

Kim H. Veltman

Frontiers in Conceptual Navigation

Partially published in: *Knowledge Organization*, Würzburg, vol. 24, 1998, n. 4, pp. 225-245.

Republished in the Author's Book: *Frontiers in Conceptual Navigation for Cultural Heritage*, Toronto: Ontario Library Association, 1999.

Abstract

This paper was prepared for a lecture at a recent meeting of the German Chapter of ISKO devoted to *Wissensorganisation mit Multimedialen Techniken [Knowledge Organization with Multimedia Techniques]*, (Berlin, October 1997), the proceedings of which will contain an abridged version. Recent advances in technology assume a separation of content and presentation with respect to data structures. In terms of access, however, there are important reasons for relating content and presentation (different views, perspectives). The paper outlines some fundamental concepts underlying a prototype for a System for Universal Media Searching (SUMS), namely, learning filters, and knowledge contexts, levels of knowledge, questions as strategy: purpose as orientation; media choices, quality, quantity, questions, space using maps and projections; multi-temporal views and integrating tools. It foresees how such a system, linked with the equivalent of a digital reference room, will provide the basis for a System for Universal Multimedia Access (SUMMA). The latter part of the paper addresses recent developments in three-dimensional interfaces. It claims that these are particularly suited for certain tasks such as visualising connections in conceptual spaces; seeing invisible differences as well as comprehension and prediction by seeing absence. It suggests also some ways in which two- and three-dimensional interfaces can be used in complementary ways.

1. Introduction
 2. Access, Content and Presentation
 3. Learning Filters and Knowledge Contexts
 4. Levels of Knowledge
 5. Questions as Strategy: Purpose as Orientation
 6. Media Choices
 7. Quality, Quantity, Questions
 8. Maps, Projections and Space
 9. Multi-Temporal Views
 10. Integrating Tools
 11. Visualising Connections in Conceptual Spaces
 12. Seeing Invisible Differences
 13. Comprehension and Prediction by Seeing Absence
 14. Conclusions
-

1. Introduction

The digitisation of knowledge is provoking a wide range of responses. On the one hand, optimists paint scenarios of an information society with knowledge workers, and even electronic agents who will do our work for us by consulting digital libraries and museums, making learning an almost automatic process. They describe a world of seamless connectivity across different operating systems on computers and also among the communications devices in myriad shapes: from televisions, video-cassette recorders, CD-ROM players and radios at home, to faxes and photocopiers in the office as well as telephones, cellular phones, and other gadgets. The advent of nomadic computing¹ will make ubiquitous computing a reality. The result, they promise will be a world where we have access to anything, anywhere, anytime, where there are self-learning environments and the world is a much better place.

On the other hand, thinkers such as Pierre Levy² argue that the new computer age is bringing a second flood, whereby we risk being drowned in the massive amounts of information. In their view systematic approaches to knowledge are a thing of the past. Their pessimistic view is that there is simply too much knowledge to cope. If they were right, one could be driven to the conclusion that the high goals of the Information Society as articulated in the Bangemann Report, illustrated in the G7 exhibitions and pilot projects have only created new problems rather than long term solutions. This paper takes a more positive view. It acknowledges that the challenges are more formidable than some of the technophiles would have us believe; that these challenges cannot be solved by simple number crunching, but can be resolved with strategies that will lead to new insights in the short term and potentially to profound advances in understanding in the long term. The hype about anything, anytime, anywhere makes it sound as if the only advance is in terms of convenience. This paper claims that much more is possible and thus concludes on a note of restrained optimism.

By way of introduction there is a brief discussion of paradoxical links between access, content and presentation. Next, basic distinctions are made between different kinds of knowledge: ephemeral and long term; static and dynamic. It is shown that decisions whether information is stored locally or remotely affect how knowledge is handled. A series of strategies for access to knowledge are then outlined: the role of purpose, questions, spatial and temporal co-ordinates, multiple classifications, authority files. It is claimed that these strategies depend in the long term on a new kind of digital reference room. The final section of the paper turns to emerging methods in three- dimensional navigation and explores some potential applications in terms of visualising connections, seeing invisible differences and predicting by seeing absence.

2. Access, Content and Presentation

The differences between knowledge in books and in digital versions are more fundamental than they might at first appear. Books are linear, which means that the method of presentation is linked with the content in a fixed way. Any decision to change presentation requires re-publication of the book. In electronic versions these constraints

may apply as in the case of CD-ROMS or HTML pages, but they need not apply. At the level of programming, the last decades have brought an increasing awareness that it is useful to separate the transport of content from its presentation³. For example, at the high level, Standardized General Markup Language (SGML), separates fully the rules for delivery of a document from the ways in which it is presented and viewed by different users. XML, which is a simplified subset of SGML, attempts to make these high level principles accessible to a larger audience.⁴ This same philosophy pertains to databases. Once the contents have been organized into fields, then any number of subsets can be viewed and presented without having to re-write, or even re-organize the original content. Solutions such as SGML, XML and databases, mean that one can write once and view in many different ways with only a minimum of effort. At the level of programming they require a strict separation between access to content and presentation of content.⁵

One might expect that what applies at the database and programming level should also apply at the viewing level, namely that the best way to make progress in access to knowledge is to separate this entirely from questions of presentation. This assumes that one can employ raw computing power to searches and not have to worry how things look. According to this approach search engines are the domain of serious programmers, while presentation is quite a different thing, the domain of users and incidental at best. Hence search engines are about computational speed, theoretically require teams of professionals, whereas presentation can be relegated to a little section on user preferences in a browser, as if all that were involved were matters of taste such as the colour of screens or the image on a screensaver.

The problem goes deeper. Searching, it is assumed, requires one set of tools, an active search engine such as *Altavista*, *Yahoo* or *Hotbot*, which then presents its results in a passive browser such as *Netscape* or *Microsoft Explorer*. Presentation requires another set of tools such as *Powerpoint* at the everyday level or more advanced tools such as *Toolbook*, *Macromedia Director* or *Authorware*. It is assumed that they are two very different problems: passive viewing of materials that have been accessed through a search engine, or active editing of materials for presentation.⁶ Hence if we are interested in both problems, we need to download the findings from our browser, via a notepad, to a presentation tool.

Alas, these assumptions account for some of the important limitations of access methods for the moment. Different views of information are much more than a matter of taste. They are crucial to understanding materials in new ways and therefore an essential element in access strategies. To take a concrete example: I am searching for titles by *Alberti* on perspective in the Library of Congress. The Library has all its records in a complex database with many fields. A typical query presents materials from some of those fields in a static form whereby the dynamic properties of the database form are effectively lost to me as a user. Hence if the Boolean feature is working, I can search on *Alberti* and *perspective* and this will give me a random set of titles answering those criteria. The latest version of a browser such as *Yahoo* allows me to arrange this list alphabetically. The titles are listed in their original language such that the same text which has as its original title, *De pictura*, is found at different points of the list, under *B*

as in *Buch von der Malerei* in German, *D* as in *Della pittura* in Italian, *M* as in (O) *Malarstwie* or *O* as in *On Painting* in English. The MARC records have a field for the standard title, but this is typically not available in searches. Downloading a full MARC record would provide me with the basic ingredients for the answer I need. But what I actually want is to have the materials from the remote database in the Library of Congress or some other institution loaded into a database within my local browser such that I can have different views of this material on the fly without needing each time to make a new search from the remote source. If I wished, for instance, to look at these same titles chronologically I could do so on the spot. This functionality becomes essential if I am making a search which produces several hundred or even several thousand titles. Hence, what is needed is a search engine and browser, with editing functions to provide multiple views of materials. We have software for isolated functions. We need an integrating software which allows us to move seamlessly from searching and access to editing and presentation. Access and presentation are connected.

There are other cases when connections between access and presentation become even more vital. I am searching for a term. It results in a text containing words which I do not understand. I want to click on one of these terms, then click on a dictionary function which takes me to a copy of *Websters* or the *Oxford English Dictionary* and provides me with an online definition of the term in question. At present we have to stop our present search, look for a dictionary, search for the term in a dictionary and then go back to the text we were reading. To expand the search it would be useful to know its synonyms and antonyms. In this case we want to click on the word, choose one of these alternatives and go on line to Roget's *Thesaurus* to find the answers. Without a close coupling of access and presentation this contextualizing of searches is difficult if not impossible. Thus, while the programming level requires a clear separation of access strategies from presentation methods, the user level requires the re-integration of these two domains in order to achieve more effective conceptual navigation.

The above examples entail an approach to searching which allows a user to have access simultaneously to a series of basic reference tools that are the electronic equivalent of a scholar writing at a desk with standard dictionaries, encyclopaedias and other reference works within arms' reach. This is one of the underlying assumptions in the prototypes for a System for Universal Media Searching (SUMS, ©1992-1997). A senior scholar working in a major library would have access to a large reference room such as the Reading Room in the British Library. An electronic improvement would be a digital reference room, which offered the cumulative resources of all the major reference rooms. This would give every user desktop access to the combined materials available in the British Library, the Bibliothèque de la France, the Vatican and other great collections. Such a tool is foreseen in the System for Universal Multi-Media Access (SUMMA, ©1996-1997, cf. Appendix 1).

Implicit in the above is a new methodology to search strategems. An old model assumed that everything could be collected in a single place. This has been a dream since the Library of Alexandria. At the other side of the spectrum, a more recent Internet model

assumes a fully distributed model. The proposed method offers a middle way between these two extremes in that it foresees a centralized digital reference room⁷ (which could of course have a number of mirror sites) serving as an electronic repository for global meta-data and pointing to distributed repositories of contents around the world.⁸

One fundamental idea underlying this new search strategy is very simple. Libraries, in particular their reference sections, have constructed enormous lists of materials, notably their (now electronic) catalogues of author names, subject headings, keywords, titles, place names, call numbers and the like. These catalogues are presentation methods to gain better access to their own collections of books, and other materials. At the same time, these catalogues are effectively authority files with standard and variant versions of authors' names, place names etc. and as such can be used as instruments for navigating through materials elsewhere. As a library catalogue, each name, subject, title and place, may point to only one specific object in that particular library. As a list of names, the same catalogue can help me refine my search for a J. Taylor in general to John Taylor of Cambridge in particular. As a list of subjects the same catalogue can help me refine my search from medicine in general to osteoporosis, cardio-vascular research or other particular branches the names of which I might have forgotten or never have known until I looked at the systematic branches of the field. Alternatively, as a list of titles, each single title has a discrete call number and so I can go to that call number to discover other titles classed in the same areas. Or browsing near that number, I discover related problems and fields. Thus the presentation tools for a given library collection can become search and access tools for other collections, and even help in refining questions about other materials which may not have organised as carefully as the library materials. Thus past efforts at organising knowledge can help us in present efforts and future searches.⁹

To take a specific and unexpected example of this idea: *Roget's Thesaurus of English Words and Phrases* is typically used as a standard reference work for finding related words. On closer inspection we find that this was much more than a handy list of words. Roget set out to divide the whole field of meaning into 1000 headings (now reduced in Longman's edition to 990), subsumed under six basic classes (figure 1). A few years ago (1992), a scholar named Day,¹⁰ used this classification system as a finding tool for Biblical verses, pointing out that this is often much more effective than either topical indexes (which are frequently redundant because there is no critical heirarchy of terms) or standard concordances based on actual words (which are usually linked to a specific edition of the Bible and thus less universal in their application). This approach, which could potentially be applied to any book, again suggests a basic principle: classes used to order the world, can be used to find materials in the world: presentation methods are keys for access methods. The WWW virtual library has begun to exploit this principle, limiting themselves to the upper levels of the classes and using them only to find web sites rather than as an integrated means for finding web sites, library titles and materials.

1	Class
1	Abstract Relations
2	Space
3	Matter

4	Intellect
5	Volition
6	Emotion

Figure 1. Six basic classes used in *Roget's Thesaurus*, under which he organized all the fields of meaning into 1000 headings. These have since been used as categories for finding Biblical passages and could potentially be applied to any book.

In one mode these searches will be manual. In another they will be aided by voice technology such that one is taken directly to the appropriate points in a list. In other modes, agent technologies will record one's search habits in order to gain a profile of the user's interests. These agents would then use this profile to search other sources. In addition, they could take a list of the user's subject interests, determine closely related subjects, use this to explore potentially relevant material and suggest these as being of possible interest to the user.¹¹ Parallel to these activities of agents at a personal level will be their role in systematically making a series of authority lists of names, places, events etc.¹² Eventually, access strategies will include a whole range of choices including: agents, filters, languages, levels of education, machine configurations, personal profiles, relations, special needs, structures and viewers.

3. Learning Filters and Knowledge Contexts

In any given field the complete corpus of knowledge is enormous. This corpus, which is the sum total of materials available in the world's libraries and research institutes is seldom understood by more than a handful of experts around the world. These experts and their colleagues use subsets of this corpus to write the curricula and subsets thereof for standard textbooks in their fields which then become the basis of courses. The courses at various levels of education from elementary school to the post-graduate level are subsets of these textbooks. Exams, in turn, are further subsets of the courses. Traditionally, students are expected to recognise the links between exam, text and course; teachers recognise further links with the curriculum and only a handful of experts can recognise all the precise links between a particular exam question, course, textbook, curriculum, and the corpus. Given the advent of computers these mental links can be translated into electronic hot links, such that even a beginning student will be able to trace the whole range of links between an exam (question) and the corpus (answer on which it is based), and thereby understand the full context. Conversely, one could begin with any part of the corpus and trace the (exam) questions it entails.

Digital libraries are about making available in electronic form the corpus of knowledge. Ministries of education and are translating curricula and learning outcomes into digital form. Faculties of education and institutions of learning as a whole are making individual courses available in electronic form. Once this process is complete, the links between these efforts can be made and there will be a whole range of subsets of every corpus corresponding in each case to a different level of education. A pre-school child will begin with the smallest of these subsets. Their list of persons (Who?) will be very short. Once they have mastered this list, they can go to the next level which has a few more names,

and they can continue in this fashion until they have reached the full list of all names at the research level.¹³ The same principle applies to subjects, key words (What?), places (Where?), events, times, chronologies (When?), methods, procedures, processes (How?) and explanations (Why?). The genius of the system lies therein that it does not have to create all this content. It uses the presentation materials of existing knowledge organisations, particularly libraries, museums and schools as the starting point for new access strategies. The novelty lies in making new use of the old through integration rather than trying to re-invent the wheel as so many search engines assume to be the case. Today's search engines aim to find a title we want. Research is about finding what we did not know existed. We need to develop re-search engines.

4. Levels of Knowledge

From a global point of view reference rooms in libraries contain mainly five kinds of materials, namely, 1) terms (classification systems, subject headings, indexes to catalogues); 2) definitions (dictionaries, etymologies); 3) explanations (encyclopaedias); 4) titles (library catalogues, book catalogues, bibliographies); 5) partial contents (abstracts, reviews, citation indexes). All of these are pointers to the books in the rest of the library, or 6) full contents which can conveniently be divided into another four classes, 7) internal analyses (when the work is being studied in its own right); 8) external analyses (when it is being compared or contrasted with other works); 9) restorations (when the work has been altered and thus has built into it the interpretations of the restorer) and 10) reconstructions (when the degree of interpretation is accordingly larger).

From this global point of view the first six of these categories are objective,¹⁴ while the last four (6-10) are increasingly subjective. The first category is also the most simple (isolated terms), and each successive level enters into greater detail: i.e. dictionary definitions range from a phrase or a few sentences; encyclopaedia explanations extend from a few sentences to a few pages etc. Once again the physical arrangement of major libraries serves as a starting point for the conceptual system; the presentation system of libraries offers another key to access into its electronic version. The heritage of experience in organising the known provinces of knowledge, offers the departure points into its unknown lands.

At the same time the electronic version offers more possibilities than its physical counterpart. For instance, the classification system of a library changes over time. In print form this is documented by successive editions of the Dewey, Library of Congress and other systems. These print forms are static and it would require my having the various editions in front of me all opened at the same section in order to trace their evolution. In electronic form the various editions can be linked to a time scale such that the categories change dynamically as I slide the time scale. This offers new means of understanding the history of a field. In the longer term one will be able to go from a term in any given classification scale to the same term or its nearest equivalent in other classification schemes. If one entered all the major systems traced by Samurin¹⁵ in his standard history of classification systems, shifts along the time scale would allow one to see the gradual

branching of knowledge over time and trace how a general category such as medicine has led to thousands of specialty topics.

5. Questions as Strategy: Purpose as Orientation

There is the popular game of twenty questions. In simple cases we can learn a great deal with only six questions. The first is the purpose, or Why? Secondly, we ask about the means, or How? Thirdly, we determine the time of its occurrence or When? Fourthly we ask whether it is local, regional, national or international or Where? Fifthly, we ask about the precise subject, or What? Sixthly, we ask about the persons involved or Who? This sequence of questions is part of the method. There is the well known episode in *Alice in Wonderland* where she asks the cat for directions only to be told that this depends on where she wants to go. To navigate effectively the totality has to be reduced to navigable subsections. To this end questions offer a strategy.¹⁶

Knowing the purpose in terms of basic categories such as everyday, business, health, law, religion and leisure determines the main thrust of the information and knowledge we are seeking. Everyday includes classifieds, news, sports, traffic and weather. If our goal is business oriented we will be concerned with a very different subset of knowledge about a city and a person visiting it for leisure as a tourist. The purpose (Why?) thus provides a first way of determining the scope of the search and thus narrows the field. Knowing the means by which the goal is to be accomplished (How?) further narrows this scope. For instance, if our interest in everyday news is limited to television, then we can ignore radio and newspapers. Knowing the temporal boundaries provides further limits (the old *terminus ante quem* and *post quem* or simply When?) Knowing the geographical boundaries of our interests (Where?) further limits the scope of the search. If we are interested in leisure and tourism specifically in Italy or India, then we can ignore the leisure information for all the rest of the world. Knowing the precise subject (What?) and/or the precise persons (Who?) provide final refinements to this process of narrowing the scope.

This means that, in the case of basic searches, a simple series of six simple choices can guide a user from a vague intention to a quite specifically defined question. This slightly tedious, highly structured procedure is appropriate for beginners and some members of the general public who wish to use their search engines in the manner that they use their remotes. Accordingly the number of choices are sufficiently limited such that they could fit onto their remote: they can do question hopping as easily as channel hopping. A slight advance introduces longer lists of choices elsewhere on the screen to increase the number of alternatives.

The Internet has introduced the notion of Frequently Asked Questions (FAQ). These are typically listed without any particular order. The questions methodology foresees that these are organized by question type. Hence, having chosen a topic, by pressing Who? the user would receive all FAQ's concerning persons, by pressing When? they would have all FAQ's concerning temporal attributes. By pressing How? they would have all

FAQ's concerning function. These lists could in turn be viewed from multiple viewpoints: alphabetically, temporally, etc.

A next stage in sophistication introduces the questions as headers each which has a box below it into which users can define more precisely the details of their search. For instance, under *Who*, they might write, *Leonardo da Vinci*; under *What* they might write, *optics*; under *Where* they might write *Milan*, under *When* they might write *1505-1508* and under *why* they might write *Education*. They are not constrained to enter information under every question, except that every omission results in a more general answer. If they leave out *When* they would get all his optical works written in Milan. If they left out *Where* they would get all his optical works and if they left out *What* they would get his complete works. The elegance of this approach is that a simple word typed in by the user is in each case automatically translated into a question linked with appropriate fields of a database. The user's one word statements on the surface are translated effortlessly to formal programming queries below the surface.

Further stages of sophistication transform these random entries of general subjects to successively more comprehensive lists based on a) the user's personal list of subjects; b) headings in database fields; c) the subject lists of the Library of Congress; d) the Library of Congress classification list; e) multiple classification lists. Once again the order established in libraries and other institutions (partly for presentation purposes) helps to refine the subtlety of the access strategies and render elegant the navigation principles.

6. Media Choices

Librarians have long been concerned with careful distinctions about the medium of the message, and they have typically separated objects accordingly, books go to one place on the shelves, maps to another, prints to another. In great collections, such as the British Library, maps and prints have their own rooms. Similarly music recordings go to a special section. Very different media have gone to fully separate institutions including film libraries, television archives, art galleries and museums. Each of these have developed their own specialised techniques for cataloguing and recording. So having used the experience of libraries to create cumulative authority lists of names, subjects and places for books and related media, the records of these other institutions can be used to arrive at a more comprehensive picture of known materials. One important aspect of recent meta-data trends is the identification of media types as an essential aspect of their description. Thus a simple list of media choices serves as yet another tool in refining one's potentially infinite search of all things, to a manageable subset covering some thing(s) in some particular media.¹⁷

The net result of such strategies is equally interesting because it means that hitherto dispersed information about different media can now be accessed at once rather than have to rush from a book catalogue to a print catalogue, map catalogue etc. Hence one will be able to see at a glance how one subject inspired many media whereas others have been limited to a few of even an isolated medium. McLuhan focussed on the medium as the

message. But there are clearly histories to be told of messages which inspired a whole range of media.

7. Quality, Quantity, Questions

Until recently many persons measured search engines by their ability to find things. Today the problem with most search engines is that they find too much. The problem has more to do with quality than quantity, which depends not so much on the power of the engine as the precision with which the engine is given direction. If we ask for everything we get everything. Even in the case of something fairly specific such as Smith, everything about all the Smiths in the world is still a great deal more than we usually want to handle. Most of us have no idea that there is a great deal more available than we suspect, so when we begin searches on a global scale we need a crash course in specificity.

Bits of this are common sense. If we are searching about the latest things happening in Sydney, Australia it is wise to look first in databases in Sydney rather than searching through every site on the other continents. If we do not find it there, we can do a global search and perhaps discover that an Australian expatriate now living in Los Angeles or London, has a database on precisely what we need. If we are looking specifically for financial materials relating to banks it makes no sense to search databases of legal firms, health centres or tourist spots. As the IBM advertisement states: it is better to fish where the fish are. To this end, our course, Specificity 101, would state, as was noted above, that in the case of simple searches, a strategic sequence of Why?, How?, When?, Where?, What? and Who?, is often enough to narrow the scope of our general interest into something sufficiently specific to result in a manageable number of answers.

If the number of answers is still too large then we need further means to make them smaller. In the physical world when the number of applicants to a university is too high, standards are raised, which means that only those of a certain quality are chosen. The same principle applies in the electronic world. Until recently there was a problem that in many electronic documents the precise level of quality was undefined. A number of initiatives are underway to alleviate this, including a) new conventions for high level domain names so that one can tell at a glance whether the site is basically concerned with education, business, or some other subject; b) the W3 Consortium's Protocol for Internet Content Selection (PICS) and c) their initiatives in the direction of digital signatures. The latter of these is perhaps the most far reaching. It places the responsibility of description on the content producer, although their accuracy can then be checked by others, such that a digital signature effectively functions as a submission to peer review. Quality articles can only gain from the fact that their value has been confirmed by others, so those who refuse to provide digital signatures will effectively be disqualifying themselves.

As in the world of books, more substantial electronic projects and websites have reviews and contexts which further identify quality and remarkable achievement. Thus lists of basic qualities become additional parameters for paring down the many choices to a

manageable small list. The same is true for lists of quantitative features. These too act as tools for refining the parameters of a search.

In some cases an iterative use of the basic questions will be useful. For instance, if my primary interest is biographical then my original question would have been in terms of Who? I might have started with a generic term for Who? such as *Artists* and then narrowed this universal list to subsets of *Italian* then *Renaissance*, then *Florentine* in order to arrive at *Sandro Botticelli*. To know more about his colleagues I would again ask Who? To know more about the subjects connected with Botticelli I would again ask What? And so on. Hence, as we know from cases with celebrity interviewers, investigators and psychologists, a careful sequencing of a small number of precise questions often brings better answers than many imprecise inquisitive stabs in the dark.

8. Maps, Projections and Space

Geography is an important tool for navigation both in its literal and metaphorical applications. In its literal uses, points on a map are linked with information about those points. As Tufte¹⁸ has shown, this has a long tradition in cultures such as Japan. In the West, these techniques have evolved rapidly through the advent of Geographical Information Systems (GIS), which are being linked with Global Positioning Systems (GPS) at the world level and Area Management/Facilities Management (AM/FM) at the local level such that one can zoom from a view in space to a detail on earth, which may be static or dynamic.¹⁹ Among the more interesting projects in this context are T(erra)-Vision of Art+Com²⁰ and Geospace at MIT²¹.

As one descends from a view in space one can begin in two dimensions and change to three dimensions as the scale increases.²² The detail on earth may be a building, a room, a painting, a museum object or some other item. Each item takes one via its author, artists (Who?), subjects, title, theme (What?), place (Where?) and time (When?) to lists of the author's works and related authors' works, copies, works done on the same and related themes, other works produced in the same place and other works produced at the same time. Once again the presentation scheme of the gallery or museum serves as a starting point for an access strategy and navigation method.

In the above example one is going from a geographical mode to the textual mode of database materials. The converse is equally interesting. Any place name in a list one encounters is recognised as a place name and its position is located and presented on a map. This is another aspect of the digital reference room. I am reading a text and come across the name Uzuncaburc and may not remember exactly where in Turkey this city is. I highlight the name and am taken West of Tarsus, and just North of Silifke to a spot with an impressive Greek temple on a hillside. Calling on another source in the digital reference room, I can get a full description of the town from the equivalent of a *Guide Michelin* or *Baedeker* for Turkey.

Temporal Maps, Boundaries and Buildings

A map is a record of conditions and boundaries at a given time. These boundaries change with time and are therefore dynamic. Viewed dynamically with a temporal filter the boundaries of the Roman empire expand and then contract. The provinces of Italy change. The city limits of Rome increase, recede and then expand anew. In some cases these boundaries are a matter of controversy. Tibet's view of its boundaries with China may be very different from China's view of Tibet's boundaries. Hence, for any given place there are as many versions of the maps of boundaries as there are competing interpretations. This principle can be applied along the whole gamut of scales from global views of a country's boundaries to views of cities, complexes, and even individual monuments and buildings. The reconstruction of such sites, particularly ancient ones is rapidly becoming an industry in itself.²³

Individuals

Such temporal maps can also serve to trace the movements of an individual artist or scientist from their birthplace, to their places of study and work, including their travels until the time of their death. In the case of a famous artist such as Leonardo, such temporal maps can trace the history of ownership of a given painting from the place where it was originally painted through the various collections in which it was found. There may be debates about the movements of the artists or the history of one of their paintings. Hence each interpretation becomes an alternative map. Standard interpretations are indicated as such. The reliability of interpretations is dealt with under the heading of quality (see section 7 above).

Concepts

Such maps can equally be applied to concepts such as Romanesque churches or Gothic architecture, although in this case there are two temporal dimensions to consider. First, the number of buildings included in the corpus varies from author to author and will tend to become more detailed as we approach the present. Secondly, these authors will have different theories concerning the historical development of a given church. For instance, author A may believe that the sequence of Gothic churches was St. Denis, Chartres, Naumburg, whereas author B may claim that there was a different sequence in the history of the buildings.²⁴ These alternatives are again available under lists by author, year, place etc.

Virtual Spaces

In Antiquity virtual spaces were used as a tool in the art of memory²⁵. One imagined a house with different rooms, associated different things with each room and then returned in one's mind's eye to the house and a given room when one wished to remember a given fact. The imaginary worlds of Virtual Reality Modelling Language (VRML) are effectively translating this tradition of mental spaces into visual spaces²⁶. A fundamental difference is that whereas the mental space was private and limited to only one person the visual space is potentially public and can be shared by any number of persons. There is much rhetoric about the importance of collaborative environments. The whole question

of when such an environment is more effective and when it is not requires much further study.

9. Multi-Temporal Views

In traditional knowledge environments we typically assume a single method of time reckoning. A European expects to find a Gregorian calendar in everyday life. Exceptions to this rule are Jewish, Muslim, Hindu and persons of various religions who follow other calendars for religious purposes. Temporal navigation will allow for conversion of different calendars and chronologies similar to the conversion tools which already exist for major areas of physics such as measurement, power, force, light and heat. Hence, if I am studying a Christian event in 1066 or during the Crusades, I might wish to see what the equivalent date was in the Muslim calendar and more significantly what were the main events at the time from the viewpoint of their religion or culture.

As in other cases above, lists of the categories of time, hours of the day, days of the week, months of the year, historical periods, geological periods, all the categories which have been developed for presentation of our knowledge of temporal events can be used as means for gaining access to temporal materials.

10. Integrating Tools

As noted earlier the software industry began by creating a host of tools for writing, drawing, graphing, designing and editing, each of which appeared in separate shrink wrapped packages. Often they were incompatible with one another. More recently a number of these have been bundled into office suites such that they can be used together. A problem remains that each of these typically retains its own interface such that every operation requires a change of style. Within individual products such as *Microsoft Word* we find a function called *Tools*, which includes a series of functions including spelling and grammar, language, word count, auto summarize, autocorrect and look up reference (which could be seen as a first step in the direction of a digital reference room). The challenge is to extend this notion of tools within *Word* to include the whole spectrum of tools used in all the other software packages: to have multiple operations within a single presentation framework; a common look and feel for one's search, and access, creating, editing, and presentation tools.

Such a set of tools is listed by general categories: mathematical, scientific, simulation, verbal, and visual; each of which breaks down into further functions. For instance, verbal tools include: class, create, discuss, edit, input, output, search and translate. Input tools include e-mail, fax, media, scan, telephone, xerox. This is convergence at another level. Whether all these tools are on one's local hard drive or on a remote server through a thin client will not alter this basic principle.

11. Visualising Connections in Conceptual Spaces

Library cards present their information on a two dimensional surface and electronic catalogues initially followed this format. This was partly a reflection of the limitations of computing power which are rapidly being superceded.²⁷ Accordingly a number of individuals are assuming that two-dimensional navigation will soon be discarded as a thing of the past and we shall use exclusively three-dimensional spaces for all navigation. For example, Dr. Henry Lieberman (MIT) is exploring the use of very large navigation spaces, with new techniques which allow “zooming and panning in multiple translucent layers.”²⁸ Carnegie Mellon University has a special Visualization and Intelligent Interfaces Group.²⁹ Silicon Graphics Inc. (SGI) foresees the use of landscapes.³⁰

Dr. Stuart Card (Xerox PARC) and his team have been working on a series of tools for visualizing retrieved information using techniques such as a galaxy representation, spiral calendar, perspective wall, document lens and cone tree.³¹ There is an analogous project at the Gesellschaft für Mathematik und Datenverarbeitung (GMD) in Darmstadt called Lyberworld.³² This takes a concept, searches for related terms, links these with the concept in question and presents them spatially in a cone. Alternatively the concept in question is positioned in the centre while various related terms are placed along the circumference of a circle where they exercise the equivalent of a centrifugal gravitational force. If all these surrounding terms are equal in strength they exercise an equal force on the central concept. As one of the terms becomes more significant it exercises a greater force on the central concept. Another GMD project, SEPIA, foresees a hypermedia authoring environment with four concurrent spaces: a content space, planning space, argumentation space and rhetoric space.³³

Historically the advent of three dimensional perspective did not lead artists to abandon entirely two-dimensional (re-)presentations. There were many cases such as cityscapes where three dimensions were very useful; others where two-dimensional solutions remained a viable and even preferable alternative. Text is an excellent example, which helps explain why text-based advertisements remain predominantly two-dimensional. If we are searching for a simple name (Who?), subject (What?), place (Where?), event (When?), process (How?) or explanation (Why?), two-dimensional lists are likely to remain the most effective means for searching and access. As suggested earlier, long lists benefit from alternative presentation modes such that they can be viewed alphabetically (Who?), hierarchically in tree form (What?), geographically (Where?), and chronologically (When?) if appropriate.

Three-dimensional navigation spaces are particularly valuable for contextualising knowledge. A two-dimensional title or frontispiece of a book tells us nothing about its size. A three-dimensional scale rendering helps us to recognize at once whether it is a pocket sized octavo or a folio sized book: a slender pamphlet or a hefty tome³⁴. Hence, having chosen a title one will go to a visual image (reconstruction) of the book; see, via the map function, where it appears on the shelf of a library, do virtual browsing of the titles in its vicinity or wander elsewhere in the virtual stacks.

Libraries are relatively simple structures. In the case of more complex systems such as the London Underground it is useful to move progressively from a two-dimensional

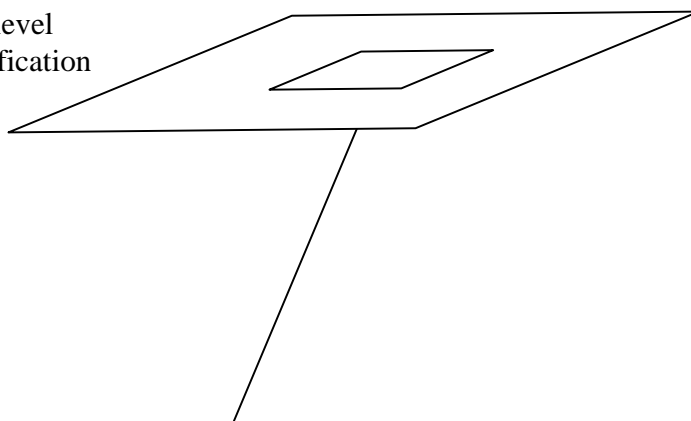
schematic simplification of the routes to a realistic three-dimensional rendering of the complete system, station by station. In the context of telecommunications the so-called physical world becomes one of seven layers in the model of the International Standards Organisation (ISO). In such cases it is useful not only to treat each of the seven layers separately but also introduce visual layers to distinguish the granularity of different views. In looking at the physical network, for example, we might begin with a global view showing only the main nodes for ATM switches. (Preliminary models for visualising the MBone already exist³⁵). A next layer might show lesser switches and so on such that we can move up and down a hierarchy of detail, sometimes zooming in to see the configuration of an individual PC, at other times looking only at the major station points. This is actually only an extension of the spectrum linking Area Management/Facilities Management (AM/FM) with Geographical Information Systems (GIS) mentioned earlier.

The third dimension has many uses beyond producing such electronic copies of the physical world. Pioneers of virtual reality such as Tom Furness III³⁶, when they were designing virtual cockpits realised that pilots were getting too much information as they flew at more than twice the speed of sound: the challenge was to decrease the amount of information, to abstract from the myriad details of everyday vision in order to recognise key elements of the air- and land-landscape such as enemy planes and tanks.

These principles are equally important in knowledge organisation and navigation. A library catalogue gives me the works of an author. Each catalogue entry tells me under how many fields a given article or book is classed. Adding these fields together leads to an alphabetical list of that author's intellectual activities. In electronic form producing such a list is theoretically simple. What I need however is a conceptual map³⁷. To what extent did an author work as a generalist in large subject fields and to what extent as a specialist? This lends itself to three dimensions. Author A is in one plane and the subject headings of their works are on other planes. These are aligned to relative positions in classification systems such that one can see at a glance to what extent this person was a generalist or a specialist (figure 2). This principle can be extended in comparing the activities of two authors.

This approach can in turn be generalised for purposes of understanding better the contributions of a group, a learned society or even a whole culture. Scholars such as Maarten Ultee have been working at reconstructing the intellectual networks of sixteenth and seventeenth century scholars based on their correspondence. A contemporary version of this approach would include a series of networks: correspondence, telephone and e-

General level
of Classification



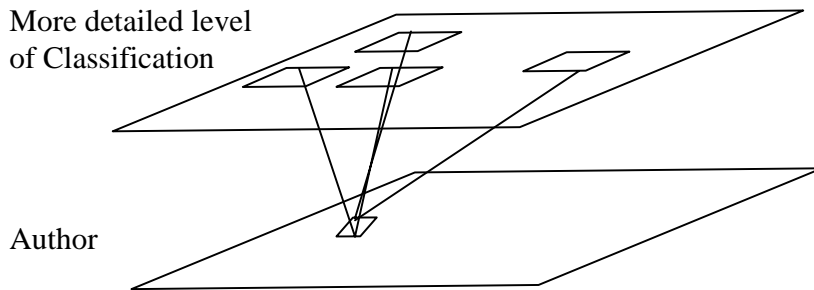


Figure 2. Visualisation of an author's activities whose specialist activities touch on four fields (three of which are closely related) and whose more generalist activities are limited to one major field. Further layers could be added to show how the same concepts recur in different places in various classification systems.

e-mail which would help us in visualising the complexities of remarkable individuals be it in the world of the mind, politics or business.

The geographical aspects of these intellectual networks could be visualised using maps. Conceptually the subjects of the letters, (and the e-mails to the extent that they are kept), could be classed according to the layers outlined above such that one gained a sense of the areas on which they focussed. For instance, what branches of science were favoured by members of the Royal Society? Did these change over time? It is a truism that Renaissance artists were also engineers and scientists. What particular fields did they favour? Can one perceive significant differences between artist-engineers in Siena, Florence, Rome and Venice? We could take the members of a learned society or some other group and trace how many layers in the classification system their work entailed and then study how this changed over time. Are the trends towards specialisation in medicine closely parallel to those in science or are there different patterns of development?

Alternatively by focussing on a given plane of specialization we could trace which authors contributed to this plane, study what they had in common in order to understand better which individuals, networks of friends and which groups played fundamental roles in the opening of new fields. Such trends can in turn be linked with other factors such as research funding or lack thereof. In addition to universities, major companies now have enormous research labs. Nortel, for instance, has over 17,000 researchers. Hitachi has over 30,000. We need maps of Who? is doing What? and Where? In our century we could also trace where the Nobel and other prizes have gone both physically and conceptually. Navigation provides virtual equivalents of journeys in the physical world. It is also a means of seeing new patterns in the conceptual world through systematic abstraction from everyday details in order to perceive new trends.

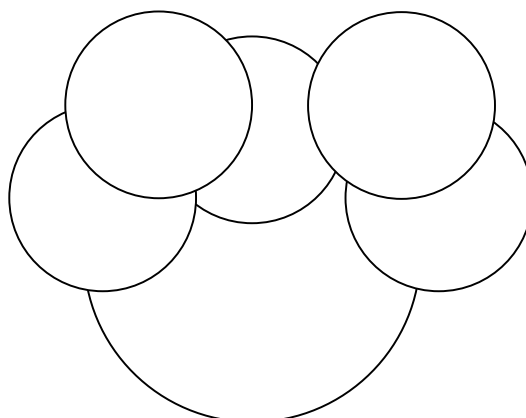
Contextualising also entails seeing relations between one subject and another. When we are studying a subject, we typically want to know about related subjects. In the past we went to a library catalogue, found a title and saw the related topics at the bottom of the

card. Electronic versions thereof exist. Recent software such as Apple's *Hotsauce* allows us to go from a traditional two-dimensional list of terms, choose one, and then see all the related topics arranged around it. These related subjects evolve with time, so with the help of a simple time scale we can watch evolution of a field's connections with other subjects. This idea can easily be extended if we translate the main topic into a circle and the related subjects into other (usually smaller) circles intersecting the main one to create a series of ven diagrams. This visualisation allows us to choose subsets common to one or more related fields, which is important if one is trying to understand connections between fields (figure 3).

We can go further still by treating the image on the screen as a plane, make this transparent, rotate it downwards by 90 degrees such that it becomes the top surface of a transparent box. The x axis now becomes the time axis such that we can literally trace the cone-like growth of various subjects and see when they first intersected. Such an example outlines a means of moving seamlessly from two-dimensional lists to three-dimensional conceptual maps of subjects with their related topics and also offers a whole new way of seeing interdisciplinarity.

If we were trying to trace the evolution of a new field, we could begin by using a dynamic view of classification systems described above. We could also use combinations of these intersecting ven diagrams. For example, the last generation has seen the emergence of a new field of bio-technology. This has grown out of two traditional fields, biology and technology. These could be represented as large circles surrounded by smaller ones representing, in this case, their related branches and specialities. Any academic work would be represented in the form of a line which thickens in proportion as the connections increase. These connections are of differing kinds. Initially they tend to be in the form of sporadic articles, conferences, or isolated research projects which have no necessary continuity.

Later there are books, journals, professorships, research institutes and spin-off companies which bring a conscious cumulative growth to the new field. Each of these phases could be identified with different colours so arranged that one can distinguish clearly between sporadic and cumulative activities (figure 4).



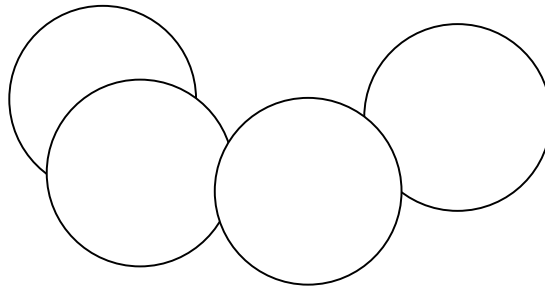


Figure 3. Ven diagram of a subject and its related subjects shown as intersecting circles. In addition to regular searches by subject, this visualisation allows a user to choose subsets common to one or more related fields, which is important if one is trying to understand interdisciplinary relationships.

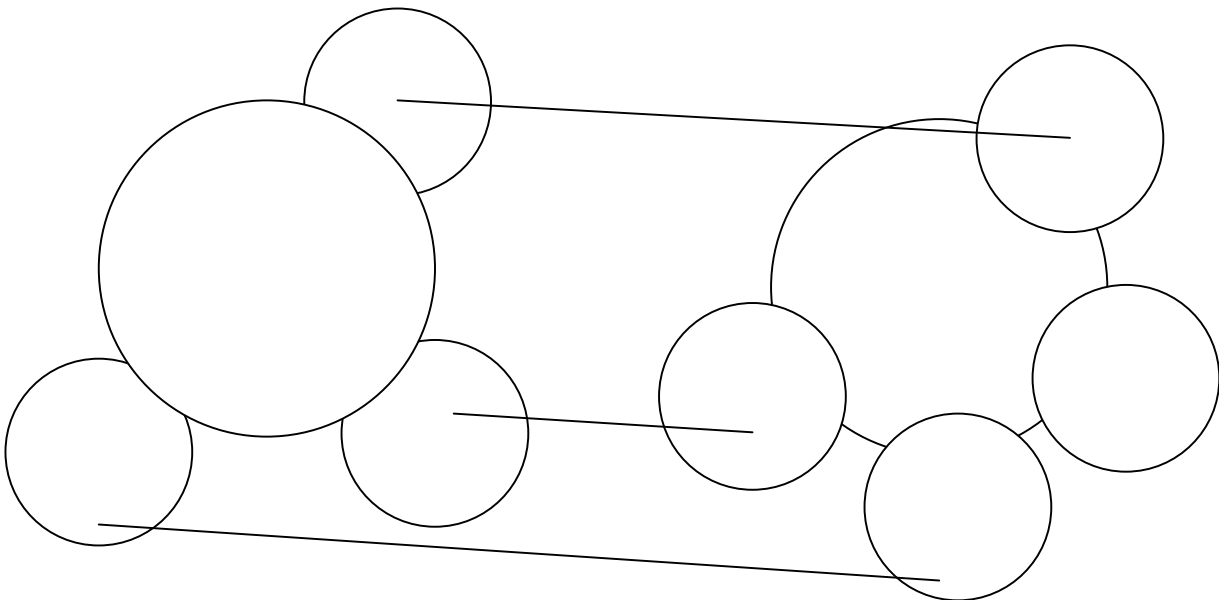
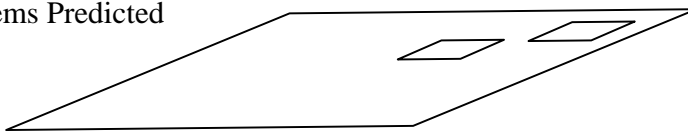


Figure 4. In this diagram the large circles again represent two fields and the smaller circles represent branches of these fields. The lines joining them represent work linking hitherto different branches. These lines thicken as the amount of articles and other research activities increase and thus become a new means of tracing the growth of an emerging field.

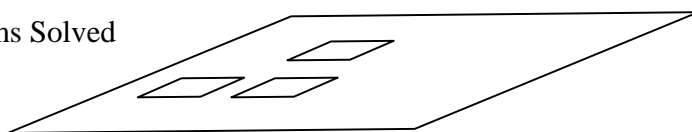
Plane 4

Problems Predicted



Plane 3

Problems Solved



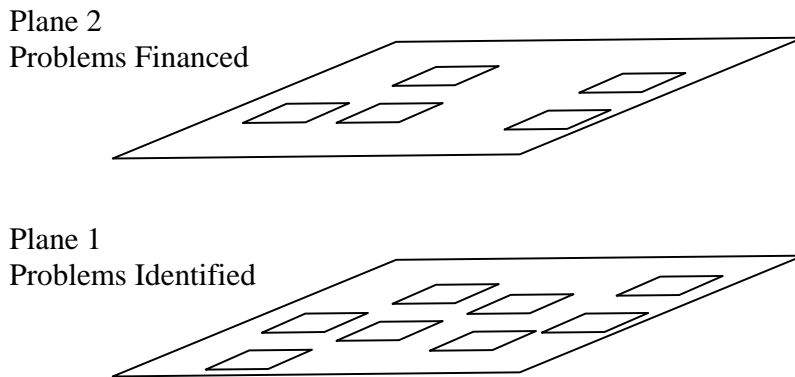


Figure 5. Using spatial arrangements of concepts to map problems identified and to visualise which subsets thereof were financed as research projects, which were solved in the sense of producing patents, inventions and products and which led to new predictions in the form of hypotheses and projections.

We can integrate these circles within the context of frames as described above. For example, the two fields of biology and technology could be on one plane. Major universities could be on a second plane. We could then trace which universities are producing the most papers of these two fields and specifically on what sub-sections thereof. On another plane we could list the major research institutes in order to determine other trends. Are these areas being studied more in the United States than Europe or Japan? If so what are the percentages? Which companies dominate the field? What links are there between research institutes and universities? Are these increasing or decreasing?

This notion of planes can be extended to see further patterns. Plane one can list all the known problems or potential research areas in a given area of science. Plane two lists which subset of these problems is presently being studied. Plane three shows which problems have been solved, or rather have some solutions in the form of inventions, patents, trademarks and products. Plane four lists a further subset for which solutions are predicted or which have hypotheses for their solution (figure 5).

Such comparative views will help scientists and decision-makers alike to understand more clearly trends in rapidly developing fields. Such matrices of problems can in turn be submitted to problem structuring methodologies whereby technical, practical and emancipatory dimensions are submitted to frameworks in order to discern where they fit into a Methodology Location Matrix.³⁸

12. Seeing Invisible Differences

During the Renaissance the discovery of linear perspective brought new skill in visualising the physical world, but it began by illustrating episodes from the lives of saints which none of the artists had witnessed personally. Hence it helped simultaneously in expanding the horizons of the visible world of nature and the invisible world of the mind. This dual development continues to our day. Three-dimensional visualisations, especially using virtual reality are helping to illustrate both the visible and invisible, and introduce many new possibilities.

If, for instance we take the Library of Congress classification, as above, and link each layer in its hierarchy with a different layer, then we arrive at a truncated pyramidal shape beginning with twenty initial topics at the top and increasing to many thousands as we descend. Say we are interested in total publications in the Library of Congress. At the top level, these publications can be linked to each of the twenty basic fields, such that each major subject is represented proportionately as a square or circle. We can thus see at a glance to what extent the number of books on science is greater than those in fine arts. By going down a level in the hierarchy we can see how those figures break down, e.g. to what extent there are more books on physics than chemistry or conversely. At another level we can see whether and if so to what extent astro-physics has more publications than bio-physics or quantum physics. We are thus able to see patterns in knowledge which we could not see simply by looking at the shelves, although they could give us some hint that one topic has more books than another.

A slightly more refined version would link this approach to book catalogues such that we can trace how these trends in publications change over time. From a global point of view we could witness the rise of the social sciences in the nineteenth century. At a greater level of detail we could see the rise of psychology as a field. This same approach can also be applied to usage patterns as studied by scholars in reception theory. In future usage patterns by on line readers will become important.

In our quest to see significant patterns it will sometimes prove useful to have agents examine trends and draw our attention only to cases where there are considerable changes, of say 10 or 20%. This will be another way to discover emerging subjects. This same methodology has a whole range of other applications including marketing, advertising, stock markets and even network management. Say, for example, that we want to monitor potential trouble spots on the system. Agent technologies measure usage at every major node of the system in terms of a typical throughput and usage. When these ratios change significantly the system identifies where they occur, and introduces standard adjustment measures. If these fail the system visualises relevant points in the neighbourhood of the node such that operators can see remotely where the problem might lie and take appropriate action. Hence, instead of trying to keep everything visible at all times the system only brings to our attention those areas where trouble could occur: an electronic equivalent of preventative medicine. Such strategies will no doubt be aided by the further development of sensory transducers whereby significant changes in heat within the system would also be rendered visible. Seeing the otherwise invisible is a key to navigating remotely through complex environments.

13. Comprehension and Prediction by Seeing Absence

In the early days of the scientific revolution there was great emphasis on the importance of inductive as opposed to deductive research, which entailed an emphasis on experience, experiment, often on a trial and error basis. As scientists gained a better understanding of the field to the extent that they were able to create physical and conceptual maps of their objects of study, it became possible to deduce what they had not yet observed. For example, from a careful observation of the motions of the known planets, astronomers were able to predict the location of Uranus and subsequently other planets. In combination with induction, deduction regained its role as an important ingredient in science.

The same proved true in chemistry. Once there was a periodic table, chemists found that the known chemicals helped them to chart the positions of as yet unknown compounds. Once we have a matrix we can see where there is activity and where activity is missing. By now, chemistry is expanding with enormous speed. It is estimated that every week over 14,000 new chemical combinations are discovered. As in the case of pilots flying at twice the speed of sound we need methods for abstraction from the day to day details, new ways of seeing patterns. Access, searching and navigation are not just about seeing what we can find, but also about strategies such that we see the appropriate subsets at any given time.

Until a generation ago mainframe computers typically relied on punch cards with holes. Each hole represented a specific configuration of a subject or field. Rods or wires were then used to determine which cards had the same fields. Early versions of neural networks adopted a virtual version of the same principle by shining virtual lights through configurations of holes. When the holes co-incided the subjects were the same. Database indexes effectively accomplish the same thing with one fundamental difference: we see the results but have no means of seeing the process.

To see a bigger picture we need to be able to see how the tiny details fit into the larger categories of human endeavour so that we can discern larger patterns. Roget as we saw had six basic classes (figure 1). Dewey had ten: 0) generalities; 1) philosophy and related disciplines; 2) religion; 3) social science; 4) language; 5) pure sciences; 6) technology; 7) the arts; 8) literature; 9) general geography and history. The Library of Congress has twenty such fundamental classes. Beneath these universal headings are many layers of subordinate categories hierarchically arranged. If we treat each of these layers as a plane, and have a way of moving seamlessly from one plane to the next, then operations performed at one level can be seen at various levels of abstraction.

Suppose, for example, that we have been searching for Renaissance publications by Florentine authors. Moving up to the highest level we can see on which fields they wrote: religion, science, art and so on. Moving back down a few levels we can identify which particular branches of science and art concerned them most. Going back to the top level we can also see that there were many topics which they did not discuss. The Renaissance view was supposedly universal in its spirit. In practice, it often had distinct limitations. If

we have access to multiple classification systems, then we can see how these patterns change as we look at them say, through the categories of Duke August and Leibniz at Wolfenbüttel or through the categories of Ranganathan's system. These approaches become the more fascinating when we take a comparative approach. How did German Lutheran cities differ from Italian Catholic or Swiss Calvinist cities in terms of their publications? How does religion influence the courses of study and research? What different cultural trends are evident from a comparison of publications in India, China, Turkey, Russia and Japan?

Such book catalogues and classification systems are the most important efforts at bringing order to the world in terms of subjects. But subjects in isolation are still only somewhat ordered information. Meaning which brings knowledge and wisdom requires more, namely a systematic ordering of these subjects in terms of their logical and ontological relations. Efforts in this direction go back at least to the *I Ching*. Aristotle, Thomas Aquinas, Ramus, Francis Bacon and Roget were among the many contributors to this tradition. In our generation, Dr. Dahlberg presented these fundamental categories in a systematic matrix³⁹. More recently these have been adapted by Anthony Judge into a matrix of nine columns and nine levels (figure 6), which generates a chart of 99 subjects. These range from fundamental sciences (00), astronomy (01) and earth (02) to freedom, liberation (97) and oneness, universality (99)⁴⁰. Anthony Judge is using this as an “experimental subject configuration for the exploration of interdisciplinary relationships between organizations, problems, strategies, values and human development”.

Heiner Benking, another speaker at the German chapter of the ISKO conference builds upon the same framework in his conceptual superstructure or cognitive Panorama Bridge, which is the basis of his Rubik's *Zauberwürfel* [*Cube of Ecology* or *Magic Cube*].⁴¹ He argues that one can use planes in order to see patterns in thought. These planes can include continua between the animate and the inanimate on one axis and between micro-, macro- and meso-scales on another axis.

	Matrix columns	Matrix levels
9	Condition of the whole	Experiential (modes of awareness)
8	Environmental manipulation	Experiential values
7	Resource redistribution	Innovative change (context strategies)
6	Communication reinforcement	Innovative change (structure)
5	Controlled movement	Concept formation (context)
4	Contextual renewal	Concept formation (structure)
3	Differentiated order	Social action (context)
2	Organized relations	Social action (structure)
1	Domain definition	Biosphere
0	Formal pre-conditions	Cosmosphere/Geosphere

Figure 6. An integrative matrix of human preoccupations by Anthony Judge (UIA) adapted from Dr. Ingetraut Dahlberg.

The planes can be used to see relations among different actions, options and strategies. They can be used to see different levels of abstraction and compare different viewpoints at a conceptual as well as at a perceptual level. A challenge remains how precisely we are to navigate between such conceptual landscapes and the knowledge structures of libraries which have been a main focus of this paper. At a programming level this should be relatively straightforward. Each of the ninety nine subjects is tagged with its equivalents in the main classification schemes. At the user level, this and similar matrices then become a natural extension of the system. When we use these categories as a filter to look at publications in the Renaissance or research trends in the late twentieth century, we have another means to comprehend which areas were the focus of attention and which were abandoned, or even forgotten. Search and access systems must help us to see absence as well as achievement, and possibly provoke us to look more closely at the spaces which are being ignored. Were they truly dead ends, have they surfaced in a new guise or do they now require new study?⁴²

In our view the question of viewpoints may be somewhat more complex than is generally imagined. Viewpoints are not just about comparing abstractions. They are also about different bodies of knowledge which are an essential ingredient of culture. An Englishman sees the world through the prisms of Shakespeare and Milton, an Italian through the verses of Dante and a German through the poetry of Goethe and Schiller.⁴³ Each of these geniuses did more than create poetry: they launched a heritage of associations which are shared by every cultured person in that tradition, such that there is a manageable corpus of phrases that is mutually recognised by all members. In order better to comprehend these shared traditions in the case of cultures completely foreign to us, it may prove useful to develop agents familiar with all the standard literature of those cultures such that they can help us to recognise quotes which are so familiar to natives that they are expressed as ordinary phrases, e.g. *To be or not to be* (Shakespeare) *Every beginning is difficult* (Goethe), *One must live to eat, not eat to live* (Molière), and yet evoke a wealth of associations which the outside visitor could not usually expect.

Here, at the end, we can only touch upon this most elusive aspect of navigation, which is not only about what they say or write. It is about what one culture asks and another does not, about which one culture discusses and the other is silent, for which one culture has a hundred terms (snow among the Inuit) and of which another culture has no experience (a nomad in parts of the Sahara). It will be a long time before all these challenges have been overcome. Yet if we recognise them clearly, there is no need to be overwhelmed by these challenges. We must continue the process of sense making and ordering the world, which began with our first libraries and schools and shall continue, we hope, forever. For in this lies our humanity.

14. Conclusions.

Amidst the sunny dreams of an information society where anyone can learn anything at anytime, there are dark clouds of those who fear that this is a futile quest. There are doomsayers⁴⁴ who claim that our comprehension of information has not kept pace with our production of information, that a new flood is upon us and that our prime concern should be to keep from drowning. These fears are not new. After all, in the Second World War, even large groups of individuals proved unable to do all the calculations necessary for determining the trajectories of heavy artillery shells. This was a chief motivation for the creation of the first supercomputer. In 1945, Vannevar Bush, himself a pioneer in computers and advisor to the President of the United States, recommended the use of hypertext as a way of mastering the problems of information. A few years later another pioneer, Douglas Engelbart, inventor of the mouse, suggested a variant of hypertext, which could be used in a collaborative environment. Since then hypertext, multimedia, collaborative work and learning have all become buzzwords. Even so, speaking in 1997 (at the Sixth World Wide Web Conference in Santa Clara), Engelbart noted with eloquent and humble wistfulness that the tools he had envisaged over 45 years ago are still not in place. On the surface the naysayers are right: we are still drowning in information.

On closer inspection there are clearly exceptions. Almost anyone can make a bank or credit card transaction anytime, anywhere, and the companies concerned have little trouble keeping track of them. The military can effectively see anyone, anytime, anywhere and any limitations in this regard will soon be overcome with the next generation of nomadic computing. It is not so much that the tools do not exist, but rather the uses to which they have been put. As McLuhan put it: the future is already here. It's just that its unevenly distributed.

If we were to take our cues from the military we would discover that information is the new battleground in more senses than one. Communications satellites, mobile computing, and meta-data are quickly replacing fighter planes, submarines and tanks as the arsenal of the future. The enormous military, political and business implications of the information society will undoubtedly be handled by the respective groups but this alone will not solve the challenges of understanding which have always been the province of scholars and members of the contemplative life. We need to fight our own spiritual battles if we are to use the technologies to increase the power of the mind.

This paper outlines a strategy for tackling the enormous problems at hand. It is at once a description of an existing prototype and a blueprint for its further development. The prototype is called a System for Universal Media Searching (SUMS) and was chosen as one of 19 projects to represent Canada at the G7 Information Society Exhibition in Brussels (February 1995), at the G7 Summit in Halifax (June 1995); as part of G7 pilot project 5 (Multimedia Access to World Cultural Heritage) at the Information Society and Developing Countries (ISAD) Conference in Midrand (May 1996); as an invited signatory to the Memorandum of Understanding for Multimedia Access to Europe's Cultural Heritage and as a member of the European Commission's MOSAIC project, designated as the museums section of the Trans European Networks (TEN) Project.

There are two fundamental premises: first, that the experience of libraries, museums, archives and similar institutions in organising, ordering, classing and accessing knowledge is an obvious point of departure for serious strategies of search and access; and second, that the methods used for presentation of knowledge offer valuable clues for a coherent access interface and strategy. This is a key to a common look and feel for all our activities, be it creating, classing, publishing or accessing.

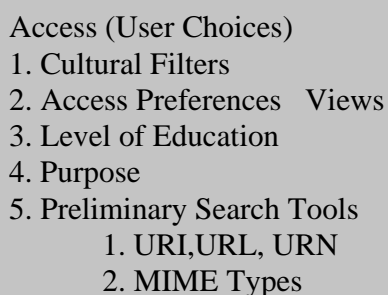
Following from these premises is a new approach to the traditions of knowledge collection, organisation and retrieval. At one end of the spectrum there has been a dream that everything could be collected in one centralized institution. This inspired the Library of Alexandria, the British Museum and a host of other efforts. At the other end of the spectrum there has been an assumption that everything could be decentralised in a completely distributed system. Our claim is that neither of these extremes can work, which means that a new model is called for: a centralised repository of meta-data, a digital reference room which is effectively a cumulative collection of all existing reference sections in libraries and museums.

Most thinkers have assumed that agents will play the role of electronic butlers at a personal level, producing tailor-made selections of newspaper clippings in keeping with our individual tastes, but perhaps they have a more universal role to play as electronic versions of library assistants in this digital reference room. All this will not be without effort or cost, but faced with an alternative of information chaos the investment is relatively modest.

The basic interface relies on three elements 1) six familiar questions: Who?, What?, Where?, When?, How?, Why?; 2) lists of choices which begin with ten basic notions: access, learning, levels (of knowledge), media, quality, quantity, questions, space, time, tools; 3) maps as a way of navigating spatially. The main body of the paper outlined the role of each of these, focussing on the power of strategic use of simple questions, the significance of purpose as orientation, the use of quality and quantity as means of refining queries; the role of maps and geographical modes, and the use of integrating tools.

Traditional presentation and access methods have been two-dimensional. A comprehensive system must incorporate these traditions and develop means of moving seamlessly from two to three-dimensional spaces. The latter part of the paper explored some of the potentials of these emerging technologies in order to visualise connections in conceptual spaces and for purposes of comprehension and prediction by seeing absence as well as achievement.

Appendix 1 Key Elements of the SUMMA Model (©1997) as a Framework for a Meta-Data Digital Reference Room

- 
- Access (User Choices)
 - 1. Cultural Filters
 - 2. Access Preferences Views
 - 3. Level of Education
 - 4. Purpose
 - 5. Preliminary Search Tools
 - 1. URI, URL, URN
 - 2. MIME Types

Notes

¹ Cf. the work of Leonard Kleinrock.

See : <http://www.lk.cs.ucla.edu>.

² Levy, Pierre, (1996), *The Second Flood-Report on Cyberculture*, Council of Europe: Strasbourg, CC-CULT, 27D. Cf. Levy, Pierre, (1990), *Les Technologies de l'intelligence*, Paris: La Découverte.

³ For a further discussion of this trend see: Veltman, Kim H. (1997), Why Computers are Transforming the Meaning of Education, *ED-Media and ED-Telecom Conference, Calgary, June 1997*, ed. Müldner, Tomasz, Reeves, Thomas C., Charlottesville: Association for the Advancement of Computing in Education, vol. II, 1058-1076 available electronically at <http://www.web.net/~akfc/sums/articles/Education.html>

⁴ Recently, XML has become part of a more complex architecture strategy of the W3 Consortium which includes Resource Description Framework (RDF), Protocol for Internet Content Selection (PICS 2.0) and privacy initiatives (P3P).

See: <http://www.w3.org/TR/NOTE-rdfarch>.

⁵ On this topic see: <http://www.parc.xerox.com/spl/projects/mops/existing-mops.html>
<http://www.parc.xerox.com/spl/projects/oi/default.html>

⁶ By this reasoning, each exercise requires its own software. Hence writing requires its own software, e.g. *Word Perfect* or *Microsoft Word*; drawing requires another software, such as *Corel Draw*, or *Adobe Photoshop*; design requires other software again: e.g. *Alias-Wavefront* or *Softimage*.

⁷ For a more detailed disussion of this concept see: Veltman, Kim H. (1997), Towards a Global Vision of Meta-data: A Digital Reference 'Room', *2nd International Conference. Cultural Heritage Networks Hypermedia, Milan, September 1997*, Milan: Politecnico di Milano (in press).

⁸ Such a methodology has implications for hardware and network strategies. It means, for instance that users will typically engage at least two connections simultaneously, one to the location they are searching, a second to the on-line digital reference room. In the case of everyday searches where a smaller set of reference materials would suffice, it is perfectly possible to imagine these on the hard drive for ready reference. At present the great debates of personal computers versus network computers are ostensibly about purely technical questions: whether one uses a regular hard drive with resident software or a thin client which relies on software mainly on a remote server, the assumption being that users could readily have everything on remote servers and thus effectively be able to do without hard drives. Perhaps we need more on the home front than some programmers suspect and more connectivity than they foresee.

⁹ A sceptic may rightly object that there is a fundamental flaw in this approach, namely, that these ordered lists in libraries can never pretend to cover the whole of knowledge. We would agree that this point is well taken, but insist that this does not diminish the legitimacy of using the assets of libraries to the extent that they are applicable. These

limitations result in part from different kinds of knowledge. Libraries traditionally focus on books of enduring value or long term knowledge, which is relatively static in nature. By contrast, materials of passing interest were often classed under *Ephemera*, as a way of avoiding the problems of materials where categories were dynamic and constantly changing.

It is instructive to note that when the *Association Internationale de Bibliographie* was founded in the latter nineteenth century, it soon split into two organisations, one which became ISO TC 37, focussed on the categories and classing of established knowledge, whereas the *Union Internationale des Associations* (UIA) focussed its efforts on classing fields that were emerging and not yet clearly defined. It is therefore no coincidence, that the Director of Communication and Research of the UIA, Mr. Anthony Judge, is such a pioneer in the classing of nascent subjects such as world problems. See, for instance, Benking Heiner, Judge, Anthony J. N., (1994), *Design Considerations for Spatial Metaphors: Reflections on the Evolution of Viewpoint and Transportation Systems*, Position Paper: *ACM-ECHT94 Workshop on spatial Metaphors. Workshop at the European Conference on Hypermedia Technology, Edinburgh, 18-23 September 1994*, available electronically at <http://www.lcc.gatech.edu/~dieberger/ECHT94.WS.Benking.html> . For a more thorough listing, see the homepages of the authors at <http://newciv.org/members.benking> and <http://www.uia.org> and more specifically under *Research on Transdisciplinary Representation and Conceptual Navigation* at <http://www.uia.org/uiares/resknow.htm>.

The Internet is producing electronic versions of both enduring content found in books and ephemeral materials, but unlike libraries there is not yet a coherent method for distinguishing between them. A whole number of initiatives have been undertaken to remedy this situation, including high level domain addresses, more precise URLs, URNs and URIs, meta-data tagging in http protocols, and other meta-data schemes which have been reviewed by the author in the article cited in note 1. As these new methods bring greater discipline to materials on the net they will become more amenable to the methods used by libraries. In the meantime, the emphasis on sheer number crunching which some assume as a complete solution for all electronic knowledge should perhaps be applied particularly to these undisciplined portions of the net.

In some cases, the on-line ephemera have characteristics, which are rarely found at all in libraries and are valuable precisely and only because they are available to some persons hours, minutes or sometimes even seconds before they are available to others, notably, stock exchange figures, sports and race track information and the like involving bets. Some of this material is so fleeting that it loses almost all its monetary value within twenty-four hours. Search strategies for such ephemera are predictably different than those of the eternal truths.

¹⁰ Day, A. Colin, (1992), *Roget's Thesaurus of the Bible*, San Francisco: Harpers San Francisco.

¹¹ Such strategies are, of course, not without their dangers. One has to be very careful to distinguish the user's interests as a professional from their leisure interests. A nuclear physicist might well do searches on isotopes and quarks in one capacity and turn to sports or sex in the other. If such professional and leisure modes were mixed the resulting search strategies might be more than mixed.

A more important potential role for agents lies in translating general lists of search terms to more controlled lists which can be co-ordinated with subject lists and classification systems. In studying a topic we typically make a list of terms or keywords which interest us. For instance, a user may be interested in adaptive modelling, complex adaptive modelling and conceptual modelling and write these terms sometimes in this form, sometimes in reverse as modelling, adaptive etc. An agent would recognize that the user is interested in modelling, adaptive, complex adaptive and conceptual . It would create authority lists with controlled vocabularies, check which terms are found in subject lists of standard classifications and thus arrive gradually at a distinction between those terms which link directly with recognized fields of enduring knowledge, reflected in library catalogues, and those terms which represent new areas of study to date perhaps only recorded in citation indexes. In so doing one would create bridges between simple lists, thesauri and classification systems. The discipline of these more controlled lists could then be used to call up synonyms, broader, narrower and other related terms.

¹² Some will argue that the making of lists is an outdated exercise because the power of computers is now so great it is easier to search everything available through brute force number crunching than to bother with the niceties of lists. As is so often the case in life brute force has limitations, which are overshadowed by intelligence.

¹³ The philosophy behind this aspect of the system is simple. Names remain the same, but it makes little sense to overwhelm children with a list of millions of names (such as the Library of Congress authority list) when they are still learning to read their first names. Hence beginners are given minimal subsets which are gradually increased as their horizons expand. This is effectively a simulation in electronic form of traditional experience. A child going to the resource centre or library in a kindergarten would have a very small list of names. In an elementary and secondary school the catalogue would increase accordingly. In a university library the names would be larger again and continue to expand as the research student was introduced to the catalogue of the world's great libraries (National Union Catalogue, British Library and Bibliothèque Nationale).

¹⁴ Dictionaries are objective in the sense that there is only one definition that corresponds to the definition in the *Oxford English Dictionary (OED)*. The potential definition of a word remains subjective to the extent that there are definitions other than those in the *OED*. The same principle applies to classification systems and encyclopaedias.

¹⁵ Samurin, E. I., (1977), *Geschichte der bibliotekarisch-bibliographischen Klassifikation*, [History of Library and Bibliographical Classification], Munich: Verlag Dokumentation. Translated from the original Russian: *Ocerki po istorii bibliotечно-bibliograficeskoj klassificacii*, (1955-1959), Moscow, 2 vol.

¹⁶ These basic questions may have unexpected applications. In a very stimulating article, Professor Clare Beghtol, suggested that text types might prove an important way of classing materials. See: Beghtol, Clare, (1997), Stories: Applications of Narrative Discourse Analysis to Issues in Information Storage and retrieval, *Knowledge Organization*, 24 (2), 64-71.

	Who	What	Where	When	How	Why
Ruthrof 1981	Personae	Events (non human) Acts (human)	Space	Time		
Brewer	Personae	Events	Setting	Setting	Narrator	Resolution

1985			(Location)	(Time)		
Halasz	Existents	Events	Existents			
1987	(Characters)		(setting)			
Polkinghorne	People	Events				Voice
1988		Actions (human)				
Lamarque	Characters	Structure				
1990		(plot, including events)		Time		Voice
Rigney	Actors	Events	Place	Time		
1990	(individual and/or collective)					
Clark	Agents	Plot	Scene			Narrator End
1995						Voice

Figure 7. Adaptation of Begthol's chart of narrative elements from non-ISAR fields in keeping with six basic questions.

Following the typology of Egon Werlich, Egon, (1975), *Typologie der Texte: Entwurf eines textlinguistischen Modells zur Grundlegung einer Textgrammatik* [The Typology of Texts: An Outline of a Text-Linguistic Model for the Establishment of a Grammar of Text], Heidelberg: Quelle und Meyer, Begthol suggested a basic distinction between narrative texts and non-narrative types: e.g. description, exposition, argumentation and instruction. Combining the research of six authors she proposed a table of narrative elements from non-ISAR fields (p.66). A slight adaptation of these fields shows how these can be aligned in terms of the six basic questions (figure 7).

This suggests a definition of narrative as a text type which applies to all six questions in ways that the others do not. A more thorough study would need to relate these questions to the typological work of Lanser, Susan S., (1981), *The Narrative Act. Point of View in Prose Fiction*, Princeton: Princeton University Press; Lintvelt, Jaap, (1981), *Essai de Typologie Narrative: Le point de vue. Théorie et Analyse*, [Essay of Narrative Typology: Point of View. Theory and Analysis], Paris: Librairie José Corti; Lindemann, Bernhard (1987), *Einige Frage an eine Theorie der Sprachlichen Perspektivierung*, [Some Questions concerning a Perspectival Treatment of Language], and *Perspektivität in Sprache und Text*, [Perspectivity in Language and Text], Hrsg. Peter Canisius, Bochum: Verlag Dr. Norbert Brockmeyer, 1-51. In keeping with the method outlined below the various questions might each be placed on an independent plane. Different text types would then link differing numbers of planes.

¹⁷ There will, of course, be more complex instances, as when a researcher wishes to determine all instances of a subject in all media, in order to study the relative significance of particular media. Was the theme used more in books and written materials or primarily in paintings? Did the advent of film and television increase the use of the theme or lead to its demise? These are cases where agent technologies will increasingly serve to do most of the preparatory work. In the past researchers had research assistants to find the raw material. Agents will translate this process into an electronic form. The challenge of making sense of the raw data will remain with the researcher.

¹⁸ Tufte, Edward R., (1990), *Envisioning Information*, Cheshire, Conn.: Graphics Press. For an early discussion of these themes in terms of computer graphics see Benking, Heiner and Steffen, Hinrich, (1985), *Computer Graphics for Management, Processing*,

Analysis and Output of Spatial Data (Corporate, Administrative, Facilities and Market), WCGA-CAMP '85 (*World Computer Graphics Association and Computer Graphics for Management and Productivity*) Conference, Section C.7.2 ,*Future Trends*, Berlin, 440-458.

¹⁹ For instance, these maps can be linked via Global Positioning Systems (GPS) to moving objects such as cars or even moving mail and package containers such that one can trace their movements. If this were applied to all valuables, stealing and robbery could soon be outmoded bad habits.

²⁰ See: http://www.artcom.de/t_vision/welcome.en.

²¹ Lukuge, Ishanta, Ishizaki, Suguru, (1995), Geospace. An Interactive Visualization System for Exploring Complex Information Spaces, *CHI 95 Proceedings: Conference on Human Factors in Computing Systems: Mosaic of Creativity, May 7-11 1995, Denver, Co.* See: http://www.acm.org.sigchi/chi95/Electronic/documents/papers/il_bdy.htm.

²² A number of companies are active in this domain. The Environmental Systems Research Institute (ESRI) is creating maps of the world. Autodesk has already created such maps for the world, North America and the Netherlands (cf. <http://www.mapguide.com>). For a review of these developments see Potmesil, Michael, (1997), *Maps Alive: Viewing Geospatial Information on the WWW*, *Bell Laboratories, Lucent Technologies TEC 153*, Holmdel, New Jersey, 1-14 or electronically on the web at <http://www6.nttlabs.com/HyperNews/get/PAPER/30.html>.

²³ See for instance, Forte, Maurizio, ed., (1997), *Archéologie virtuelle. Le passé retrouvé*, Paris: Arthaud, based on the Italian, (1996), *Archeologia, percorsi virtuali nelle civiltà scomparse*, [Archaeology, *Virtual Journeys through Lost Civilisations*], Milan: Mondadori.

²⁴ See, for instance, the differences between Professors Willibald Sauerländer and Martin Gosebruch.

²⁵ See: Yates, Frances, Dame, (1966), *The Art of Memory*, London: Routledge and Kegan Paul.

²⁶ See, for instance, the work of Tony Judge in *Visualising World Problems, Organisations, Values at*: <http://www.uia.org/uiademo/VRML/vrmldemo.htm>.

²⁷ In the last five years the processing power of computers has evolved from 33 Herz to over 200 Herz. Within the next three years this power will increase to at least 1000 Herz, by which time computers will have the power to operate in three-dimensional environments.

²⁸ Massachusetts Institute of Technology, (1995), *Media Laboratory. Projects. February 1995*, Cambridge, Mass.: MIT, 6.

²⁹ See: <http://almond.sru.cs.cmu.edu/afs/cs/project/sage/mosaic/samples/sage/3d.html>.

³⁰ A more interesting application is in the context of Collaborative Integrated Communications for Construction (CICC) available electronically at

<http://www.hhdc.bicc.com/people/dleever/papers/cycleof.htm>, which envisages a cycle of cognition in which the landscape is but one of six elements, namely, map, landscape, room, table, theatre, home. For an excellent summary of some of the major systems presently available see Peter Young (1997), *Three Dimensional Information Visualisation* available electronically at

<http://rvprl.cs.uml.edu/shootout/viz/vizsem/3dinfoviz.htm>.

³¹ See: Rao, Ramana, Pedersen, Jan O., Hearst, Marti A., Mackinlay, Jock D., Card, Stuart K., Masinter, Larry, Halvorsen, Per-Kristian, Robertson, George C., (1995), Rich Interaction in the Digital Library, *Communications of the ACM*, New York, April, 38 (4), 29-39. Card, Stuart (1996), Visualizing Retrieved Information, *IEEE Computer Graphics and Applications*.

³² Kling, Ulrich (1994), Neue Werkzeuge zur Erstellung und Präsentation von Lern und Unterrichtsmaterialien [New Tools for the Production and Presentation of Learning and Instructional Materials], *Learntec 93. Europäischer Kongress für Bildungstechnologie und betriebliche Bildung*, ed. Beck, Uwe, Sommer, Winfried, Berlin: Springer Verlag, 335-360. Cf. <http://www-cui.darmstadt.gmd.de/visit/Activities/Lyberworld>.

The GMD also organizes research on Foundations of Advanced Three Dimensional Information Visualization Applications (FADIVA) and Visual Information Retrieval Interfaces (VIRI).

See: <http://www-cui.darmstadt.gmd.de/visit/Activities/Viri/visual.html>.

³³ Ibid., 336-340. Cf. Streitz, N., Hannemann, J., Lemke, J. et al., (1992), SEPIA: A Cooperative Hypermedia Authoring Environment, *Proceedings of the ACM Conference on Hypertext, ECHT '92*, Milan, 11-22.

³⁴ See, for instance, the methods being developed by Lucent in their Live Web Stationery at <http://medusa.multimedia.bell-labs.com/LWS/>.

³⁵ Munzner, Tamara, Hoffman, Eric, Claffy, K., Fenner, Bill, (1996), Visualizing the Global Topology of the Mbone, *Proceeding of the 1996 IEEE Symposium on Information Visualization, San Francisco, October 28-29*, 85-92 available electronically at:

<http://www-graphics.stanford.edu/papers/bone>.

³⁶ See: <http://www.hitl.washington.edu/>.

³⁷ The only work in this direction, of which the author is aware, is the metadatabase research at Rensselaer Polytechnic in conjunction with Metaworld Enterprises entailing a Two Stage Entity Relationship (TSER) in the context of an Information Base Modelling System. See <http://viu.eng.rpi.edu/overview2.html> and <http://viu.eng.rpi.edu/IBMS.html>.

³⁸ Martin, Steve, Clarke, Steve, Lehaney, Brian, (1996), Problem Situation Resolution, and Technical, Practical and Emancipatory Aspects of Problem Structuring Methods, *PARM '96, Practical Aspects of Knowledge Management, First International Conference, Basel, 30-31 October 1996*, 179-186. I am grateful to Heiner Benking for this reference.

³⁹ Any attempt at ontological structuring will inevitably inspire critics to claim that a slightly different arrangement would have been closer to the true heirarchy. While such debates have their value, it is important to recognize that even if there is no complete agreement about a final configuration, the conflicting versions can still contribute to new insights, by challenging us to look at trends from a more universal level.

⁴⁰ See: <http://www.uia.org/webints/aaintmat.htm>.

⁴¹ See: Benking, Heiner, (1997), Understanding and Sharing in a Cognitive Panorama. *Culture of Peace and Intersymp 97. 9th International Conference on Systems Research, Informatics and Cybernetics, August 18-23*, Baden-Baden, available electronically at <http://www3.informatik.uni-erlangen.de:1200/Staff/graham/benking/index.html>.

Cf. Benking, Heiner, (1992), Bridges and a Master Plan for Islands of Data in a Labyrinth

of Environmental and Economic Information, *Materials and Environment. Databases and Definition Problems: Workshop M. and System Presentation. 13th ICSU-CODATA Conference in collaboration with the ICSU-Panel on World Data Centers, Beijing, October 1992.*

⁴² If we look, for instance, at classifications of the Middle Ages there were no categories for science (as we now know it) or psychology. What we call *science* was typically (*natural*) *philosophy* or was included under the rubric of the *quadrivium* (arithmetic, geometry, music and astronomy). Psychology was often in literature such as the *Roman de la Rose*.

⁴³ For a further discussion of these problems see Veltman, Kim H., (1997), *Why Culture is Important [in a World of New Technologies]*, 28th *Annual Conference: International Institute of Communications Conference, October 1997*, London: International Institute of Communications, 1-10.

⁴⁴ For a recent example see Shenk, David, (1997), *Data Smog. Surviving the Information Glut*, New York: Harper Collins.

Acknowledgements

This paper is dedicated to Professor Dr. Ingetraut Dahlberg who has generously encouraged the author for the past fifteen years. I am deeply grateful for her gentle inspiration. I am very grateful to Deans Wiebe Bijker (Maastricht) and Hans Koolmees (IDM, Maastricht) for provocative questions during a recent visit which helped me to shape the discussions about access versus presentation. 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 kindly read the manuscript and offered suggestions.

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 for their enduring support and encouragement.

Author's Mailing Address:
Dr. Kim Henry Veltman
Director, Perspective Unit
University of Toronto
100 Lombard Street, Suite 303
Toronto, Ontario M5C 1M3

Canada
tel. (416) 303-4033
fax. (416) 941-9581
E-mail kim_veltman@hotmail.com
Home page: <http://www.web.net/~akfc>