile the IETF, is exploring Uniform Resource Agents (URA's) [119]: "as a means of specifying composite net-access tasks. Tasks are described as "composite" if they require the construction and instantiation of one or more Uniform Resource Locators (URL's) or Uniform Resource Names (URN's), and/or if they require transformation of information returned from instantiating URL's/URN's."[120]

Precisely, how all these initiatives should be integrated is still a matter of conjecture. For example, the Internet Anonymous File Transfer Protocol Archives Working Group (IAFA),[121] initially worked on Templates for Internet data. This became a new group called Integration of Internet Information Resources Working Group (IIIR).[122] This group also worked toward Query Routing Protocol (QRP), which they abandoned in favour of working on a Structured Text Interchange Format (STIF).[123] More significantly, they also set out to integrate WAIS, ARCHIE, and Prospero into a Virtually Integrated Information Service (VUIS). To this end they introduced four Requests for Comments.[124] Of these, the Integrated Internet Information Service (IIIS) foresees the integration of some of the major types of information used on the internet (figure 8):

Another attempt by the IETF at creating an integrated strategy for meta-data on the internet is their Common Indexing Protocol[125] (CIP), which foresees a combination of four elements: a client, a protocol for the front-end, an indexing object and a database backend or query protocol (figure 9, cf. Appendix 2 which provides a glossary of some of key technical terms). While undoubtedly essential, such attempts are focussed mainly on information available on the Internet and do not yet address more complex challenges of other knowledge available in museums and libraries.

Gopher	WAIS	WWW	Archie	Others
Resource discovery system perhaps based on Whois++				
Uniform Resource Name to Uniform Resource Locator Mapping System based on Whois++ or X.500				
	Transponder	Transponder	Transponder	
	Resource	Resource	Resource	

Figure 8. Basic Scheme from RFC 1727 showing how various protocols would be integrated using Whois++ and X.500.

Client	Protocol Front End	Indexing Object	Database Backend or Query Protocol
Whois++	Whois++	SQL or Indexer API	Z39.50
PH	PH		or GNU DBM
LDAP	LDAP		or…GDBM

Figure 9. Basic Scheme concerning Common Indexing Protocol (CIP).

Work is also progressing on an Application Configuration Access Protocol (ACAP, RFC 2244).[126] Meanwhile other groups within the IETF are addressing more wide-ranging solutions. One group, for instance, is working on World Wide Web Distributed Authoring and Versioning[127] (WebDAV), which will deal with meta-data, name space management, overwrite prevention and version management, and has become part of the W3's Resource Description Framework (RDF, see below).

ii) World Wide Web Consortium (W3) [128]

If the IETF is the chief body concerned with developing a pipeline for the Internet, the W3 Consortium is the main body devoted to integrating meta-data with respect to content on the Internet. It is, for instance, developing a convention for embedding meta-data in HTML.[129] When an IETF committee working on a Universal Resource Indicators (URI) was disbanded for want of agreement, the problem was taken up by W3C, who are tackling all the existing addressing schemes.[130] The result of these efforts will be to create a universal solution for the stopgap measures outlined above in figure 8.

One of the key activities of the W3 Consortium has been in the context of markup languages. As was noted earlier, languages such as Standard Generalized Markup Language (SGML),[131] helped the aims of meta-data by separating form from content: separating different views or presentation methods from the underlying information. The advent of Hyper Text Markup Language (HTML)[132] as an interim pragmatic solution temporarily obscured this distinction. Since then the consortium has been working on a subset of SGML, which is adequate for dealing with simpler documents and re-establishes the distinctions between form and content. This Extensible Markup Language (XML)[133] is also being submitted to the ISO (10179:1996).

Markup Languages	Chemical Markup Language	CML
	Handheld Device Markup Language	HDML
	Mathematical Markup Language	MML
	Precision Graphics Markup Language[134]	PGML
Description Languages	Hardware Description Language	HDL
	Web Interface Description Language	WIDL
Formats	Channel Definition Format	CDF
	Resource Description Format	RDF

Figure 10. Special markup and description languages and formats linked with XML.

It is foreseen that XML will form a basis to which one will add Cascading Style Sheets (CSS)[135] as part of a Document Style Semantics and Specification Language (DSSSL).[136] Similarly one can then add specialized markup languages, decription languages and formats (figure 10). XML will serve as the underlying structure for a comprehensive scheme,[137] which includes Protocol for Internet Content Selection (PICS), Digital Signatures (Dsig), Privacy Information (P3P) within a Resource Description Framework (RDF). PICS initially began as a means of restricting access for children to pornographic and other dangerous contents. PICS is evolving into a common platform for labelling online resources and a system for describing content using a restricted vocabulary. The PICS labels (metadata) for Internet resources[138] have five aims:

1	Resource Descriptive Schemas
2	Organizational Management
3	Discovery and Retrieval
4	Intellectual Property Rights
5	Privacy Protection Tasks

PICS entails three kinds of metadata:[139] i) embedded in content; ii) along with, but separate from content and iii) provided by an independent provider (label bureau). In a next phase PICS will become part of a larger Resource Description Framework[140] (RDF), which aims at machine understandable assertions of web resources in order to achieve:

1	Resource Discovery	
2	Cataloging Catalogue Information	
3	Intelligent Software Agents	
4	Content Rating Endorsement Information	[PICS]
5	Intellectual Property Rights	
6	Digital Signatures Information about Sets of Documents	[Dsig]
7	Privacy Information	[P3P]
8	Information About Sets of Documents and Document Management	[Web DAV]

RDF will have at least three vocabularies, namely a Protocol for Internet Content Selection (PICS) rating architecture; the Dublin Core (DC) elements for digital libraries and Digital Signatures (Dsig) for authentication. RDF uses a Document Object Model[141] (DOM), and a Resource Description Messaging Format[142] (RDMF). Implicit in this approach is the possibility of mapping a subject in the Dublin Core Framework, with subjects in one of the main classification schemes (e.g. Library of Congress, Dewey, Göttingen) and a version in everyday language.

XML will thus serve as an underlying structure for simple web documents, while SGML continues to be used for complex information such as the repair manuals for aircraft carriers or large jets.[143] It is important to recognize that the W3's approach to meta-data is constantly evolving and is likely to change considerably in the course of the next few years.[144] For instance, the director of the W3 consortium, Tim Berners Lee, in a keynote to WWW7 (Brisbane, April 1998), recently outlined his vision of a global reasoning web, whereby every site would also be classed in terms of its veracity or truth value.

iii) Z39.50[145]

Complementing these efforts of the Internet community are those of the library world, which have focussed almost exclusively on interoperability among libraries and have left aside the more complex elements of Internet information. Chief among these is Z.39.50. This is the American National Standard for Information (ANSI) Retrieval. It is based on two ANSI-NISO documents (1992[146] and 1995[147]), which led to a network protocol,[148] that is session oriented and stateful, in contrast to http and gopher, which are stateless. An early version ran on WAIS. The new version runs over TCP/IP. It uses an Object Identifier (OID). Z39.50 has the following six attribute sets:

1	Bibliographic 1	Bib-1[149]
2	Explain	Exp-1
3	Extended Services	Ext-1
4	Common Command Language	CCL-1
5	Government Information Locator Service	GILS
6	Scientific and Technical Attribute Set (Superset of Bib-1)	STAS

In addition it offers six record syntaxes, namely:

	Explain	
2	Extended Services	
3	Machine Readable Card including national variants	MARC
4	Generic Record Syntax	GRS-1
5	Online Public Access Catalogue	OPAC
6	Simple Unstructured Text Record Syntax	SUTRS

The Library of Congress has become the central library site for Z39.50 developments. The solution is being used in the European Commission's OPAC Network (ONE), a project, which includes the British Library (BL), the Danish National Library (DB), the Dutch Electronic libraries project (PICA), an Austrian initiative (Joanneum Research) and the Swedish National Library. It is also being used in the Gateway to European National Libraries (GABRIEL).

Meanwhile the Z39.50 protocol has been accepted as a basic ingredient by the Consortium for the Interchange of Museum Information (CIMI), which in turn has been supported as a part of the European Commission's Memorandum of Understanding for Access to Europe's Cultural Heritage. Hence, while some technologists may lament that the solution lacks elegance, it has the enormous advantage of having been accepted by virtually all the leading players in the international library and museum scene and thus needs to be considered as one of the elements in any near future solution.

iv) Dublin Core

Major libraries and museums typically have highly professional staff and therefore assume that records will be in a MARC format, or possibly with more complex methods such as SGML or EAD, or the variations provided by TEI and CIMI. Smaller libraries cannot always count on access to such resources. To this end, the Online Computer Center (OCLC) based in Dublin, Ohio in conjunction with the National Center for Supercomputing Applications (NCSA), sponsored an initial Metadata Workshop (1-3 March 1995),[150] at which 17 elements of the Dublin Core (DC)[151] also known as Monticello Core Elements (Mcore) were proposed (see figure 11 below) as well as three types of qualifiers,[152] namely, language, scheme and type. Since this was the first of a series of meetings it is frequently referred to as Dublin Core 1.

A second meeting (Dublin Core 2), which took place in Warwick, England, produced the Warwick Framework.[153] This provided containers for aggregating packages of typed meta-data and general principles of information hiding. A third meeting (Dublin Core 3) held in Dublin, Ohio focussed on images.[154] A fourth meeting (Dublin Core 4) took place in Canberra[155]and a fifth (Dublin Core 5) in Helsinki.[156]

Title	Format
Creator	Identifier
Subject	Identifier
Description	Source
Publisher	Language
Contributors	Relation
Date	Coverage
Type	Rights

Figure 11. List of the fifteen Dublin Core (DC) or Monticello Core (Mcore) elements, seen as a basic subset of more complex records such as MARC, SGML, TEI etc.

The Dublin Core has nine working groups: rights management, sub-elements, data model, DC Data, DC and Z39.50; relation type, DC in multiple languages, coverage, format and resource types. The Dublin Core is being applied to the Nordisk Web Index and the European Web Index (NWI/EWI). One of the reasons why it is so significant is because it is being linked with a number of other meta-data formats, namely, HTML 2.0/3.2 META Elements, WHOIS ++ Document Templates, US MARC, SGML and possibly MCF.[157] These meta-data records may be bibliographic, but can also relate to administration, terms/conditions as well as ratings (figure 12).

MD Bibliographic	MD Administration	MD Terms/Conditions	MD Ratings
MD Dublin Core	MD MARC		

Figure 12. Basic scheme showing how meta-data (MD) pertaining to bibliographic records can be linked with administration, terms/conditions and ratings.

Digital Object	Metadata Container	
	Metadata Package	
Handle	Content Container	
Metadata Container	Content Element	Content Element
Content Container		Metadata Container
		Content Package

Figure 13. A more generalized scheme showing relations of meta-data sets to their various parts[158].

The true power of this approach is that it can readily be expanded into a more general method for handling, interchange and ultimately marketing of information and/or knowledge packages, which helps to explain why firms such as IBM have become very seriously interested in and supportive of this approach. It offers a new entry point for their e-business vision of the world (figure 13).

As the above figures reveal, it is foreseen that the Dublin Core elements from personal sites and smaller institutions will interact with the more elaborate formats of major institutions (MARC etc.). Hence while the Dublin Core may, at first glance, appear to be merely a quick and dirty solution to a problem, it actually offers an important way of bridging materials in highly professional repositories with those in less developed ones. Moreover, while the Dublin Core in its narrow form is primarily a method for exchanging records about books and other documents, within this more generalized, expanded context, it offers a method for accessing distributed contents.

How will the extraordinary potentials of the technologies outlined above be developed? Any attempt at a comprehensive answer would be out of date before it was finished. For the purposes of this paper it will suffice to draw attention to a few key examples. One of the earliest efforts to apply these new tools is the Harvest Information Discovery and Access System[159] The Harvest method uses the Summary Object Interchange Format (SOIF),[160] which employs the Resource Description Message Format (RDMF), in turn a combination of IAFA templates and BibTex[161] which is part of the Development of a European Service for Information on Research and Education (DESIRE)[162] project linked with the European Commission's Telematics for Research Programme. It has been applied to Harvest, Netscape, and the Nordisk Web Index (NWI). This includes a series of attributes,[163] a series of template types[164] and other features.[165] While this method is limited to Internet resources, it represents an early working model.

The challenge remains as to how these tremendously varied resources can be integrated within a single network, in order that one can access both new web sites as well as classic institutions such as the British Library[166] or the Bibliothèque de la France. One possible solution is being explored by Carl Lagoze[167] in the Cornell Digital Library project. Cornell is also working with the University of Michigan on the concept of an Internet Public Library.[168] Another solution is being explored by Renato Iannella[169] at the Distributed Technology Centre (DSTC). This centre in Brisbane, which was one of the hosts of the WWW7 conference in 1998, includes a Resource Discovery Unit. In addition to its Basic URN Service for the Internet (BURNS) and The URN Interoperability Project (TURNIP), mentioned earlier, it has an Open Information Locator Project Framework[170] (OIL). This relies heavily on Uniform Resource Characteristics (including Data,[171] Type, Create Time, Modify Time, Owner). In the Uniform Resource Name (URN), this method distinguishes between a Namespace Identifier (NID) and Namespace Specific String (NSS). This approach is conceptually significant because it foresees an integration of information sources, which have traditionally been distinct if not completely separate, namely, the library world, internet sources and telecoms. (figure 14).

urn:isbn:……………	publishing	ISBN no.
inet:dstc.edu.au…….	internet servers	listname
telecom:……………	telecom	telephone no.

Figure 14. Different kinds of information available using the Open Information Locator Project Framework (OIL).

Yet another initiative is being headed by the Open Management Group (OMG).[172] This consortium of 660 corporations has been developing a Common Object Request Broker Architecture (CORBA),[173] which links with an Interoperable Object Reference (IOR). One of its advantages is that it can sidestep some of the problems of interaction between hyper text transfer protocol (http) and Transfer Control Protocol (TCP). It does so by relying on Internet Inter Object Request Broker Protocol (IIOP). It also uses an Interface Repository (IR) and Interface Definition Language (IDL, ISO 14750)[174.] CORBA has been adopted as part of the Telecommunications Information Networking Architecture (TINA).[175]

Some glimpse of a growing convergence is the rise of interchange formats designed to share information across systems. The (Defense) Advanced Research projects Agency's (ARPA's) Knowledge Interchange Format (KIF) and Harvester's Summary Object Information Format (SOIF) have already been mentioned. NASA has a Directory Interchange Format (DIF). The Metadata Coalition has a Metadata Interchange Specification[176] (MDIS).

At the university level, Stanford University has a series of Ontology Projects.[177] The California Institute of Technology (Caltech) has a project called Infospheres concerned with Distributed Active Objects.[178] Rensselaer Polytechnic has a Metadatabase which includes an Enterprise Integration and Modeling Metadatabase[,179] a Visual Information Universe Model,[180] a Two Stage Entity Relationship Metaworld (TSER) and an Information Base Modelling System (IBMS)[181]

Meanwhile, companies such as Xerox have produced Metaobject Protocols[182] and Meta Data Dictionaries to Support Heterogeneous Data.[183] Companies such as Data Fusion (San Francisco), the Giga Information Group (Cambridge, Mass.), Infoseek (Sunnyvale, California),[184] Intellidex[185] Systems LLC, Pine Cone Systems[186] and NEXOR[187] are all producing new software and tools relevant to metadata.[188]

Vendors of library services are also beginning to play a role in this convergence. In the past, each firm created its own electronic catalogues with little attention to their compatibility with other systems. In Canada, thanks to recent initiatives of the Ontario Library Association (OLA), there is a move towards a province wide licensing scheme to make such systems available to libraries, a central premise being their compatibility and interoperability.

10. Global Efforts

Technologists engaged in these developments of meta-data on the Internet are frequently unaware that a number of international organizations have been working on meta-data for traditional sources for the past century. These include the Office Internationale de Bibliographie, Mundaneum,[189] the International Federation on Information and Documentation (FID[190]), the International Union of Associations (UIA[191]), branches of the International Standards Organization (e.g. ISO TC 37, along with Infoterm), as well as the joint efforts of UNESCO and the International Council of Scientific Unions (ICSU) to create a World Science Information System (UNISIST). Indeed, in 1971, the UNISIST committee concluded that: “a world wide network of scientific information services working in voluntary association was feasible based on the evidence submitted to it that an increased level of cooperation is an economic necessity”.[192]

In 1977, UNISIST and NATIS, UNESCO's concept of integrated national information concerned with documentation, libraries and archives, were merged into a new Intergovernmental Council for the General Information Programme (PGI).[193] This body continues to work on meta-data.

Some efforts have been at an abstract level. For instance, the ISO has a subcommittee on Open systems interconnection, data management and open distributed processing (ISO/IEC JTC1/SC21). The Data Documentation Initiative (DDI), has been working on a Standard Generalized Markup Language (SGML) Document Type Definition (DTD) for Data Documentation.[194] However, most work has been with respect to individual disciplines and subjects including art, biology, data, education, electronics, engineering, industry, geospatial and Geographical Information Systems (GIS), government, health and medicine, library, physics and science. Our purpose here is not to furnish a comprehensive list of all projects, but rather to indicate priorities thus far, to name some of the major players and to convey some sense of the enormity of the projects already underway. More details concerning these initiatives are listed alphabetically by subject in Appendix 3.

The most active area for meta-data has been in the field of geospatial and Geographical Information (GIS).[195] At the ISO level there is a Specification for a data descriptive file for geographic interchange (ISO 8211),[196] which is the basis for the International Hydrographic Organization’s transfer standard for digital hydrographic data (IHO DX-90).[197] The ISO also has standards for Geographic Information (ISO 15046)[198] and for Standard representation of latitude, longitude and altitude (ISO 6709),[199] as well as a technical committee on Geographic Information and Geomatics[200] (ISO/IEC/TC 211), with five working groups.[201] At the international level the Fédération Internationale des Géomètres (FIG) has a Commission 3.7 devoted to Spatial Data Infrastructure. The International Astronomical Union (IAU) and the International Union of Geodesy and Geophysics (IUGG) have developed an International Terrestrial Reference Frame (ITRF).[202]

At the European level, geographical information is being pursued by two technical committees, European Norms for Geographical Information (CEN/TC 287)[203] and European Standardisation Organization for Road Transport and Traffic Telematics (CEN/TC 278),[204] notably working group 7, Geographic Data File (GDF).[205] At the national level there are initiatives in countries such as Canada, Germany, and Russia. The United States has a standard for Digital Spatial Metadata,[206] a Spatial Data Transfer Standard (SDTS)[207]and a Content Standard Digital Geospatial Metadata[208] (CSDGM). [209] Meanwhile, major companies are developing their own solutions, notably Lucent Technologies,[210] IBM (Almaden),[211] which is developing spatial data elements[212] as an addition to the Z39.50 standard, Arc/Info, Autodesk and the Environmental Systems Research Institute (ESRI).

Related to these enormous efforts in geospatial and geographical information have been a series of initiatives to develop meta-data for the environment. At the world level, the United Nations Environmental Program (UNEP) has been developing Metadata Contributors.[213] In the G 8 pilot project dedicated to environment, there is a Metainformation Topic Working Group[214] (MITWG) and Eliot Christian has developed a Global Information Locator Service (GILS).[215] There is a World Conservation Monitoring Centre,[216] a Central European Environmental Data Request Facility (CEDAR). Australia and New Zealand have a Land Information Council Metadata[217] (ANZIC). In the United States, the Environmental Protection Agency (EPA) has an Environmental Data Registry.[218] Efforts at harmonization of environmental measurement have also occurred in the context of G7 and UNEP.[219]

In the field of science, the same Environmental Protection Agency (EPA) has a Scientific Metadata Standards Project.[220] The Institute of Electrical and Electronic Engineers (IEEE)[221] has a committee on (Scientific) Metadata and Data Management. In the fields of physics and scientific visualisation, the United States has a National Metacenter for Computational Science and Engineering[222] with the Khoros[223]project. In biology there are initiatives to produce biological metadata[224] and the IEEE has introduced a Biological Metadata Content Standard. In the United States there is a National Biological Information Infrastructure[225](NBII) and there are efforts at Herbarium Information Standards.

In industry, the Basic Semantic Repository[226] (BSR), has recently been replaced by BEACON,[227] an open standards infrastructure for business and industrial applications.
In engineering, there is a Global Engineering Network (GEN) and, as was noted above there are a number of consortia aiming at complete interoperability of methods. In the United States, which seems to have some meta-association for almost every field, there is a National Metacenter for Computational Science and Engineering.[228] In the case of

electronics, the Electronic Industries Association has produced a CASE Data Interchange Format (CDIF).

In the field of government, Eliot Christian's work in terms of the G7 pilot project on environment has inspired a Government Information Locator Service[229] (GILS). In health, the HL7 group has developed a HL7 Health Core Markup Language (HCML).
In education, there is a Learning Object Metadata Group,[230] a Committee on Technical Standards for Computer Based Learning (IEEE P1484) and Educom has a Metadata Tool as part of its Instructional Management Systems Project. In art, the Visual Resources Association (VRA) has produced Core Categories Metadata.[231]

Not surprisingly, the library world has been quite active in the field of metadata. At the world level, the International Federation of Library Associations (IFLA) has been involved, as has the Text Entering Initiative (TEI), the Network of Literary Archives (NOLA), and the Oxford Text Archive (OTA). At the level of G8, it is a concern of pilot project 4, Biblioteca Universalis.[232] At the European level there is a list of Library Information Interchange Standards (LIIS).[233] In Germany, there is a Metadata Registry concerned with metadata and interoperability in digital library related fields.[234] In the United States, there is an ALCTS Taskforce on Metadata and a Digital Library Metadata Group (DLMG).

In the United Kingdom, the Arts and Humanities Data Service (AHDS) and the United Kingdom Office for Library and Information Networking (UKOLN)[235] have a Proposal to Identify Shared Metadata Requirements,[236] a section on Metadata[237]and for Mapping between Metadata Formats.[238] They are concerned with Linking Publishers and National Bibliographic Services (BIBLINK) and have been working specifically on Resource Organization and Discovery in Subject Based Services (ROADS)[239] which has thus far produced gateways to Social Science Information (SOSIG), Medical Information (OMNI)[240] and Art, Design, Architecture, Media (ADAM). They have also been active in adopting basic Dublin Core elements. A significant recent by Rust has offered a vision provided by an EC project, Interoperability of Data in E-Commerce Systems (INDECS), which proposes an integrated model for Descriptive and Rights Metadata in E-Commerce.[241] This concludes the detour announced twelve pages ago.

Standing back from this forest of facts and projects, we can see that there are literally hundreds of projects around the world all moving towards a framework that is immensely larger than anything available in even the greatest physical libraries of the world. Tedious though they may seem, these are the stepping stones for reaching new planes of information, which will enable some of the new scenarios in knowledge explored earlier. They are also proof that the danger of a second flood in terms of information as foreseen by authors such as Pierre Lévy is not being met only with fatalistic, passive, resignation.

Steps have been taken. Most of the projects thus far have focussed on the pipeline side of the problem. How do we make a database in library *A* compatible with that of library *B* such that we can check references in either one, and then, more importantly, compare references found in various libraries joined over a single network? Here the Z39.50

protocol has been crucial. As a result, networks are linking the titles of works in a number of libraries spread across a country, across continents and potentially around the world. Examples include the Online Computer Center (OCLC), the Research Library Information Network (RLIN) based in the United States and PICA based in the Netherlands. The ONE[242] project, in turn, links the PICA records with other collections such as Joanneum Research and the Steiermärkische Landesbibliothek (Graz, Austria), the Library of the Danish National Museum, Helsinki University Library (Finland), the National Library of Norway, LIBRIS (Stockholm, Sweden), Die Deutsche Bibliothek (Frankfurt, Germany), and the British Library. Some of these institutions are also being linked through the Gateway to European National Libraries Project (GABRIEL).[243] The German libraries are also working on a union catalogue of their collections. In the museum world there are similar efforts towards combining resources through the Museums Over States in Virtual Culture (MOSAIC)[244] project and the MEDICI framework of the European Commission. In addition, there are projects such as the Allgemeine Künstler Lexikon (AKL) of Thieme-Becker, and those of the Getty Research Institute: e.g. Union List of Author Names (ULAN) and the Thesaurus of Geographic Names (TGN)[245].

What are the next steps? The Maastricht McLuhan Institute, a new European Centre for Knowledge Organization, Digital Culture and Learning Technology, will focus on two. First, it will make these existing distributed projects accessible through a common interface using a System for Universal Media Searching (SUMS). The common interface will serve at a European level for the MOSAIC project and at a global level as part of G8, pilot project five: Multimedia access to world cultural heritage.

A second, step will be to use these resources as the basis for a new level of authority lists for names, places and dates. In so doing it will integrate existing efforts at multilingual access to names as under development by G8 pilot project 4, Biblioteca Universalis, and earlier efforts of UNEP, to gain new access to variant names. In the case of terms, it will make use of standard classifications (e.g. Library of Congress, Dewey, Göttingen and Ranganathan[246]), as well as specialized classification systems for art such as Iconclass[247] and the Getty Art and Architectural Thesaurus.[248] As such the research project will in no way be in competition with existing projects. Rather it will integrate their efforts as a first step towards a new kind of digital reference room.[249]

Access to knowledge, which deals with claims about information, requires more than keywords in free text and natural language. Systematic access to knowledge requires a) authority files for names, subjects, places with their variants as outlined above; b) maps of changing terms and categories of knowledge in order to access earlier knowledge collections; c) systematic questions. If one takes basic keywords, translates these into standardized subject terms (what?), and combines these questions with those of space (where?), time (when?) and process (how?), one has a simple way of employing the Personality, Matter, Energy, Space and Time (PMEST) system of Ranganathan. With some further division these questions also allow a fresh approach to Aristotle's substance- accident system (figure 15). In very simple terms: isolated questions provide

Who?	What?	How?	Where?	When?
Personality (P)	Matter (M)	Energy (E)	Space (S)	Time (T)
Being	Substance	Relation	Position	Time
	Matter	Activities[250]	Dimension	
		Quantity	Place	
		Quality		

Figure 15. Five basic questions related to the five key notions of Ranganathan's PMEST system and the ten basic categories of Aristotle's substance accident system.

access to data and information. Combinations of questions provide access to structured information or knowledge.

One of the major developments over the past thirty years has been a dramatic increase in different kinds of relations. Perrault[251] in a seminal article introduced a method of integrating these systematically within UDC. The Medical Subject Headings (MESH) has five kinds of relations. Systems such as Dewey are too primitive to allow a full range of relations. Nonetheless, if the Dewey subjects are mapped to the UDC system where these connections have been made, then one can integrate relations within the search strategies.[252] Thus relations such as broader-narrower offer further search stratagems.

In order to ensure that the scope of the project becomes more universal than merely universally daunting, the digital reference room will begin with a subset of the whole, creating the cultural section of a future comprehensive reference room. The research function of the Institute will focus initially on extending the web of co-operation with other cultural institutions in order to prevent duplication of efforts and reinvention of the wheel. On this basis the cultural digital reference room will gradually be expanded to include links to corresponding digital texts from the great institutions. The institute itself will not attempt to replicate physically any of these collections. Rather it will serve as a centralized list of authority names, places and dates linked with a distributed collection of reference sources.

This seemingly narrow focus on art and culture will lead quite naturally to other fields. Paintings typically entail narratives. Hence the reference room must expand to include literature. As was already noted, to study the location of paintings and other museum objects, requires systematic treatments of scale and thus the reference room will expand to include the fields of geo-spatial and geographical information systems. In a subsequent phase, research will turn to expanding the scope of the digital reference room from this focus on culture, from the arts to the sciences, to the full range of human knowledge. As this occurs the common interface will be linked with the digital reference room to produce a System for Universal Multi-Media Access (SUMMA).

11. Emerging Scenarios

These authority lists of names, places and dates will, in the first instance, serve as the basis for a new level of interoperability among collections, at the content level as opposed to the basic pipeline connectivity. This entails considerably more than simple access to titles or even the full contents of materials listed in contemporary author and subject catalogues of libraries. On the one hand, it entails links to dictionaries and encyclopaedias, which will provide searchers with related terms. It also involves cross-references to citation indexes, abstracts and reviews.

Reference rooms, as the collective memory of civilization's search methods, also contain a fundamental historical dimension. To take a concrete example: today a book such as Dürer's *Instruction in Measurement* (*Underweysung der Messung*) is typically classed under perspective. In earlier catalogues this book was sometimes classed under proportion or more generally under geometry. As digital library projects extend to scanning in earlier library and book publishers' catalogues, a new kind of retrospective classification can occur, whereby titles eventually have both their old classes and their modern ones. This will radically transform future historical research, because the catalogues will then lead scholars into the categories relevant for the period, rather than to those that happen to be the fashion at the moment. Links to on-line versions of appropriate historical dictionaries will be a next step in this dimension of the digital reference room. Eventually there can be the equivalents of on-line etymologies on the fly.

There are, of course, many other projects concerning digital libraries. Some, such as the Web Analysis and Visualization Environment (WAVE) specifically attempt to link interoperable meta-data with facetted classification.[253] This project is important because it links methods from traditional library science (e.g. classifications) with those of mathematics (concept analysis). Even so this and other systems are focussed on access to contemporary information. What sets the MMI project apart from these initiatives is that it sets out from a premise of concepts and knowledge as evolving over time, as an historical phenomenon.

It will take decades before the digital library and museum projects have rendered accessible in electronic form all the documents and artifacts now stored in the world's great libraries, museums and galleries. By that time the enormous growth in computing power and memory, will make feasible projects that most would treat as science fiction or madness today. In the past decades we have seen the advent of concordances for all the terms in the Bible, Shakespeare and other classic texts. A next step would be to transform these concordances into thesauri with formally defined terms, such that the relations and hierarchies therein become manifest. This principle can then gradually be extended to the literature of a school, a particular decade, a period or even an empire.

This will allow us to look completely afresh at our past and ask whole new sets of questions. Rather than speaking vaguely of the growth of vernacular languages such as English or Italian, we can begin to trace with some quantitative precision, which were the crucial periods of growth. This will lead to new studies as to why the growth occurred at just that time. We shall have new ways of studying the history of terms and the changing associations of those terms. We shall move effectively to a new kind of global citation index.

It is said that, in the thirteenth century, it took a team of one hundred Dominican monks ten years of full time work to create an index of the writings of St. Thomas Aquinas. With a modern computer that same task can theoretically be accomplished in a few minutes. (Cynics might add that this would be after having spent several months writing an appropriate programme and a few weeks debugging it). In the past, scholars also typically spent days or months tracing the sources of a particular passage or crucial text. Indeed, scholars such as Professor M. A. Screech, who sought to trace the sources of major authors such as Montaigne or Erasmus, discovered that this was a lifetime's work. In the eye's of some this was the kind of consummate learning that epitomized what knowledge in the humanities was all about. For a reference to Aquinas might lead to a reference to Augustine, who alluded to Plotinus, who was drawing on Plato. To understand a quote thus took one into whole webs of cumulative philosophical, religious and cultural contexts, which make contemporary hypertext efforts look primitive indeed.

If we can trace quotes, we should also be able to trace stories and narrative traditions. Ever since the time of Fraser's *Golden Bough*,[254] we have been aware of the power of recurrent themes in poems, epics, legends, novels and other writings. Indeed much of academic studies of literature are based on little else. Trace the theme of x through this author or that period often dominates assignments and exams. If tracing these themes were automated, it would open up new approaches in much more than literature. For instance, if we were standing in front of Botticelli's *Story of Griselda* (London, National Gallery), and were unfamiliar with the story, we could just point our notepad computer and have it show and/or tell us the story.

In the case of direct quotations, machines can theoretically do much of this work today. Often, of course, the references are more indirect than direct or they are allusions that could lead to twenty other passages. It is well known that each of us has their favourite terms, words, which are imbued with special significance, as well as preferred words or sounds that serve as stopgaps in our thought. (Many of us, for instance, have met an individual who bridges every sentence or even phrase with an "um," or peppers their speech with an "actually," "indeed" or some other semantically neutral stopgap). An average person often gets by with a vocabulary of only about a thousand words. By contrast there are others in the tradition of Henry Higgins with vocabularies in the tens of thousands of words. Will we some day have the equivalent of fingerprints for our vocabularies, such that we can identify persons by their written and spoken words? Will the complexities of these verbal maps become a new way for considering individual development? Will we begin studying the precise nature of these great verbalizers? Some languages are more substantive (in the sense of noun based) whereas other such as Arabic are more verbal (in the sense of verb based)? Will the new "verbal" maps help us to understand cultural differences in describing the world and life? Will such maps become a basic element of our education?[255]

In the past, famous individuals wrote guidebooks and travelogues, which provided maps of their physical journeys and they wrote autobiographies to offer maps of their mental and emotional journeys. In the past generation, personalities such as Lord Kenneth Clark produced documentaries such as *Civilization* to accomplish this in the medium of film. At the Sony laboratories in Paris, Dr. Chisato Namaoka[256] is engaged in a Personal Experience Repository Project, which aims to record our memories as they occur while we visit a museum or significant tourist site, and to use that captured information for further occasions. Individuals such as Warren Robinett or Steve Mann have gone much further to speculate on the possibility of having a wearable camera that records everything one ever did in one's life: another take on the scenarios presented in the movie *Truman Show* (1998). Such developments could readily transform our conceptions of diaries and other memory devices.

They also introduce possibilities of a new kind of "experience on demand" whereby any visit to a tourist site might be accompanied with the expressions of famous predecessors. In the past, the medium determined where we could have an experience: books tended to take us to a library, films to a cinema, television to the place with a television set. In future, we can mix any experience, anywhere, anytime. How will this change our patterns of learning and our horizons of knowledge?

All this assumes, of course, that computers can do much more than they can today. This is not the place to ponder at length how soon they will be able to process semantic and syntactical subtleties of language to the extent that they can approach deep structure and elusive problems of meaning and understanding. Nor would it be wise to speculate in great detail or to debate about what precisely will be the future role of human intervention in all this.

Rather, our concern is with some more fundamental problems and trends. One of the buzzwords about the Internet is that it is bringing "disintermediation,"[257] which is used particularly in the context of electronic commerce to mean "putting the producer of goods or services directly in touch with the customer." Some would go further to insist that computers will soon allow us to do everything directly: order books via sites such as Amazon.com without needing to go to bookstores; go shopping on-line without the distractions of shopping-malls. In this scenario, computers will make us more and more active and we shall end up doing everything personally. At the same time, another group claims that computers will effectively become our electronic butlers, increasingly taking over many aspects of everyday life. In this scenario, computers will make us more and more passive and we shall end up doing less and less personally. Indeed, some see this as yet another move in the direction of our becoming complete couch potatoes.

In our view, there is no need to fear that computers will necessarily make us exclusively active or passive. That choice will continue to depend on the individual, just as it does today. Nonetheless, it seems inevitable that computers will increasingly play an inter-mediating role, as they become central to more and more aspects of our lives. In the past decade, the concept of agents has evolved rapidly from a near science fiction concept to an emerging reality. There is now an international Foundation for Intelligent Physical

Agents (FIPA).[258] There are emerging fields devoted to user-modelling and user adapted interaction, entailing person-machine interfaces, intelligent help systems, intelligent tutoring systems and natural language dialogues.[259]

Leading technologists such as Philippe Quéau, have predicted the advent of tele-virtuality,[260] whereby avatars[261] will play an increasing role as our virtual representatives in the Internet. Recently, in Paris, there was a first international conference on Virtual Worlds (July 1998), attended by those at the frontiers of two, hitherto quite separate fields: virtual reality and artificial life. Some predict that self-evolving artificial life forms will soon be integrated into avatars. Some of the early virtual worlds began simply by reconstructing actual cities such as Paris[262] or Helsinki.[263] Others such as Alpha World[264] are creating a new three-dimensional virtual world based on elements familiar from the man-made environment. Potentially these worlds could be synthetic ones, or purely imaginary ones, no longer subject either to the physical laws or even the spatial conditions of planet earth. At Manchester, Professor Adrian West,[265] has begun to explore the interactions of virtual worlds, parts of which are subject to different laws of physics.

In a world where the Internet Society is planning to become interplanetary[266], assigning addresses for different planets, the solar system and eventually other galaxies, the question of navigation is becoming much more acute and “Where in the world?” is becoming much more than a turn of phrase. We shall need new methods to discern whether the world we have entered is physically based or an imaginary construct; whether our avatar has embarked on a “real” trip or almost literally into some flight of phantasy. In the past generation, children have grown up expecting the realism of video-games to be considerably poorer than that of realistic films or the real world. Within the next generation, such easy boundaries will increasingly blur and then disappear almost entirely. Is it reality? Is it a game? Is it playful reality or realistic playfulness? Such questions will become ever more difficult to discern.

In light of all this, some activities of scholars will certainly remain: reflecting on what sources mean, weighing their significance, using them to gain new insights and to outline new analyses, goals, dreams, visions, even utopias. Meanwhile, it is likely that many of the activities which preoccupied scholars for much of their lives in the past will become automated within the next generations, namely, hunting for texts, tracking down quotes and looking for sources.

At the same time many new activities will emerge. Before the advent of space travel and satellites no one could imagine precisely what it would be like to look at the earth from space. Within a single generation we have developed methods for zooming systematically from such satellite images down to a close up of an object on earth in its original scale, and even how to descend to microscopic levels to reveal biological, molecular and atomic properties. We need to create the equivalents of such zooms for our conceptual worlds, moving systematically from broader terms to narrower terms. We need new ways of visualizing how the horizons of our conceptual worlds grow. At the simplest level this entails demonstrating how we have shifted from a Ptolemaic to a Copernican worldview.

Much more elusive and difficult is to find ways of showing how our mental horizons have expanded over time. What impact did major changes in travel such as the crusades, pilgrimages, and the grand tour, have on vocabularies or inventions? Most of the future questions to be asked cannot yet be formulated because we cannot yet see ways of collecting, ordering and making sense of the vast materials that would be required to formulate them.

At present, the frontiers of scientific visualization are focussed on helping us to see phenomena such as the flow of air in a jet at supersonic speeds, the development of storms and tornadoes, the dispersal of waste in Chesapeake Bay, changes in the ozone layer, and many other events that we could not begin to know until we had methods for seeing them. Computers are transforming our knowledge because they are helping us to see more than we knew possible. The physical world opens as we make visible its unseen dimensions.[267] The mental world awaits a similar journey and as with all journeys we must remember that what we see is but a small part of the immensity that is to be known, experienced, sensed or somehow enters our horizons.

Amidst all these developments there are also new problems. In the United States, there is a tendency to reduce cultural heritage to a commodity as merely another aspect of consumerism. This tendency implies the erosion of a long standing distinction between public good versus private interests. The public good, as protected by governments, is a long term investment. Private, business interests are concerned with short term-profits. Culture is more than a series of paintings in museums and books in libraries. In each generation a subset of those paintings and books serve as an informal corpus of works which any given country or group uses for their identity. In the Netherlands, for instance, painters such as Rembrandt, Vermeer and Van Gogh help define not only Dutch painting but what it means to be Netherlandish, just as Shakespeare and Milton, help persons in England define what it means to be English. Private companies may have the resources to buy a given manuscript or painting but they could never afford to buy all the cultural artefacts, by which French, German, Italian or any of the other greater cultures is defined. Needed, therefore, is an ongoing political commitment towards a cumulative cultural memory, whereby the individualities of culture will be fostered, for precisely therein lie the secrets of why tourists are so fascinated by the uniqueness of Italy or the mysterious *je ne sais quoi* of France.[268] While some may predict that corporations could or will rule the world, a deeper challenge lies in assuring that we retain a world that is coherent and self-conscious enough to be worth ruling. An unruly world without cultural identities would not only be much less poorer spiritually but would also be much less attractive from an economic standpoint.

12. Conclusions

This paper began from the premise that every new medium changes our definitions of, approaches to and views of knowledge. It claimed that networked computers (as enabled by the Internet), cannot be understood as simply yet another medium in a long evolution that began with speech and evolved via cuneiform, parchment, manuscripts to printed books and more recently to radio, film, and video. Computers offer a new method of translating information from one medium to another, wherein lies the deeper meaning of the overworked term multimedia. Hence, computers will never create paperless offices. They will eventually create offices where any form of communication can be transformed into any other form.

In the introduction we raised questions about an excellent article by Classen concerning major trends in new media.[269] He claimed that while technology was expanding exponentially, the usefulness[270] of that technology was expanding logarithmically and that these different curves tended to balance each other out to produce a linear increase of usefulness with time. In our view, simpler explanations are possible. First, technologists have been so concerned with the pipeline aspects of their profession (ISO layers 1-6 in their language), that they have ignored the vast unexplored realms of applications (ISO layer 7). Second, phrases such as "build it and they will come" may sound rhetorically attractive, but unless what is built actually becomes available, it can neither be used nor useful. Rather than seek elusive limits to usefulness, it is much more effective to make things available. In short, a more effective formulation might be: let it be useable and used and usefulness will follow.

Any attempt at a systematic analysis of future applications (cf. Appendix 1) would have required at least a book length study. For this reason the scope of the present paper was limited to exploring some of the larger implications posed by the new media. We claimed that there are at least seven ways in which networked computers are transforming our concepts of knowledge. First, they offer new methods for looking at processes, how things are done, which also helps in understanding why things are done in such ways. Second, and more fundamentally, they offer tools for creating numerous views of the same facts, methods for studying knowledge at different levels of abstraction. Third, they allow us to examine the same object or process in terms of different kinds of reality. Fourth, computers introduce more systematic means of dealing with scale.

Fifth, they imply a fundamental shift in our methods for dealing with age-old problems of relating universals and particulars. Analysis thereof pointed to basic differences between the arts and sciences and the need for independent historical approaches to reflect these, all the more so because computers, which are only concerned with showing us the latest version of our text or programme, are a direct reflection of this scientific tradition. We need a richer model that also shows us layered, cumulative versions. Sixth, computers transform our potential access to data through the use of meta-data. Seventh and finally, computers introduce new methods for mediated learning and knowledge through agents.

While the main thrust of the paper was focussed on the enormous potentials of networked computers for new approaches to knowledge, some problems were raised. These began with some of the limitations in the technology that is actually available today, with respect to storage capacity, processor speeds, bandwidth and interoperability. The dangers of making normative models, which then affect the future evidence to be considered, as in the case of the human genome project, were touched upon. So too were the dangers underlying some of the rhetorically attractive, but equally misleading assumptions behind come contemporary approaches to complex systems.

At the outset of the paper, mention was also made of the dangers articulated by Pierre Lévy that we are in danger of a second flood, this time in the form of a surfeit of information, as a result of which we can no longer make sense of the enormity of materials descending upon us. Partly to counter this, a section of the paper entered into considerable detail on worldwide efforts concerning meta-data as a means of regaining a comprehensive overview of both the immense resources that have been collected already and the ever increasing amounts which are being added daily. Sense making tools are an emerging field of software.

A half century ago pioneers such as Havelock, Innis and McLuhan recognized that new media inevitably affect our concepts of what constitutes knowledge. The mass media epitomized this with McLuhan's pithy phrase: "The medium is the message." Reduced and taken in isolation, it is easy to see, in retrospect, that this obscured almost as much as it revealed. The new media are changing the way we know. They are doing so in fundamental ways and they are inspiring, creating, producing, distorting and even obscuring many messages. New machines make many new things possible. Only humans can ensure that what began as data streams and quests for information highways become paths towards knowledge and wisdom.

Acknowledgements

I am very grateful to Dr. Harald Kraemer for inviting me to pursue these questions of computers and knowledge for the inaugural address at his international conference: *Euphorie Digital? Aspekte der Wissensvermittelung in Kunst, Kultur und Technologie* held at the Siemens Nixdorf Museum (Paderborn, 28-30 October). I am grateful to Laurie McCarthy of the Lockheed Martin Advanced Technology Center (Palo Alto), Dr. Flaig of the Fraunhofer Gesellschaft (Darmstadt) and John T. Edmark of Lucent Technologies (Holmdel, NJ) for kindly sending videos of their work, which were used for the lecture.

The larger framework of this paper has grown out of discussions with friends and colleagues such as Dr. Rolf Gerling, Dipl. Ing. Udo Jauernig, Eric Dobbs, John Orme Mills, O.P., and Professor André Corboz and many years of experience at research institutions including the Warburg Institute (London), the Wellcome Institute for the History of Medicine (London), the Herzog August Bibliothek (Wolfenbüttel), where Frau Dr. Sabine Solf played an essential role, the Getty Center for the History of Art and the Humanities – now the Getty Research Institute (Santa Monica), the McLuhan Program in Culture and Technology at the University of Toronto and more recently the Ontario Library Association, where Larry Moore has been most encouraging. I am grateful to the individuals at all of these centres.

Finally I am grateful to members of my own team, who have been both generous and supportive, notably, Rakesh Jethwa, Andrew McCutcheon, Greg Stuart, Hugh Finnegan, John Bell, Elizabeth Lambden and John Volpe. I am grateful to Dr. Anthony Judge (UIA) and Heiner Benking (Ulm) for challenging me to think more thoroughly about problems in moving from two-dimensional to three-dimensional navigation. Mr. Benking, Professor Corboz and Father Orme-Mills kindly read the manuscript and offered suggestions.

This essay has grown partly out of work for Eric Livermore (Nortel, Advanced Networks), and Stuart McLeod (CEO, Media Linx), so I express my thanks to the two companies an in particular to these two individuals who have been very loyal in their support and encouragement.

It is planned that this article will be republished by the Ontario Library Association as the fourth chapter of a book, *Frontiers in Conceptual Navigation for Cultural Heritage*.

Appendix 1. Basic Application Areas not Systematically Discussed in this Paper.

Content Supplier

1. Describe, Analyse
2. Class
3. Index
4. Scan, Capture
5. Create
6. Database
7. Display
8. Store, Archive
9. Preserve

Broker

10. Retrieve
11. Restore
12. Translate
13. Transform, Morph
14. Encrypt
15. Copyright (Legal)
16. Transact (Financial)
17. Administer
18. Service Bureaus

User Task

18. Work
19. Design
20. Manufacture
21. Output, Broadcast
22. Forecast

User Discipline

23. Libraries
24. Museums
25. Military
26. Industry
27. Government
28. Education
29. Health
30. Entertainment
31. Evaluate

Appendix 2. Glossary of Key Elements in Internet Metadata

Client
Whois ++
PH
Lightweight Directory (LDAP)

Middleware Platform
Common Object Request Broker Architecture (CORBA)
JAVA
    Remote Method Interface (RMI)
    A Java-Corba Alliance is in the works (May 18, 1998)[271]
DCOM

Protocol Front End
a) Protocols

1) Hyper Text Transfer Protocol (HTTP)
    There is discussion within W3 that the work of CORBA's IIOP by the OMG might be integrated with that of http.

2) Internet Inter-Object Request Broker Protocol (IIOP)

b) Directory Services

1. Finger
   Queried only one database at a time

2. Whois
   Stateless query response protocol
   Queried only one database at a time

3. X.500
   User Friendly Names as URN
   uses an abstract naming system.[272]

4. Whois++[273]
   TCP based protocol, which stems from whois
   Centroids: collection of attributes and unique values for these attributes taken from contents of server = lightweight full-text index to determine whether server has any relevant info.
   Generated automatically from gopher, www, ftp via IAFA templates
   Harvested by index servers

5. WebPh[274] (PH)
   is a client which scans strings of text which appear to be

http, gopher, ftp or e-mail and converts these on the fly
into clickable hyperlinks.

6. Lightweight Directory Access Protocol[275] (LDAP)
Access a LDAP directory service or
one backended by X.500

7. Simple Object Lookup Protocol (SOLO)[276]

8. Simple Discovery Protocol (SDP)
a) nameservers use SDP to communicate with one another
b) clients use SDP directly.

c) Get

1. Wide Area Information Server (WAIS)
2. Verity
3. Fulchrum
4. Gopher
5. Z39.50

d) Put/Post

1. File Transfer Protocol (FTP)
2. News Transfer Protocol (NTP)

Indexing Object

Indexer API

Database Backend or Query Protocol

Structured Query Language (SQL)

Z39.50

GNU’s Not Unix Database Manager (GNU DBM)

GDBM

IBM DBaseII

Oracle

Sybase

Appendix 3. Metadata in Individual Disciplines

Please Note:
Only items not specifically documented in the text have footnotes in the following list.

Art
Visual Resources Association (VRA)
    Core Categories Metadata

Biology
Biological Metadata

IEEE
    Biological Metadata Content Standard

Herbarium Information Standards

National Biological Information Infrastructure (NBII)
    Current Status

Schemas
Terrestrial Biological Survey Data, Australia

Data
Data Documentation Initiative (DDI)
    SGML DTD for Data Documentation

Education
Learning Object Metadata Group

Committee on Technical Standards for Computer Based Learning (IEEE P1484)

Educom
    Instructional Management Systems Project
        Metadata Tool

Electronics
Electronic Industries Association (EIA)
    CASE Data Interchange Format (CDIF)

Engineering
Global Engineering Network (GEN)

National Metacenter for Computational Science and Engineering

Industry
Basic Semantic Repository (BSR)

Replaced by

BEACON

Open standards infrastructure for business and industrial applications

Environment

United Nations

United Nations Environmental Program	(UNEP)
Metadata Contributors	
Environmental Information Systems	
Environmental Database Systems	
Environmental Protection Agency	(EPA)
Environmental Data Registry	
Central European Environmental Data Request Facility	(CEDAR)

World Conservation Monitoring Centre[277]

Australia

Australia New Zealand Land Information Council Metadata	(ANZIC)
Environment Resources Information Network	(ERIN)
Geospatial and Geographical Information	(GIS)
A. International Standards Organization	(ISO/IEC)
Open systems interconnection, data management and open distributed processing	(ISO/IEC JTC1/SC21)
Specification for a data descriptive file for geographic interchange	(ISO 8211)
Basis for International Hydrographic Organization transfer standard for digital hydrographic data	(IHO DX-90)
Geographic Information	(ISO 15046)
Standard representation of latitude, longitude and altitude	(ISO 6709)
Geographic Information and Geomatics	(ISO/IEC/TC 211)

WG1. Framework and reference model

WG 2. Geospatial data models and operators

WG 3. Geospatial data administration

WG 4. Geospatial Services

WG 5. Profiles and Functional Standards

B. International

Fédération Internationale des Géomètres	(FIG)
Commission 3.7 Spatial Data Infrastructure	

International Terrestrial Reference Frame	(ITRF)
C. Multi-National	
European Standardisation Organization for Road Transport and Traffic Telematics	(CEN/TC 278)
WG 1. Framework for standardisation	
WG 2. Models and Applications	
WG 3. Data Transfer	
WG 4. Location Reference Systems	
WG 7 Geographic Data File	(GDF)
European Norms for Geographical Information	(CEN/TC 287)
European Policy Framework for Geographical Information	
European Geographical Information Infrastructure	
European Umbrella Group for Geographical Information	(EUROGI)[278]
European Spatial Metadata Infrastructure	
GRIded Binary	(GRIB)
Standard set of codes for the storage and transmission of meteorological data	
S-57[279]	
SQL-MM[280]	
Geographic Tag Image File Format	(GeoTif)[281]
DG XIII Open Information Interchange	(OII)
GIS Standards[282]	
European Umbrella Organisation for Geographical Information	(EUROGI)
Open Geographical Information Systems Consortium Inc	(OGC)
Open Geographic Data Committee	
Open Geodata Interoperability Specification	(OGIS)
C. National	
Canada	
Geographic Data BC, Ministry of Environment, B.C.[283]	
Spatial Archiving and Interchange Format	(SAIF)[284]
Germany	
Deutsche Dachverbund für Geoinformation	(DDGI)[285]
Russia	

Standards List[286]

United States
American Society for Testing and Materials (ASTM)[287]
Digital Spatial Metadata[288]
Spatial Data Transfer Standard (SDTS)[289]

Federal Geographic Data Committee (FGDC)[290]
Content Standard Digital Geospatial Metadata[291] (CSDGM)[292]
Metadata Standards[293]

Institute for Land Information (ILI/LIA)

Global Positioning Systems (GPS)
John Abate[294]

F. Major Companies
ARC/INFO
ArcView

Autodesk
ARX Object Technology

Bell Laboratories, Lucent Technologies

Environmental Systems Research Institute (ESRI)

IBM Almaden
Z39.50 + spatial data elements

Government
Government Information Locator Service (GILS)
Cf. Global Information Locator Service (GILS)

Health and Medicine
HL7 Health Core Markup Language (HCML)

Library
ALCTS Taskforce on Metadata
Digital Library Metadata Group (DLMG)

Library Information Interchange Standards (LIIS)

Network of Literary Archives (NOLA)

Oxford Text Archive (OTA)

Text Entering Initiative (TEI)

United Kingdom
Arts and Humanities Data Service and (AHDS)
United Kingdom Office for Library and Information Networking (UKOLN)
Proposal to Identify Shared Metadata Requirements
Metadata
Mapping between Metadata Formats
Linking Publishers and National Bibliographic Services (BIBLINK)

Nordic Metadata Project (NMP)

Physics (Scientific Visualisation)
National Metacenter for Computational Science and Engineering
Khoros

Notes on Paper

Science
Environmental Protection Agency (EPA)
Scientific Metadata Standards Project

Institute of Electrical and Electronic Engineers (IEEE)
(Scientific) Metadata and Data Management.

Notes

[1] It is instructive to note that although the impulses for this research came from various centres, notably, Cambridge, many of the key ideas developed at the University of Toronto in the context of classical studies, history, English literature and media studies.

[2] Eric A. Havelock, *Preface to Plato*, Cambridge: Belknap Press, Harvard University *Press, 1963.*

[3] In the past generation scholars such as Jack Goody (Cambridge) have explored the implications of this phenomenon in the context of developing cultures, particularly, Africa. See: , for instance, Jack Goody, *The Domestication of the Savage Mind*, Cambridge: Cambridge University Press, 1977
cf. http://gopher.sil.org/lingualinks/library/literacy/GFS812/cjJ360/GFS3530.htm.) See: also the work edited by him, *Cultura escrita en sociedades tradicionales*, Barcelona: Gedisa, 1996
cf. http://www.ucm.es/info/especulo/numero5/goody.htm).

[4] Marshall McLuhan, *The Gutenberg Galaxy, The Making of Typographic Man*, Toronto: University of Toronto Press, 1962; *Understanding Media: The Extensions of Man*, New York: McGraw-Hill,1964. (cf. http://www.mcluhanmedia.com/mmclm005.html).

[5] Harold Adams Innis, *Empire and Communications*, (1950), ed. David Godfrey, Victoria, B.C.: Press Porcepic, 1986 and *The Bias of Communication*. Introduction Marshall McLuhan. Toronto: University of Toronto Press, 1964

(http://www.mala.bc.ca/~soules/paradox/innis.htm). Cf. Judith Stamps, *Dialogue, Marginality and the Great White North. Unthinking Modernity: Innis, McLuhan and the Frankfurt School*, Montreal/Kingston: McGill Queens UP, 1995.

[6] W. Terence Gordon, *Marshall McLuhan, Escape into Understanding*, Toronto: Stoddart, 1997.

[7] See: http://www.daimi.aau.dk/~dibuck/hyper/bush.html.

[8] See: http://www2.bootstrap.org/.

[9] See: http://www.sfc.keio.ac.jp/~ted/

[10] See, for instance, Derrick de Kerckhove, *The Skin of Culture: Investigating the New Electronic Reality*, Toronto: Somerville House Publishing, 1995, and *Connected Intelligence: The Arrival of the Web Society*, Toronto: Somerville House Publishing, 1997.

See: http://www.mcluhan.toronto.edu/derrick.html).

[11] Pierre Lévy, *L'Intelligence Collective: Pour une Anthropologie du Cyberspace*, Paris: Éditions La Découverte, 1994. Translation: *Collective Intelligence: Mankind's Emerging World in Cyberspace*, translated by Robert Bononno, Plenum Press, 1998. See: also Ibid., *The Second Flood*, Strasbourg: The Council of Europe, 1996.

See: http://www.unesco.or.kr/culturelink/mirror/research/21/cl21_levi.html

See: http://www.georgetown.edu/grad/CCT/tbase/levy.html).

[12] Michael Giesecke, *Der Buchdruck in der frühen Neuzeit. Eine historische Fallstudie über die Durchsetzung neuer Informations- und Kommunikationstechnologien*, Frankfurt am Main: Suhrkamp, 1991. For the mediaeval period See: the masterful study by Brian Stock, *The Implications of Literacy, Written Language and Models of Interpretations in the Eleventh and Twelfth Centuries*, Princeton: Princeton University Press, 1983.

[13] Armand Mattelart, *Transnationals and the Third World. The Struggle for Culture*, South Hadley, Mass.:

Bergin & Garvey, 1985; Ibid, *Mapping World Communication: War, Progress, Culture*, by Armand Mattelart (University of Minnesota Press, 1994; Armand Mattelart y Michèle Mattelart, *Historia de las teorías de la comunicación*, Barcelona: Paidós, 1997.

See: http://www.geoscopie.com/guide/g717opi.html.

[14] “Can Museum Computer Networks Change Our Views of Knowledge?", *Museums and Information. New Technological Horizons. Proceedings*, Ottawa: Canadian Heritage Information Network, (1992), pp. 101-108.

[15] “Frontiers in Conceptual Navigation,” *Knowledge Organization*, Würzburg, vol. 24, 1998, n. 4, pp. 225-245.

[16] Computers and the Transformation of Knowledge", *The Challenge of Lifelong Learning in an Era of Global Change, Couchiching Institute on Public Affairs, 1993 Conference Proceedings*, Toronto, pp. 23-25. “Why Computers are Transforming the Meaning of Education,” *ED-Media and ED-Telecomm Conference, Calgary, June 1997*, ed. Tomasz Müldner, Thomas C. Reeves, Charlottesville: Association for the Advancement of Computing in Education, 1997, vol. II, pp. 1058-1076.

[17] Thoughts on the Reorganization of Knowledge", *Automatisierung in der Klassifikation eV*, ed. Ingetraut Dahlberg (Teil I), Königswinter/Rhein, 5-8. (April 1983), (Frankfurt: Indeks Verlag, 1983), pp.141-150. *(Studien zur Klassifikation*, Bd. 13, SK 13); New Media and New Knowledge", *Proceedings of the Third Canadian Conference on Foundations and Applications of General Science Theory: Universal Knowledge Tools*

and their Applications, Ryerson, 5-8 June 1993, Toronto: Ryerson Polytechnic University, 1993, pp. 347-358.

[18] Dr. Theo Classen, "The Logarithmic Law of Usefulness", *Semiconductor International*, July 1998, pp.176-184. I am grateful to Eric Livermore (Nortel) for this reference.

[19] The definition of usefulness could readily detour into a long debate. For the purposes of this article we shall take it in a very broad sense to mean the uses of computers in terms of their various applications.

[20] Ibid, p.184.

[21] The ISO identifies seven basic layers to the telecommunications network: three which belong to the network layer (physical, data-link, network), one which belongs to the transport layer (transport) and a further three which belong to the user service layer (session, presentation and application cf. figure 15). These seven layers have been applied to computers. With respect to the Internet discussions usually focus on the bottom three layers. These seven layers focus on pipelining and while this is of fundamental value it does not differentiate sufficiently the many elements on the application side:

ISO Layer	Function
1. Network	Physical
2. " "	Data-Link
3. " "	Network
4. Transport	Transport
5. Technical Service	Session
6. " "	Presentation
7. " "	Application.

Figure 15. The seven layers of the ISO.

[22] Laurie McCarthy, Randy Stiles, "Enabling Team Training in Virtual Environments, "*Collaborative Virtual Environments'98*, Manchester, June 1998, pp. 113-121.
See: http://www.isi.edu/vet.

[23] James Pycock, Kevin Palfreyman, Jen Allanson, Graham Button, "Envisaging Collaboration: Using Virtual Environments to Articulate Requirements," *Collaborative Virtual Environments '98*, Manchester: University of Manchester, 1998, pp. 67-79.

[24] See: Steve Mann at http://n1nlf-1.eecg.toronto.edu/index.html.

[25] The potential problems with such responsive environments are actually quite considerable. It is all fine and well to have the television turn on to channel two when A enters the room. But what if B enters the room at the same time, who has programmed the same device to display channel three. What then is the decision strategy? Is it in favour of the older rather than the younger, the owner rather than the guest?

[26] See: http://gen.net/index.htm

[27] See: http://www.npac.sgr.edu/users/gcf/asopmasterB/foilsephtmldir/001HTML.html
http://www.npac.sgr.edu/users/gcf/asopmasterB/fullhtml.html

[28] See: http://ce-toolkit.crd.ge.com

[29] See: http://www.interex.org/hpuxvsr/jan95/new.html#RTFToC33

[30] See: http://www.anxo.com/

[31] Thomas Flaig, "Work Task Analysis and Selection of Interaction Devices in Virtual Environments," in *Virtual Worlds*, ed. Jean Claude Heudin, Berlin: Springer Verlag, 1998, pp. 88-96.

[32] John Mylopoulos, I. Jurisica and Eric Yu, "Computational Mechanisms for Knowledge Organization," in: *Structures and Relations in Knowledge Organization*, ed. Widad Mustafa el Hadi, Jacques Maniez and Steven Pollitt, Würzburg: Ergon Verlag, 1998, p. 126. (Advances in Knowledge Organization, Volume 6).
[33] Charles. S. Pierce, *Collected Papers*, 5400.
[34] T. Sarvimäki, *Knowledge in Interactive Practice Disciplines: An Analysis of Knowledge in Education and Health Care*. Helsinki: University of Helsinki, Department of Education, 1988.
[35] Birger Hjørland, *Information See: king and Subject Representation. An Activity Theoretical Approach to Information Science*, Westport: Greenwood Press, 1997 (New Directions in Information Management, no. 34). He draws also on the ideas of Michael K. Buckland, *Information and Information Based Systems*, New York: Greenwood, 1993.
[36] L. Hjelmslev, *Prolegomena to a Theory of Language*, Madison: University of Wisconsin Press, Madison, 1961.
[37] Jürgen Habermas, Knowledge and Human Interests, London: Heinemann, 1972. German original, 1965.
[38] Hanne Albrechtsen and Elin K. Jacob, "The Role of Classificatory Structures as Boundary Objects in Information Ecologies," *Structures and Relations in Knowledge Organization*, ed. Widad Mustafa el Hadi, Jacques Maniez and Steven Pollitt, Würzburg: Ergon Verlag, 1998, pp. 1-3. (Advances in Knowledge Organization, Volume 6). Cf. Hannah Albrechtsen, *Domain Analysis for Classification of Software, M. A. Thesis*, Stokholm: Royal School of Librarianship, 1993.
[39] Thomas Davenport, *Information Ecology*,xx.
[40] Susan Leigh Star, "The Structure of ill-structured solutions: boundary objects and heterogeneous distributed problem solving," in: *Distributed artificial intelligence*, ed. By L. Gasser and M.N. Huhns, London: Pitman, 1989.
[41] John Law, Multiple Laws of Order, xx.
[42] M. C. Norrie and M. Wunderli, "Coordinating System Modelling," in: *13th International Conference on the Entity Relationship Approach*, Manchester, 1994.
[43] Motschnig Pitrik, R. and John Mylopolous, "Classes and Instances," *International Journal of Intelligent and Cooperative Systems*, 1(1), 1992, pp. xx; John Mylopolous, in: *Information Systems Handbook*, xx:xx, 199xx; Yannis Bubenko, *Conceptual Modelling*, xx:xx, 1998.
[44] See: http://ricis.cl.uh.edu/virt-lib/soft-eng.html. Cf. Kevin Kelly, "One Huge Computer," *Wired*, August 1998, pp. 128-133, 168-171, 182 re: developments in JINI.
[45] Peter P.S. Chen, "The Entity-Relationship Model: Towards a Unified View of Data," *ACM Transactions on Database Systems*, 1 (1), 1976, pp. 9-37. According to F. Miksa (personal communication), this system was further developed while Chen was a professor of computer science at Louisiana State University in 1980.
[46] For some discussion of the philosophical and sometimes subjective assumptions underlying such methods See: : W. Kent, *Data and Reality: Basic Assumptions in Data Processing Reconsidered*. Amsterdam: North-Holland, 1978; H.K. Klein and R.A. Hirscheim, "A Comparative Framework of Data Modelling Paradigms and Approaches," The *Computer Bulletin*, vol. 30, no. 1, 1987, pp. 8-15 and Alan Phelan, "Database and Knowledge Representation: The Greek Legacy," in: *Structures and Relations in Knowledge Organization*, ed. Widad Mustafa el Hadi, Jacques Maniez and Steven Pollitt,

Würzburg: Ergon Verlag, 1998, pp. 351-359. (Advances in Knowledge Organization, Volume 6).

One might expect that librarians, whose lives are dedicated to organizing knowledge should be very sensitive to these problems. In fact, their professional lives are typically spent cataloguing and dealing with materials concerning which the reality is not in question. Each call number applies to a physical book. If there is no physical book in evidence, then it is "because the book is missing," which is typically "a user problem." Their daily work engages them in simple realism. This helps to explain why librarians have frequently accepted and in most cases continue to accept naïve systems such as the entity-relationship model.

[47] Mandelbrot, for instance, noted how the length of the coast of England changed as one changed scale. See: : Benoit Mandelbrot, "How long is the coast of England? Statistical Self-Similarity and Fractal Dimension," *Science*, London, vol. 155, pp.636-638. These ideas were dramatically developed in his major book: Cf. Benoit B. Mandelbrot, *The Fractal Geometry of Nature*, New York, NY: W. H. Freeman and Company, 1982.

[48] This is being developed in the context of the Joint Picture Experts Group (JPEG). See: http://www.jpeg.org and http://www.periphere.be/lib/jpeg.htm; and http://www.gti.ssr.upm.es/~vadis/faq_MPEG/jpeg.html. This is CCITT/ISO(JTC1/SC2/WG10 and has the following standards:

T.80 Common components for image and communication-basic principles
T.81 Digital compression and encoding of continuous tone still image
T.82 Progress compression techniques for bi-level images
T.83 Compliance testing

As well as Still Picture Interchange File Format[48] (SPIFF), Registration Authority (REGAUT) and JPEG Tiled Image Pyramid (JTIP), their spokesperson, Jean Barda, has developed a System of Protection for Images by Documentation iDentification and Registration of digital files (SPIDER), which combines two important elements:

(1) a system for registering ownership over an image

(2) metadata tags embedded within the image (header and directory) that identify the image and its owner. SPIDER is one of the first applications to employ SPIFF, the newly developed ISO standard designed to supersede the JFIF/JPEG file storage format. AVELEM, the company that developed SPIDER, also has built a system called Saisie numerique et COnsultation d'images PYRamidales (SCOPYR), i.e. Digital image capture and exploitation of pyramidal images.

See: http://www.sims.berkeley.edu/courses/is290-1/f96/Lectures/Barda/index.html.

[49] Bruce Damer, *Avatars!, Exploring and Building. Virtual Worlds on the Internet*, Berkeley: Peachpit Press, 1998.

[50] See: http://www.cs.man.ac.uk/mig/giu/.

[51] See: http://madsci.wustl.edu/~lynn/VH/

[52] See: http://www.igd.fhg.de/www/igd-a7/Projects/OP2000/op2000_e.html

[53] See: http://www2.igh.cnrs.fr/HUM-Genome-DB.html

[54] Carol A. Bean, "The Semantics of Hierarchy: Explicit Parent-Child Relationships in MeSH tree structures," *Structures and Relations in Knowledge Organization*, ed. Widad Mustafa el Hadi, Jacques Maniez and Steven Pollitt, Würzburg: Ergon Verlag, 1998, pp. 133-138. (Advances in Knowledge Organization, Volume 6).

[55] To take a hypothetical example, suppose it was decided that a "normal" male is 6 feet in height. Hence, the whole range of variation from small pygmies (c.3 feet) to very tall persons (over 7 feet) would require "modification."

[56] Jmankoff@cc.gatech.edu

[57] Jennifer Mankoff, "Bringing People and Places Together with Dual Augmentation," *Collaborative Virtual Environments*, Manchester, June 1998, pp. 81-86.

See: http://www.cc.gatech.edu/fce/domisilica.

[58] For an alternative and more subtle classification See: : Didier Verna, Alain Grumbach, "Can we Define Virtual Reality? The MRIC Model", *Virtual Worlds '98*, ed. J.C. Heudin, Berlin: Springer Verlag, 1998, pp. 29-41.

[59] Gianpaolo U. Carraro, John T. Edmark, J. Robert Ensor, "Pop-Out Videos," in: *Virtual Worlds*, Berlin: Springer Verlag, pp.123-128. The author distinguishes between two distinct types integrated video space and complementary videos. In an integrated video space a CAD space and a video space are merged. In a complementary video one might be watching a golf player whose swing interests one. A CAD version of the player would allow one to view the person who had been merely two-dimensional in the video image from all directions in the model. See: also: Gianpaolo U. Carraro, John T. Edmark, J. Robert Ensor, "Techniques for handling Video in Virtual Environments," *SIGGRAPH 98, Computer Graphics Proceedings, Annual Conference Series*, New York, 1998, pp. 353-360.

[60] For a more balanced assessment See: : William Mitchell, *The Reconfigured Eye*, Cambridge, Mass.: MIT Press, 1992.

[61] Yaneer Bar-Yam, *Dynamics of Complex Systems (Studies in Nonlinearity)*, Reading, MA: Perseus Press, 1997. Cf. http://www.necsi.org/mclemens/viscss.html. For a rather different approach to complex systems See: : John L. Casti, *Would-Be Worlds. How Simulation is Changing the Frontiers of Science*, New York: John Wiley and Sons, 1997.

[62] See: http://www.necsi.org/html/complex.html

[63] This is not to say of course that there cannot be a history of science and technology. There definitely is and it is essential that we continue to foster awareness of that history. Without a clear notion of the steps required to reach our present machines for working in the world and models for understanding that world, it would be all but impossible to understand many aspects of present day science, and we would be in sore danger of constantly re-inventing the wheel.

[64] For a brief history See: the excellent study by Willian Ivins, Jr., *Prints and Visual Communication*, Cambridge, Mass.: Harvard University Press, 1953.

[65] See: André Chastel, *Le grand atelier de l'Italie*, Paris: Gallimard, 1965.

[66] Benjamin R. Barber, *Jihad vs. McWorld*, New York: Times Books, 1995.

[67] It may be true that the masterpieces of art also represent a selection from the many particulars, but the masterpieces are not generalizations of the rest: they remain individuals per se.

[68] I owe this distinction to my colleague and friend André Corboz, who notes that although sculpture and architecture are static in themselves, they require motion on the part of the observer in order to be See: n completely from a number of viewpoints.

[69] There are of course histories of these dynamic subjects but their contents are limited to verbal descriptions and give no idea of the richness of performances. In the case of dance there have been some attempts to devise printed notations, which can serve as summaries of the steps involved. In the case of music there are of course recordings. More recently there are also films and videos to cover performances of dance and theatre.

[70] China, Japan and India also had rich traditions of theatre and dance which, for the reasons being discussed, were typically ignored until quite recently.

[71] Victor Mair, *Painting and Performance,* Honolulu: University of Hawaii Press, 1988. On this topic I am grateful to Niranjan Rajah who gave a lecture "Towards a Universal Theory of Convergence. Transcending the Technocentric View of the Multimedia Revolution," at the Internet Society Summit, Geneva, July 1998.

See: niranjan@faca.unimas.my).

[72] The Metadata Coalition at http://www.metadat.org is a group of 50 software vendors and users including Microsoft with a 7 member council that has voted to support Microsoft Repository Metadata (Coalition) Interchange Specification[72](MDIS).

See: http://www.he.net/~metadata/papers/intro97.html.

[73] This includes Arbor, Business Objects, Cognos, ETI, Platinum Tech, and Texas Instruments (TI).

See: http://www.cutech.com/newmeta.html.

[74] See: http://environment.gov.au/newsletter/n25/metadata.html

[75] For basic articles on meta-data See: : Francis Bretherton, "A Reference Model for Metadata" at http://www.hensa.ac.uk/tools/www.iafatools/references/whitepaper/whitepaper.bretherton.html;
"WWW meta-indexes" at http://www.dlr.de/search-center-meta.html and Larry Kirschberg, "Meta World: A Quality of Service Based Active Information Repository, Active Data Knowledge Dictionary", at http://isse.gmu.edu/faculty/kersch/Vita-folder/index.html. For books See: : *Computing and Communications in the Extreme. Research for Crisis Management and Other Applications at* http://www.nap.edu/readingroom/books/extreme/chap2.html.

[76] For basic definitions of metadata

See: http://204.254.77.2/bulletinsuk/212e_1a6.htm and the Klamath Metadata Dictionary at http://badger.state.wi.us/agencies/wlib/sco/metatool/kmdd.htm. A basic taxonomy of metadata is available.

See: http://www.1bl.gov/~olken/EPA/Workshop/taxonomy.html. See: also the CERA metadata model at http://www.dkrz.de/forschung/reports/reports/CERA.book.html.

[77] These are defined in the Generic Top Level Domain Memorandum of Understanding at http://www.gtld_mou.org/ and include:

Two letter country codes	(ISO 3166)
Generic Top Level Domains	(gTLD)
com	
net	
org	

new: firm
shop
web
arts
rec
info
nom
International Top Level Domain (iTDL)
int
Special US only
gov
mil
edu
Internal Systems
arpa

[78] The URL began as a simple resource locator for basic internet protocols, such as:
file
Gopher
http
news
telnet.
Figure 16. Basic categories covered by URLS.

A more universal approach to resource location is foreSee: n in the evolving Uniform Resource Indicators (URI) as listed in note 95.

[79] See: http://www.iso.ch/cate/d6898.html
[80] See: http://www.iso.ch/cate/d18931.html
[81] See: http://www.iso.ch.iso/cate/d18506.html
[82] See: http://www.acl.lanl.gov/URN/FPI-URN.html
[83] See: http://www.issn.org
[84] See: http://www.cisac.org/eng/news/digi/ensymp972.htm
[85] In draft in ISO TC46sc9.
[86] This includes specification of ISRC related metadata which is linked with MUSE, an EC funded initiative of the record industry which is due to announce (c. October 1998) a secure means for encoding and protecting identifiers within digital audio.
[87] See: http://www.tlcdelivers.com/tlc/crs/Bib0670.htm
[88] See: http://www.elsevier.nl/inca/homepage/about/pii. On these identifiers from the publishing world. Cf. an article by Norman Paskin.
See: http://www.elsevier.co.jp/inca/homepage/about/infoident.
[89] See: http://www.handle.net/doi/announce.html. Cf. http://pubs.acs.org/journals/pubiden.html and http://www.doi.org. This began in the book and electronic publishing industry but is attracting wider membership.
[90] See: http://purl.oclc.org
[91] See: http://www.cs.princeton.edu/~burchard/gfx/bg.marble.gif
[92] Concerning HTML 3
See: http://vancouver-webpages.com/Vwbot/mk-metas.html. Meta Tags include:
Title

Description
Keywords
Subject
Creator
Publisher
Contributor
Coverage: Place Name
Coverage: xyz
Owner
Expires
Robots
Object Type
Rating
Revisit .

[93] See: http://hegel.ittc.ukans.edu/topics/internet/internet-drafts/draft-l/draft-leach-uuids-guids-00.txt.

Cf. http://www.icsuci.edu/~ejw/authoring/rd_tri.gif.

[94] See: http://www.hensa.ac.uk/tools/www.iafatools/slides/01.html. This is being developed by the Internet Anonymous FTP Archives Working Group, whose templates on Internet Data are based on whois++ and include:

URI
File System
Contents
Author
Site Administrator
Another Metadata Format.

This model is being applied to ROADS.

[95] See: http://www.dbr/~greving/harvest_user_manual/node42.html

[96] The purposes of MCF are:

1. Describe structure of website or set of channels
2. Threading e-mail
3. Personal Information Management functions (PIM)+
4. Distributed annotation and authoring
5. Exchanging commerce related information such as prices, inventories and delivery dates.

It will use a Directed Linked Graph which contains:

URL
String
E-mail
Author
Person
Size
Integer

It will also use Distribution and Replication Protocol (DRP) developed by Netscape and Marimba.

The MCF began at Apple Computers. See: : R.V. Guha, “Meta Content Framework,” *Apple Technical Report 167,* Cupertino, March 1997.

See: http://mcf.research.apple.com/mcf.html. Guha then moved to Netscape and developed the Meta Content Framework with Tim Bray of Textuality.

See: http://www.textuality.com/mcf/NOTE-MCF-XML.html.

[97] Web Collections will include:

Web Maps
HTML e-mail Threading
PIM Functions
Scheduling
Content Labelling
Distributed Authoring
It uses XML to provide hierarchical structure for this data.

See: http://www.w3.org/TR?NOTE-XML.submit.html.

[98] For IFLA metadata

See: http://www.nlc-bnc.ca/ifla/II/metadata.htm.

[99] International Standard Book Description (ISBD) has eight basic features:

1. Title and Statement of Responsibility Area
2. Edition Area
3. Place of publication) specific area
4. Publication Distribution etc. area
5. Physical Description Area
6. Series
7. Note Area
8. Standard Number (or alternative) and terms of availability

It should be noted that ISBD has a series of other divisions, namely:

Antiquarian	ISBD (A)
Monographs	ISBD (M)
Non Book Materials	ISBD (NBM)
Printed Music	ISBD (PM)
Serials	ISBD (S).

[100] See: http://omni.nott.ac.uk

[101] See: http://lcweb.loc.gov/marc

[102] The MARC/UNIMARC format uses ISO Z39.50. It is applied to OCLC’s Netfirst. There are plans to link this with a URC to create a MARC URC. The MARC record comes in many variants including:

Machine Readable Record of Bibl. Info.		MARBI
US “	“	USMARC
UK	“	UKMARC
UNI “	“	UNIMARC
INTER”	“	INTERMARC
Canadian “	“	CANMARC
Danish “	“	DANMARC
Finnish “	“	FINMARC
Libris “	“	LIBRISMARC
Norwegian MARC	“	NORMARC

South African " " SAMARC
Iberian (i.e. Spanish)" " IBERMARC
Norway " " NORMARC

[103] See: Bernhard Eversberg, *Was sind und was sollen Bibliothekarische Datenformate*. Braunschweig: Universitätsbibliothek der TU, 1994. See: http://ubsun01.biblio.etc.tu-bs.de/acwww25/formate/formate.html.

[104] The Association for Library Collections and Technical Services (ALCTS:DA) has a Committee on Cataloging Description and Access at http://www.lib.virginia.edu/ccda and is engaged in mapping of SGML to MARC and conversely.
See: http://darkwing.uoregon.edu/mnwation/ccdapage/index.html.

[105] See: http://www.loc.gov/rr/ead/eadhome.html. Berkeley is also involved in EAD with a view to creating UEAD URC.
See: http://sunsite.Berkeley.EDU/ead.

[106] Anne Gilliland-Swetland, "Defining Metadata," in: *Introduction to Metadata. Pathways to Digital Information*, ed. Murtha Baca, Los Angeles: Getty Information Institute, 1998, pp. 1-8. The author also lists eight attributes of metadata: source of metadata, method of metadata creation, nature of metadata, status, structure, semantics and level. Also in this booklet is a useful attempt to map between some of the major metadata standards: Categories for the Description of Works of Art (CDWA), Object ID, the Consortium for the Interchange of Museum Information (CIMI), Foundation for Documents of Architecture/Architectural Drawings Advisory Group (FDA/ADAG), Museum Educational Site Licensing (MESL) project, Visual Resources Sharing Information Online (VISION), Record Export for Art and Cultural Heritage (REACH), United States Machine Readable Cataloging (US MARC) and the Dublin Core. While providing an excellent survey of existing efforts towards standards in the United States, this list does not reflect a comprehensive picture of efforts around the world.

[107] See: http://nt.comtec.co.kr/doc/uri/urc-s/html/scenarios_2.html

[108] Chris Weider
Bunjip Information Systems
2001 S. Huron Parkway
Ann Arbor Michigan 48104
USA
1-313-971-2223.

[109] See: http://whirligig.ecs.soton.ac.uk/~ng94/project/names/urndefl.htm

[110] See: http://www.acl.lanl.gov/URC/

[111] It is not surprising that major projects have plans to link their projects with Uniform Resource Characteristics (URC), notably:

IAFA Templates	(IAFA URC)
Machine Readable Record	(MARC URC)
Encoded Archival Description[111]	(EAD URC)
Text Entering Initiative	(TEI URC)
Consortium for Interchange of Museum Information	(CIMI URC).

[112] This uses whois++ for URN resolution; relies on DNS; element ID=Naming Authority; uses Whois ++ for resource identification. It is similar to IAFA templates but they have different contents depending on type of object.

URC Method: Whois++[112] URC + dns mapping

URN	Uniform Resource Name	(URN)
URL	Uniform Resource Locator	(URL)
LIFN	Location Independent File Name	(LIFN)

Author Author Name
Title Resource Title
Abstract Short Description
HTTP headers
Content Length
Content Type
Content Language

See: Weider Mitra, M. Mealing, "Uniform Resource Names (URN)". Internet Draft (work in progress) IETF, November 1994.
See: ftp://ds.internic.net/internet-drafts/draft-ietf-uri-resource-names-03.txt,

113 This uses http for URN resolution, relies on DNS, separates Element ID into: Authoring ID and Request ID. Browsers are encouraged to support URCs returned as HTML or plain text. Associated URC format used to return results of a URN lookup in a form suitable for automatic processing.

URC Method: Trivial URC +x-dns-2 mapping
ftp://………..
Abstract……….
ftp: mirror site
ftp: " "
ftp: " "
Language
Character
Uses http.

Cf. Ron Daniel, Los Alamos, IETF Draft 1995. Paul E. Hofman, Ron Daniel, "Trivial URC Syntax:urc0". Internet Draft (work in progress), IETF May 1995.
See: ftp://ds.internic.net/internet-drafts/draft-ietf-uri-urc-trivial-00.txt

114 Ron Daniel, T. Allen, "An SGML based URC service". Internet Draft (work in progress), IETF, June 1995.
See: ftp://ds.internic.net/internet-drafts/draft-ietf-uri-urc-sgml-00.txt

115 Daniel LaLiberte and Michael Shapiro, IETF draft 1995.
See: http://www.hypernews.org/~liberte/www/path.html.
Other competing solutions to this challenge described by Martin Hamilton entail directory services such as SDP or SOLO.

116 See: http://www.dstc.edu.au/RDU/Apweb96/index.html

117 See: http://www.dstc.edu.au/RDU/TURNIP

118 See: http://me-www.jrc.it/~dirkx/ewse-urn-turnip.html. This is headed by Dirk-Willem van Gulik:
Dirk.vanGulik@jrc.it.

119 See: http://archie.mcgill.ca/research/papers/

120 See: http://archie.mcgill.ca/research/papers/1995/uradraft.txt

121 See: http://www.hensa.ac.uk/tools/www.iafatools/slides/01.html

122 See: http://ds.internic.net/ietf/iiir/iiir-character.tx

[123] See: http://www.imc.org/ietf-calendar/stif.txt

[124] RFC 1625 WAIS over Z39.50:1988
1729 Z39.50 in TCP/IP (Lynch)
1728 Resource Transponders (Weider)
1727 Integrated Internet Information Service.

[125] See: http://services.bunyip.com:8000/research/papers/1996/cip/cip.html;
cf. http://www.ietf.org/html.charters/find-charter.html

[126] See: http://ds.internic.net/rfc/rfc2244.txt

[127] See: http://www.ics.uci.edu/~ejw/authoring/

[128] See: http://csvu.nl/~eliens/wwww5/papers/Meta.html

[129] See: http://www.oclc.org:5046/~weibel/html-meta.html

[130] The list of resources called by the URI includes the following:

about
callto
content id (cid)
cisid
data
file
finger
file transfer protocol (ftp)
gopher
CNRI handle system (hndl)
hyper text transfer protocol (http)
hyper text transfer protocol over secure sockets layer (https)
inter language unification (ilu)
internet mail protocol (imap)
Internet Object Request (IOR)
internet relay chat (irc)
java
javascript
java database connectivity (jdbc)
lightweight directory application protocol (ldap)
location independent file name[130] (lifn)
livescript
mailto
mailserver
md5
message id (mid)
mocha
Network File System (NFS)
network news transport protocol (nntp)
path
phone
prospero
rwhois
rx

Short Message Service (Service)
Session Initiation Protocol (SIP)
Session hyper text transfer protocol (shttp)
Stable Network Filenames (STANF)
telnet
tv
Enhanced Man Machine for Videotex and Multimedia (VEMMI)
videotex
view-source
wide area information servers (wais)
whois++
whodp
z39.50r
z39.50s
For a full list See: http://www.w3.org/Addressing/schemes.

[131] Sometimes called Structured Graphic Markup Language (SGML). This grew out of IBM's Generalized Markup Language (GML) and Gen Code of the Graphic Communications Association (GCA). There have been projects to map SGML to MARC as mentioned in note 84 above.

[132] HyperText Markup Language (HTML) is now under ISO/IEC JTC1/SC 18/WG8 N1898. See: http://www.oml.gov/sgml/wg8/document/1898.html. Specialized versions include compact HTML for Small Information Appliances at http://207.82.250.251/cgi-bin/start. Attempts to expand the scope of HTML have led to Simple HTML Ontology Extensions (SHOE) at http://www.cs.umd.edu/projects/plus/SHOE and a Dictionary of HTML Meta Tags[132] See: http://vancouver-webpages.com/META/metatags.detail.html.

[133] See: http://www.cogsci.ed.ac.uk/~ht/new-xml-link.html

As is so frequently the case, Microsoft has copied the ideas and created its own proprietary versions: Microsoft Extensible Markup Language (XML)
XML based data transfer (Microsoft XML)
Extensible Style Language (XSL)
Microsoft Channel Description Format (CDF).

For basic literature concerning XML See: :
Jon Bosak, "XML, Java and the Future of the Web" at http://sunsite.unc.edu/pub/sun/-info/standards/xml/why/xmlapps.htm; Rohit Khare, "XML. The Least you need to Know" at http://www.cs.caltech.edu/~adam/papers/xml/tutorial and Richard Light, *Presenting XML*, S. Net, August 1997. For XML Software Tools See: http://www.cs.caltech.edu/~adam/local/xml.html. For Java Object Stream to XML Packages See: http://www.camb.opengroup.org/~laforge/jsxml/. For a Lark Non-Validating XML Parser in Java See: http://www.textuality.com/Lark/. XML links with the Document Object Model. There is also an XML API in Java (XAPI-J) and XML Typing.
Cf. Extensible Hyper Language (EHL) at http://www.cogsci.ed.ac.uk/~ht/new-xml-link.html.

[134] See: http://207.82.250.251/cgi-bin/start

[135] For a second version of Cascading Style Sheets no. 2 (CSS2)
See: http://207.82.250.251/cgi-bin/start

[136] This is being developed into Extensible Style-Document Style Semantics and Specification Language (XS-DSSSL).
[137] Schema of key architecture elements in the W3 plan for linking XML with other elements described at http://www.w3.org/TR/NOT-rdfarch:

XML application RDF application PICS 2.0 P3P

RDF-semantics
XML-structure

[138] See: http://www.w3.org/PICS/NG
[139] See: http://www.w3.org/Talks/9707/Metadata/slide8.htm
[140] See: http://www.w3.org/TR/WD-rdf-syntax
[141] This includes:

Document
Element
Attribute
Text
Comment
Processing Instruction (PI).
See: http://www.w3.org/DOM.

[142] RDMF entails :

RD Retrieval
RD Submission
Server Description
Schema Description
Taxonomy Description
Status Retrieval.

See: http://www.w3.org/TR/NOTE-rdm.html

[143] Some observers such as Khare propose a different way of looking at the roles of the different markup languages:

Syntax	SGML
Style	CSS/XSL
Structure	HTML
Semantics	XML.

[144] The following are Basic W3 Metadata Plans as of May 1998:

Metadata Syntax Specification (RDF)
Language for RDF schemata
Language for expressing filters (simple boolean functions of) RDF
Algorithm for canonicalizing an RDF expression for digital signature
Syntax for digitally signing RDF expressions
Vocabulary for expressing PICS labels in RDF and a
Conversion algorithm from PICS 1.1.

cf. http://www.ics.forth.gr/ICS/acti/netgroup/documents/TINAC/

A fifth method of URN to URL Mapping (Resource Discovery) has been developed by ARPA , called Handle, which uses a Repository Access Protocol (RAP).

ARPA has also been very active in the development of a Knowledge ARPA Knowledge Sharing Effort (ARPA KSE). (See: http://www.cs.umbc.edu/agents/kse.shtml), which entails both a Knowledge Query Manipulation Language (KQML.
See: http://cs.umbc.edu/kqml) and a Knowledge Interchange Format (KIF.
See: http://logic.stanford.edu/kif/kif.html).

[145] See: http://renki.helsinki.fi/z3950/3950pr.html. There is a great deal of information available on Z39.50. For the Maintenance Agency homepage
See: http://lcweb.loc.gov/z3950/agency/. Hosts available for testing are at http://lcweb.loc.gov/z3950/agency/objects/iso-pub.html; Implementor Agreements at http://lcweb.loc.gov/z3950/agency/objects/agree.html; Interoperability Testing at http://lcweb.loc.gov/z3950/agency/objects/testbed.html; Implementors Register at http://lcweb.loc.gov/z3950/agency/register.html.

In addition the National Institute of Science and Technology (NIST) has Implementation Papers at http://lcweb.loc.gov/z3950/agency/nist.html; Object Identifiers at http://lcweb.loc.gov/z3950/agency/iso-pub.html; Registered objects and other definitions at http://lcweb.loc.gov/z3950/agency/objects.html; SQL Extensions at http://www.dstc.edu.au/DDU/projects/ZINC/proposal3.ps; Version 4 Development at http://lcweb.loc.gov/z3950/agency/version4.html. There is both a Z39.50 Implementors Group (ZIG)[145] and a Z39.50 Users Group (ZUG). Profiles are available at http://lcweb.loc.gov/z3950/agency/profiles/about.html; further information about at http://lcweb.loc.gov/z3950/agency/profiles/about.html; in an Internet environment at ftp://ds.internic.net/rfc/frc1729.txt; Access to digital collections at http://lcweb.loc.gov/z3950/agency/profiles/collections.html; Access to digital library objects at http://lcweb.loc.gov/z3950/agency/profiles/dl.html; and a CIMI Profile at http://lcweb.loc.gov/z3950/agency/profiles/cimi2.html.

[146] See: http://lcweb.loc.gov/z3950/agency/1992doc.html

[147] See: http://lcweb.loc.gov/z3950/agency/1995doc.html

[148] See: http://www.cni.org/pub/NISO/docs/Z39.50-1992/www/50.brochure.toc.html

[149] This uses 63 attributes including:

Personal Name
Conference Name
Title
Title Uniform
ISBN
ISSN
LC Card Number
Relation
Position
Truncation
Structure
Completeness.

[150] See: http://www.oclc.org:5046/oclc/research/conferences/metadata
http://purl.oclc.org/net/eric/DC/syntax/metadata.syntax.html

[151] See: the Dublin Core Home Page at http://purl.org/metadata/dublin_core. The Meta2 Archive by thread is at http://www.roads.lut.ac.uk/lists/meta2/. Important contacts are Dr.

Clifford Lynch at http://www.sciam.com/0397/issue/0397lynch.html; Stuart Weibel at http://www.dlib.org/dlib/July95/07/weibel.html and John Kunze.

[152] See: http://www.roads.lut.ac.uk/Metadata/DC-SubElements

[153] See: http://www.oclc.org:5046/oclc/research/conferences/metadata2/
http://wwwblib.org/dlib/july96/07weibel.html

[154] See: http://www.oclc.org:5046/conferences/imagemeta/index.html. This problem has also been pursued elsewhere by Clifford Lynch.
cf. http://www.sciam.com/0397/issue/0397lynch.html) through the
Coalition of Networked Information (CNI), which held a Metadata Workshop on Networked Images.
See: http://purl.oclc.org/metadata/image. For other projects on image metadata
See: http://www.dlib.org/dlib/january97/oclc/01weibel.html.

[155] See: http://www.dstc.edu.au/DC4

[156] See: http://www.linnea.helsinki.fi/meta/DC5.html.
cf. http://www.ariadne.ac.uk/issue12/metadata.

[157] Ron Knight, "Will Dublin form the Apple Core?" at http://www.ariadne.ac.uk/issue7/mcf.

[158] Carl Lagoze, Clifford Lynch, Ron Daniel, Jr., "The Warwick Framework: A Container for Aggregating Sets of Metadata."
See: http://cs-tr.cs.cornell.edu/~lagoze/container.html:

Container
Package
Dublin Core

Package
MARC Record

Package
Indirect reference ------→ Package
Terms and Conditions

Relationship for a Container
Relationship Package
Dublin Core Record
Content Package
Access Core List
URN
Other Package
MARC Record

Figures 17-18. Two further diagrams showing possible ways of integrating Dublin Core Elements with those of MARC records and other standard library formats.

[159] See: http://harvest.transarc.com

[160] See: http://www.dbr/~greving/harvest_user_manual/node42.html. This was developed by Michael Schwartz.

[161] See: http://www.ukoln.ac.uk/metadata/DESIRE/overview/rev_02.htm

[162] See: http://www.ukoln.ac.uk/metadata/DESIRE/overview

[163] Attribute List
Abstract
Author

Description
File Size
Full Text
Gatherer Host
Gatherer Name
Gatherer Port
Gatherer Version
Update Time
Keywords
Last Modification Time
MD5 16 byte checksum of object
Refresh rate
Time to Live
Title

[164] Harvest Template Types:

Archive
Audio
FAQ
HTML
Mail
Tcl
Troff
Waissource.

[165] Other harvest features include:

Identifier
Value
Value Size
Delimiter
URL References

[166] For insight into the British Library's efforts See: : *Towards the Digital Library. The British Library's Initiatives for Access Programme*, ed. Leona Carpenter, Simon Shaw, Andrew Prescott, London: The British Library, 1998. Readers are also referred to the library's publication: *Initiatives for Access. News.*

[167] See: http://www2.cornell.edu/lagoze/talks/austalk/sld014.htm

[168] See: http://www.ipl.org

[169] See: http://www.dstc.edu.au/RDU/pres/nat-md/

[170] See: http://www.dstc.edu.au/RDU/pres/www5

[171] URC Data includes:

Title
Author
Identifier
Relation
Language.

[172] See: http://ruby.omg.org/index.htm

[173] See: http://www.omg.org/corbserver/relation.pdf

[174] This is connected with the Inter-Language Unification (ILU) project of Xerox PARC at http://parcftp.parc.xerox/pub/ilu/ilu.htm, which is producing an Interface Specification Language (ISL). It is not to be confused with Interactive Data Language (IDL) See: http://ftp.avl.umd.edu/pkgs/idl.html.

[175] See: http://www.omg.org/docs/telecom/97-01-01.txt
Cf. http://www.igd.fhg.de/www/igd-a2/conferences/cfp_tina97.html

[176] See: http://www.he.net/~metadata/papers/intro97.html

[177] See: http://mnemosyne.itc.it:1024/ontology.html. Cf. http://www-ksl.stanford.edu/kst/wahat-is-an-ontology.html.

[178] See: http://www.infospheres.caltech.edu

[179] See: http://viu.eng.rpi.edu

[180] See: http://viu.eng.rpi.edu/viu.html

[181] See: http://viu.eng.rpi.edu/IBMS.html

[182] See: http://www.parc.xerox.com/spl/projects/mops

[183] See: http://dri.cornell.edu/Public/morgenstern/MetaData.html
cf. http://dri.cornell.edu/pub/morgenstern/slides/slides.html

[184] See: http://www2.infoSee: k.com/

[185] See: http://www.intellidex.com. Cf. John Foley, “Meta Data Alliance”, *Information Week*, Manhasset, NY, January 27 1997, p. 110.

[186] See: http://www.carleton.com/metacnt1.html

[187] See: http://localweb.nexor.co.uk

[188] For a further list of software
See: http://www.ukoln.ac.uk/metadata/software-tools. For a list of tools
See: http://badger.state.wi.us/agencies/ulib/sco/metatool/mtools.htm.

[189] See: http://www.pastel.be/mundaneum/. Cf.W. Boyd Rayward, *The Universe of Information. The Work of Paul Otlet for the Documentation and International Organisation*, Moscow, 1975.

[190] Based on its French name: Fédération Internationale de la Documentation

[191] Based on its French name: Union Internationale des Associations.
See: http://www.uia.org.

[192] UNISIST. *Synopsis of the Feasibility Study on a World Science Information System*, Paris: UNESCO, 1971, p. xiii.

[193] See: http://www.unesco.org/webworld/council/council.htm

[194] See: http://www.icpsr.umich.edu/DDI. Cf. the Association for Information and Image Management International (AIIMI) which organizes the Document Management Alliance at http://www.aiim.org/dma.
Cf. also the European Computer Manufacturers Association (ECMA) which has produced the Script Language Specification (ECMA 262) at http://www.ecma.ch/index.htm and http://www.ecma.ch/standard.htm.

[194] See: http://www.sdsc.edu/SDSC/Metacenter/MetaVis.html#3 which provides electronic addresses for all of the above.

[195] See: http://www.sdsc.edu/SDSC/Metacenter/MetaVis.html#3 which provides electronic addresses for all of the above.

[196] See: http://www.ru/gisa/english/cssitr/format/ISO8211.htm

[197] See: http://www2.echo.lu/oii/en/gis.html#IHO

[198] See: http://www2.echo.lu/oii/en/gis.html#ISO15046

[199] See: http://www2.echo.lu/oii/en/gis.html#ISO6709
[200] See: http://www.stalk.art.no/isotc211/welcome.html
[201] See: http://cesgi1.city.ac.uk/figtree/plan/c3.html.
ISO/IEC/TC 211):
WG1. Framework and reference model
WG 2. Geospatial data models and operators
WG 3. Geospatial data administration
WG 4. Geospatial Services
WG 5. Profiles and Functional Standards.
[202] See: http://www2.echo.lu/oii/en/gis.html#ITRF
[203] See: http://ilm425.nlh.no/gis/cen.tc287
[204] See: http://www2.echo.lu/impact/oii/gis.html#GDF
[205] See: http://www2.echo.lu/vii/en/gis.html#GDF
[206] See: file://waisvarsa.er.usgs.gov/wais/docs/ASTMmeta83194.txt
[207] See: http://sdts.er.usgs.gov/sdts/mcmcweb.er.usgs.gov/sdb
[208] See: http://fgdc.er.usgs.gov/metaover.html
[209] See: http://geochange.er.usgs.gov/pub/tools/metadata/standard/metadata.html
cf. http://www.geo.ed.ac.uk/~anp/metaindex.htm
[210] Michael Potmesil, "Maps Alive: Viewing Geospatial Information on the WWW," *Sixth WWW Conference*, Santa Clara, April 1997, TEC 153, pp.1-14. See: http://www6.nttlabs.com/HyperNews/get/PAPER/30.html
[211] See: http://www.research.ibm.com/research/press
[212] See: http://ds.internic.net/z3950/z3950.html which provides a list of available implementations.
[213] See: http://www.grid.unep.no/center.htm
[214] See: http://gelos.ceo.org/free/TWG/metainofrmation.html
[215] See: http://info.er.usgs.gov/gils. Cf. Eliot Christian, "GILS. Where is it where is it going?"
See: http://www.dlib.org/dlib/december96/12christian.html.
[216] See: http://www.wcmc.org.uk/
[217] See: http://www.erin.gov.au/general/discussion-groups/ozmeta-l/index.html
[218] See: http://www.epa.gov/edu
[219] On Harmonization of Environmental Measurement See: : : Keune, H., Murray, A. B, Benking, H.in: *GeoJournal*, vol. 23 no. 3, March 1991, pp.149-255 available on line at:
http://www.ceptualinstitute.com/genre/benking/harmonization/harmonization.htm
On Access and Assimilation: Pivotal Environmental Information Challenges,
See: *GeoJournal*, vol. 26, no. 3, March 1992, pp. 323-334 at:
http://www.ceptualinstitute.com/genre/benking/aa/acc&assim.htm
[220] See: http://www.lbl.gov/~olken/epa.html
[221] See: http://www.llnl.gov/liv_comp/metadata/metadata.html
[222] See: http://www.psc.edu/Metacenter/MetacenterHome.html
[223] See: http://www.khoral.com/plain/home.html
[224] See: http://www.nbs.gov/nbii
[225] See: http://www.nbii.gov/
[226] See: http://www.cs.mu.oz.au/research/icaris/bsr.html
[227] See: http://www.cs.mu.oz.au/research/icaris/beacon.html

[228] See: http://www.sdsc.edu/SDSC/Metacenter

[229] See: http://www.usgs.gov/gils/prof_v2html#core

[230] See: http://jetta.ncsl.nist.gov/metadata

[231] See: http://www.oberlin.edu/~art/vra/core.html

[232] See: http://www.unesco.org/webworld/telematics/uncstd.htm

[233] See: http://www2.echo.lu/oii/en/library.html

[234] See: http://www.mpib-berlin.mpg.de/dok/metadata/gmr/gmrwkdel.htm

[235] See: http://www.ukoln.ac.uk/metadata/
http://www.ukoln.ac.uk/metadata/interoperability
http://www.ukoln.ac.uk/dlib/dlib/july96/07dempsey.html
http://www.ukoln.ac.uk/ariadne/issue5/metadata-masses/

[236] See: http://ahds.ac.uk/manage/proposal.html#summary

[237] See: http://www.ukoln.ac/metadata/

[238] See: http://www.ukoln.ac.uk/metadata/interoperability

[239] See: http://www.roads.lut.ac.uk/Reports/arch/

[240] See: http://omni.nott.ac.uk

[241] Godfrey Rust, "Metadata: The Right Approach. An Integrated Model for descriptive and Rights Metadata in E-Commerce," *D-Lib Magazine*, July-August 1998, at: http://www.dlib.org/dlib/july98/rust/07rust.html.

[242] See: http://www.dbc.dk/ONE/oneweb/index.html

[243] See: http://portico.bl.uk/gabriel/

[244] See: http://www.infobyte.it

[245] See: http://www.gii.getty.edu/vocabulary/tgn.html

[246] Cf. the interesting work by Dr. A. Steven Pollitt (Huddersfield University, CeDAR), "Faceted Classification as Pre-coordinated Subject Indexing" at: http://info.rbt.no/nkki/korg98/pollitt.htm. I am very grateful to Dr. Pollitt for making me aware of his work. Some believe that traditional discipline based classifications are outmoded in an increasingly interdisciplinary world. For instance, Professor Clare Beghtol, *The Classification of Fiction: The Development of a System Based on Theoretical Principles*, Metuchen, N.J.: Scarecrow Press, 1994, believes that the distinction between non-fiction and fiction is no longer relevant since both categories entail narrative. Meanwhile, Nancy Williamson, "An Interdisciplinary World and Discipline Based Classification," *Structures and Relations in Knowledge Organization*, ed. Widad Mustafa el Hadi, Jacques Maniez and Steven Pollitt, Würzburg: Ergon Verlag, 1998, pp. 116-132. (Advances in Knowledge Organization, Volume 6), although sceptical about replacing disciplines entirely, has explored a series of alternative short-term solutions. Since university and other research departments continue to be discipline based it may be sensible to maintain what has been the starting point for all classification systems for the past two millenia, and work on creating new links between, among these disciplines.

[247] See: http://iconclass.let.ruu.nl/home.html

[248] See: http://www.gii.getty.edu/vocabulary/aat.html

[249] See: the author's "Towards a Global Vision of Meta-data: A Digital Reference Room,"*Proceedings of the 2nd International Conference. Cultural Heritage Networks Hypermedia*, Milan, pp. 1-8 (in press).

[250] This Aristotle subdivided into active Operation and passive Process.

[251] J. Perrault, "Categories and Relators: a New Schema," *Knowledge Organization*, Frankfurt, vol. 21 (4), 1994, pp. 189-198. Reprinted from: J. Perrault, *Towards a Theory for UDC*, London: Bingley, 1969.

[252] A fundamental problem in the systematic adoption and interoperability of these relations is that different communities and even different members within a community use alternative terms for the same relation. For instance, what some library professionals call "typonomy" is called "broader-narrower terms" by others, "generic" by philosophers, and in computing circles is variously called "is a", "type instantiation" and "generalization." Similarly, "hieronomy" is variously called "is part of," "partitive" by philosophers and "aggregation" by computer scientists. MMI will encourage research at the doctoral level to create a system for bridging these variant terms, using as a point of departure Dahlberg's classification of generic, partitive, oppositional and functional relations.

[253] Robert E. Kent, "Organizing Conceptual Knowledge Online: Metadata Interoperability and Faceted Classification," *Structures and Relations in Knowledge Organization*, ed. Widad Mustafa el Hadi, Jacques Maniez and Steven Pollitt, Würzburg: Ergon Verlag, 1998, pp. 388-395.

See: http://wave.eecs.wsu.edu

[254] See: http://www.philo9.force9.co.uk/books10.htm

[255] This may be closer than we think. Cf. David Brin, *The Transparent Society*, Reading, Mass.: Addison-Wesley, 1998, p. 287, who also reports on trends towards proclivities profiling, p.290.

[256] See: http://www.csl.sony.co.jp/person/chisato.html

[257] See: http://150.108.63.4/ec/organization/disinter/disinter.htm. For a contrary view See: Sarkar, Butler, and Steinfield's paper (*JCMC-electronic commerce*, Vol.1 No.3).

[258] See: http://www.cselt.it/ufv/leonardo/fipa/

cf. http://drogo.cselt.stet.it/fipa/

[259] See: http://umuai.informatik.uni-essen.de/field_of_UMUAI.html

[260] See: http://www.ina.fr/TV/TV.fr.html

[261] Bruce Damer, *Avatars!*, as in note 34 above.

[262] See: http://www.chez.com/jade/deuxmond.html which represents Paris.

[263] See: Virtual Helsinki at http://www.hel.fi/infocities/eng/index.html.

[264] See: http://idt.net/~jusric19/alphalinks.html

[265] See: http://socrates.cs.man.ac.uk/~ajw/

[266] At the Internet Society Summit (Geneva, July 1998), Vint Cerf, the new Chairman, in his keynote, described how the international space agency is working on a new address scheme to be launched with the next voyage to Mars late this year.

[267] Cf. John Darius, *Beyond Vision*, Oxford: Oxford University Press, 1984; *The Invisible World, Sights Too Fast, Too Slow, Too Far, Too Small for the Naked Eye to See: :* , ed. Alex Pomaranoff, London: Secker and Warburg, 1981.

[268] For a recent call to articulate a European approach to counter the above trends See: : : *Towards a Human Information Society. People Issues in the Implementation of the EU Framework V Programme*, ed. J. P. Chester, Loughborough: Quorn Selective Repro Limited, 1998. (European Union ACTS Programme. USINACTS Project).

[269] Dr. Theo Classen, as in note 18 above.

[270] The definition of usefulness could readily detour into a long debate. For the purposes of this article we shall take it in a very broad sense to mean the uses of computers in terms of their various applications.

[271] See: : *Information Week*, May 18 1998, p. 212.

[272] See: : M. Smith, "X.500 Attribute Type and Object Class to hold Uniform Resource Identifiers" at ftp://dns.internic.net/internet-drafts/draft-ietf-asid-x500-url-01.txt.

[273] See: : : Bunyip Information Systems
310 Ste-Catherine Street West, Suite 300
Montreal Quebec H2X 2A1
Tel. 514-875-8611
Fax. 514-875-8134
Archie
WHOIS ++ Protocol for Directory Service
Chris Weider
Compatible with X.500

[274] See: http://www.middlebury.edu/~its/Software/WebPh/README.html

[275] See: http://www.umich.edu/~dirsvcs/ldap/doc/guides/slapd/1.html#RTFToC1

[276] See: ftp://zenon.inria.fr/rodeo/solo/draft-huitema-solo-01.txt

[277] See: http://www.wcmc.org.uk/

[278] See: http://www.euroigi.org.

[279] See: http://www.ru/gisa/english/cssitr/format/s-57.htm

[280] See: http://www.ru/gisa/english/cssitr/format/sql_mm.htm

[281] See: http://www2.echo.lu/oii/en/gis.html#GeoTIFF

[282] See: http://www2.echo.lu/oii/en/gis.html#GIS

[283] See: http://www.env.gov.bc.ca/gdbc/saif/toc.htm

[284] See: http://www.env.gov.bc.ca/~smb/saif.html

[285] See: http://www.ifp.uni-stuttgart.de/ddgi/ddgi-main.html

[286] See: http://www.ru/gisa/english/cssitr/format/bycountr.htm

[287] See: http://www.ihs.on.ca/astm.htm

[288] See: file://waisvarsa.er.usgs.gov/wais/docs/ASTMmeta83194.txt

[289] See: http://sdts.er.usgs.gov/sdts/mcmcweb.er.usgs.gov/sdb

[290] See: http://www.gfdc.gov/fgdc2.html

[291] See: http://fgdc.er.usgs.gov/metaover.html

[292] See: http://geochange.er.usgs.gov/pub/tools/metadata/standard/metadata.html
cf. http://www.geo.ed.ac.uk/~anp/metaindex.htm

[293] See: http://www.fgdc.govt/Metadata/metahome.html

[294] See: http://www.att.com/attlabs/people/fellows/abate.html.