Kim H. Veltman

Future Strategies and Visions for SUMS and SUMMA

Prepared for: Acts of Congress: *EVA 96'*, London, 1996

Never Published

---

1. Introduction

SUMS (Copyright 1992-1996) stands for System for Universal Media Searching. It has been described as a software for conceptual navigation in cyberspace. SUMMA (Copyright 1996) stands for System for Universal Multi Media Access. These grew out of the needs of a scholar trying to organize and display material for a standard bibliography on perspective and research on Leonardo.[1] Since this work involved materials in major libraries and museums, it gained the attention of the Canadian Heritage Information Network (CHIN), through a keynote lecture in Winnipeg in 1990[2], and to a wider audience through a keynote at the CIDOC section of ICOM (Quebec, 1992)[3]. In February, 1995, SUMS was chosen as one of 18 national projects to represent Canada at the G7 Information Society Exhibition in Brussels. It was chosen again for the World Summit in Halifax (June 1995). On that occasion the members of G7 agreed to go ahead with 11 pilot projects relating to the major areas of the information society[4]. Each of these projects had a lead country. Italy was chosen to head Pilot Project 5, Multimedia Access to World Cultural Heritage, which focusses on museums. At the Information Society and Developing Countries (ISAD) Conference in Midrand, South Africa (May 1996) four projects from Italy and Canada were chosen to represent Pilot Project 5[5] (fig. 1).

Function	Exhibit	Organization	City
Capture	3-D Laser Camera	National Research Council	Ottawa
Archive	Florence, Leonardo	Museum for History of Science	Florence
Display	Tomb of Nefertari	Infobyte, ENEL	Rome
Navigate	SUMS	SUMS Corp.	Toronto

Fig. 1. List of four sections of G7 pilot project 5.

SUMS began as an interface to local databases It was chosen for the navigation section of Pilot Project 5 and then extended to include seamless links with the Internet. It is presently HTML (Hyper text Markup Language) compliant, uses the Z39.50 protocol, and relies on SQL for its query interfaces. These protocols will change as software and hardware standards evolve. The methodology is independent of any particular technology. The underlying philosophy is to keep the tools for navigation as simple and homogeneous as possible notwithstanding changes in levels of education and enormous changes in complexity of facts. The key idea is that one can get to the same information in a variety of ways. It has a variety of search methods. One is via maps. This will in future be linked with GIS (Geographical Information Systems), AM/FM (Area Management/ Facilities Management) and GPS (Global Positioning Systems).

1. Access
2. Learning
3. Levels
4. Media
5. Quality
6. Quantity
7. Questions
8. Space
9. Time
10. Tools.

Fig. 2. List of ten basic choices in SUMS.

A second method is by asking six basic questions: who, what where, when, how and why? These are used in conjunction with ten basic choices (fig. 2). The power of SUMS comes through combinations thereof. Each of these choices breaks down into further lists of ten or less choices. There are hundreds of these lists. Combined in various ways these generate many thousands of choices. SUMS is a first step towards SUMMA, which will have four interrelated functions (fig. 3):

1) classing information to make it searchable and to create personal search engines
2) searching seamlessly in a distributed environment
3) organizing the results of one's searches
4) learning in new ways[6]

Fig. 3. Four interrelated functions in SUMS.

1. Classing

There are famous sayings about databases such as "garbage in, garbage out" and "one gets what one puts in." SUMS uses the principles behind these ideas as a point of departure for basic strategies in classing material in order to search, organize and learn. These boil down to some simple rules. If one searches indiscriminately one finds things indiscriminately. Hence, at least five different kinds of more organized search are identified (fig 4):

1) personal classification
2) personal lists based on database fields
3) standard subjects
4) standard classification system
5) multiple classification systems

Fig. 4. Five kinds of access strategies to knowledge

Personal Classification

We all class material although we do not always do so consciously. The lists we make reveal as much about our preparation as they do about the precision with which we are seeking knowledge about a given topic. Student *A* is asked to make a list of all the categories under who. This individual begins with persons, then decide that they wish to have professions as examples of which they identify explorers, firemen, pilots. Student *B*'s list includes artists, chemists, doctors, engineers, mathematicians, physicists and scientists. Student *C*'s list might be identical to *B*'s list initially, but then add a series of distinctions under physicists, namely atomic physicists, particle physicists and muon experts. What this boils down to is asking persons to write a list of terms which are important to them. This gives a personal dictionary of the terms in which they are interested.

If these lists are viewed hierarchically (fig. 5) they provide us with useful clues about the students. Given the absence of female professions it is likely that the lists were composed by males. The simplicity of the first list suggests a boy aged 6-12. The second list suggests a high school student, whereas the third list suggests a university student studying science. Ultimately the details are less important than the general principle whereby the number of levels of distinctions reveal the complexity of their author's mind. The more levels, the greater the precision with which one can search. A complete beginner sees no need for any distinctions, has no levels and is content simply to have a list of all persons under *who*. The expert has many levels. Thus each personal classification system provides a cognitive map of what they consider important.

In this example the question *who* was used. The same principle applies to all of the basic questions, namely, *what*, *where*, *when*, *how*, and *why*. For instance, a beginner will have an undifferentiated list of subjects/objects. A specialist will have many levels of distinctions. If these lists are to be compatible with other sources then it is important that they use terms accepted by the standard subject lists.

Student A WHO

Professions
Explorers
Firemen
Pilots

Student B WHO
Professions
Artists
Doctors
Engineers
Mathematicians
Scientists
Chemists
Physicists

Student C
WHO
Professions
Artists
Doctors
Engineers
Mathematicians
Scientists
Chemists
Physicists
Atomic Physicists
Particle Physicists
Muon Experts

Fig. 5. Hierarchical lists of students A, B and C.

One of the obvious applications of such a personal vocabulary of terms entails the ordering of materials found when surfing the Internet. One comes across an item. One determines whether it is a *who*, *what*, *where*, *when*, *how*, or *why* question, checks the categories one has listed under the question and classes or tags the new-found material with the appropriate category. If the material requires a new category then one adds this in the appropriate place.

Personal Lists Based on Database Fields

A second level of complexity assumes that if one wishes clues as to what one might get out of a database, the best place to look is in the categories or fields used to enter materials into the database.

WHO Artist/Maker

Culture
Ethnic Tradition
Manufacturer
Tribe

WHAT Institution
    Department
Object
    Object Name
    Object Type
Material
Medium
Number
    Accession Number
    Borden Number
    Catalogue Number
Quantity
School/Style
Support
Subject/Image
SPACE
    Height
    Width
    Length
    Outside Diameter
    Depth
    Unit_Linear
Unit
Title

WHERE Manufacturer Country
    Manufacturer Province
Origin Country
    Origin Province/Territory
Use Country
    Use Province/Territory

WHEN Begin Date 1
End Date 2

HOW
Technique

Fig. 6. List of core fields in the CHIN Human History Database arranged hierarchically in terms of basic questions. Italicized terms indicate that a broader topic has been added to cover two narrower topics. A capitalized italicized term indicates one of the basic choices in SUMS.

WHO
Collector
Collector

WHAT
Chemical Classification
Institution
Department
Discipline
Number
Acquisition Number
Accession Number
Catalogue Number
Phylum/Division
Class
Order
Family
Genus
Species
Subspecies
Variety
Specimen Nature
Quantity Specimen
Sex
Specific lithography
Lithostratigraphic

WHERE
Origin-Continent
Origin-Country
Origin-Province/Territory
Origin-Geographical Province
Origin -Ocean /Basin
Locality number
Lot
Mine
Spec Group/Region

WHEN
Age/Stage
Geological Period

Fig. 7. List of core fields in the CHIN Natural Sciences Database arranged hierarchically in terms of basic questions.

These fields are subdivided in terms of six basic questions and then organized hierarchically. To use a concrete example one might begin with core fields in the CHIN (Canadian Heritage Information Network) human history database and arrange these hierarchically (fig. 6). This same approach applies to titles in the CHIN Natural Sciences Database (fig. 7) and can readily be applied to the Museum Documentation Association (MDA) or any major database system. These lists effectively provide sets of universal terms which can then be queried in terms of individuals. The system allows users to view the lists either alphabetically (flat file) or hierarchically (tree). Such lists provide a basic orientation for those wishing to class, search, order and learn. As outlined in figure 3 above, successively more complex approaches will relate these terms to standard subject lists, standard classification systems and multiple classification systems, a detailed discussion of which is beyond the scope of the present paper.

3. Searching

To deal with more advanced searching SUMS expands the basic choices mentioned in figure one. For example, the third item, levels, is expanded into ten levels (of knowledge, fig. 8). Internal analyses then breaks down into a series of further choices. For the purposes of this paper we focus on examples which integrate some of the basic CHIN database categories in alphabetical (fig. 9).

The above cases entail searches for a specific person, thing or place. In research we often do not know exactly who or what we are looking for. We may, for instance, be studying mediaeval art and wish to know names of artists of the thirteenth century. This is merely a question of asking for a universal category and then requesting a list of particulars/ individuals under that heading. In other cases qualitative descriptors and relators will provide ways of exploring the context without necessarily knowing the details. Such meta-searching which will require fuzzy logic and neural networks is beyond the scope of this paper.

LEVELS
1. Terms
2. Definitions
3. Explanations
4. Titles
5. Partial Contents
6. Full Contents
7. Internal Analyses
8. External Analyses
9. Restorations
10. Reconstructions

Fig. 8. List of ten levels of knowledge used in SUMS.

INTERNAL ANALYSES

1. Identifications
2. Interpretations
3. Analyses

IDENTIFICATIONS
1. Art
2. Books
3. Objects
4. Persons

IDENTIFICATIONS ART
1. Function
2. Motif
3. Part
4. Study

ART STUDY
1. Administration
2. Alternate Names
3. Conservation
4. Details
5. History
6. Image History
7. Loans

ART DETAILS
Catalogue No.
Culture
School/Style
Object Type
Object Name
Material
Medium
Support
Technique
Produced in
Used in

Figure 9. Examples of how the categories used in the CHIN databases recur in the SUMS framework in the context of searching. Note that not all of the initial lists are expanded into their full form in the interests of space. Such lists provide users with details appropriate to a given interest rather than overwhelming them with the information found in all the fields at once.

4. Organizing

The combined choices of SUMS provide the equivalent of an electronic bucket for organizing the materials one finds. Because the complexity of this bucket varies with the complexity of the user it offers a dynamic rather than a static approach to organization. As object oriented programming evolves, closer connections can be established between the classing, searching and organizing functions of the system.

One possible scenario is as follows. A user defines their vocabulary for classing knowledge, which amounts to the basic terms they wish to use in doing searches and for organizing what they find. It suffices that they simply list the terms orally. The system assigns them an appropriate position in a hierarchy. The finished hierarchy becomes the parameter with which they define the limits of their searching and helps the system to adjust the complexity of the choices which the user is offered. For instance, if the user has only one level of hierarchy in their distinctions, they will not be offered choices with complex hierarchical choices. Each level of the educational system will have its own standard configurations of choices which will function in the manner of templates. Individuals with their own variants can submit these for consideration. If accepted these become available for users wishing to have alternative methods of organizing the materials. The system thus remains open while at the same time having a coherent, integrating framework which assures that every additional contribution continues to contribute to something more than the sum of the individual parts.

5. Learning

The same framework used for 1) classing, 2) searching and 3)organizing knowledge, will serve also for 4) learning in new ways. At the lower levels the same personal classification list used for the first three will be used for this fourth purpose. In more advanced levels this list will be substituted by a) database fields b) a standard subjects list (such as Library of Congress), c) a standard classification system (such as Dewey, Göttingen or Library of Congress, and d) multiple classification systems (cf. fig. 3 above). Each of these successive steps provides a greater range of vocabulary and a greater precision in terms of defining the parameters of what one wishes to learn.

From one's educational level the system will have a basic framework for knowing what it might expect of the learner. The reasons for learning also need to be established. Is this being done for leisure enjoyment or will the lives of persons depend on one's performance in a critical job such as a dam supervisor or a squadron leader? The system will trace one's progress, will provide periodic tests as needed or as wished, and will continue to add to the complexity of the materials presented in accordance with one's preparation. If a person has no knowledge of French they will not usually be shown French texts or manuscripts. As they learn languages, materials in that language become available. There are, of course occasions when they may wish to peruse a manuscript in a language which they cannot read. Provisions for overriding the usual arrangements will exist and it is likely that this would have a fee attached to it. The system will encourage all serious study while discouraging incidental browsing unless this browsing is clearly part of some larger plan.

An essential aspect of the complete project will be a new contextualization of anything learned such that one can see where any fact fits into an exam, where this fits into a text, a course, a curriculum and the corpus of knowledge in the field. A second dimension of this contextualization will be cultural. A third aspect will be historical. As a result one will, for instance, be able to choose the term China and watch how the territory associated with that country varies historically from a Chinese viewpoint and then examine how the same territory varies historically when examined from a Russian, Japanese or Indian viewpoint. Hence there will be historical, cultural dynamic maps. These maps will serve in helping to trace not only the growth of answers but also the growth of questions. Perrault[7] has offered a fascinating enumeration of categorical and relational tabulations for major thinkers including Aristotle, Lull, Kant, Mills, Pages, Kervagent, Gardin and Farradane. Such a list, linked with the six questions could provide a framework for showing how the filters of study have evolved over time.

There will also be the equivalents for ideas. There will be new kinds of spreadsheets to visualize the growth of manuscripts and books devoted to a given topic, the development of new branches within a given field. Hence just as there is gradual development in the complexity of a personal classification scheme, one will able to trace the evolving complexification of classification schemes culturally. The interest here will not be in some revival of the phylogenetic-ontogenetic analogies but rather in a new awareness of learning which is the secret behind all true growth.

6. Conclusions

This paper has outlined some features of a System for Universal Media Searching (SUMS), which is a prototype for conceptual navigation in cyberspace and serves as a first step towards a System for Universal Multi Media Access (SUMMA) with four basic features: classing, searching, ordering and learning.

SUMS is one of four parts of a larger strategy within pilot project 5 of G7, *Multimedia Access to World Cultural Heritage*, which includes sections on how to capture, archive, display as well navigate through museum materials. All four of these sections reflect new methods for approaching cultutural heritage. All four of these methods are characterized by the fact that they are not Eurocentric. They are effectively electronic receptacles, which apply equally to the cultures of Africa, Asia, South America, indeed anywhere in the world. Each culture can present their materials through their own lens such that the same object can be looked from a series of viewpoints. Hence, Jerusalem can, for instance, be studied from a Jewish, Christian or Muslim viewpoint, which will give new meanings to multiculturalism as the G7 projects evolve from an abstract vision to practical examples from all over the world.

Acknowledgements

In August 1995, The SUMS Corporation was created with Kim Veltman, Jordan Christensen, Andrew McCutcheon and Rakesh Jethwa as founding members. Jordan

Christensen has directed a team of some 20 programmers and assistants including Andrew McCutcheon, David Seale, Greg Stuart, Jimmy Woo and Miklos Fülop. Their help in creating a working demonstration is gratefully acknowledged.

The background work for these projects was done in the course of many years of scholarships, fellowships and other grants. These include the Canada Council (1971-1975), the Wellcome Trust (1975-1977), Volkswagen Foundation (1977-1979), Alexander von Humboldt Foundation (1979-1981), Thyssen Foundation (1981), Gerda Henkel Foundation (1982-1984), University of Toronto (1984-1989), Getty Trust (1986-1987) and the Social Sciences and Humanities Research Council of Canada (1987-1992), CHIN, Industry Canada, the Department of Canadian Heritage, and the Ministry of Foreign Affairs, Italy. We acknowledge the help of Professor Derrick de Kerckhove, the Director of the McLuhan Program in providing SUMS with space and basic amenities. These and the lease of equipment were made possible through a grant from Origin (formerly BSO/Origin) for which we are grateful. I thank Dr. James Helmsley for his kind invitation to present SUMS at the London conference, attendance of which was made possible through a travel grant from Bell MediaLinx.

Notes

[1] The project began as a front-end using D-Base and Toolbook as frameworks. The limitations of Toolbook led to a decision to develop our own solution using C++ with the help of a small team of young programmers. When HTML made great advances, it was decided to use this alternative as a temporary solution. For further information about the system see http://www. mcluhan.utoronto.ca.

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

[3] 1992 ICOM Conference, CIDOC section, Opening Keynote Address, "Past Imprecision for Future Standards: Computers and New Roads to Knowledge", Quebec City, September 1992. This was published as: "Past Imprecision for Future Standards: Computers and New Roads to Knowledge", *Computers and the History of Art*, London, vol. 4.1, (1993), pp. 17-32.

Cf. the following articles in books: "New Media and New Knowledge", *Proceedings of the Third Canadian Conference on Foundations and Applications of General Science Theory: Universal Knowledge Tools and their Applications*, Ryerson, 5-8 June 1993, Toronto: Ryerson Polytechnic University, 1993, pp. 347-358;"The Electronic Highway and Education: New Doors to keep Open", *Learntec 93, Europäischer Kongress für Bildungstechnologie und betriebliche Bildung, Tagungsband*, ed. Uwe Beck, Winfried Sommer, Berlin: Springer Verlag, pp. 423-434; "Databanks in Education", *The 12th E.C.O.O. and the 8th I.C.T.E. Joint Conference***,** Toronto, (May 1991), pp. 412-418; "Knowledge Packages", *The 12th E.C.O.O. and the 8th I.C.T.E. Joint Conference***,** Toronto, (May 1991), pp. 757-759; "Multidimensional Bibliography and Classification, Eröffnungsvortrag": *Anwendungen in der Klassifikation. Proceedings 8 Jahrestagung der Gesellschaft für Klassifikation eV*, ed. Rolf G. Hensler (Teil I), (Hof Geismar, 1984), (Frankfurt, Indeks Verlag, 1984), pp. 57-75.(*Studien zur Klassifikation*, Bd. 14 SK 14).

See also the following articles in refereed journals: "A System for Universal Media Searching, (SUMS)", *Computers and the History of Art*, London, 1994 (in press); "Conceptual Navigation: Views beyond Windows", *Sistema Terra*, Rome, 1993 (in press); "A Front-end for Multi-valent, Multi-cultural Searching", *ALT News 04, Applied Learning Technologies in Europe*, February 1994, pp. 8-9,14; Guest Editor of the first issue of *Knowledge Organization* (formerly *International Classification*), including an editorial "Computers and the Visual Arts" and an article "Electronic Media and Visual Knowledge", Frankfurt, vol. 20, no. 1, 1993, pp. 2-3, 47-53; "McLuhan, Museums and Education", *Museums and Technology*: special issue of *The Muse*, Ottawa, vol. IX, no. 2, (Summer-Fall, 1991), pp. 78-85; "Computers and a new Philosophy of Knowledge", *International Classification*, Frankfurt, vol. 18, (1991), pp. 2-12.

[4] The eleven pilot projects of G7 are as follows:

1. Global Inventory (of Projects)
2. Global Interoperability
3. Cross Cultural Education and Training
4. Bibliotheca Universalis
5. Multimedia Access to World Cultural Heritage
6. Environment
7. Global Emergency
8. Government On-line
9. Global Healthcare
10. Global Marketplace for SME's
11. Maritime Information Systems.

[5] The particular examples at the ISAD exhibition were chosen because they represented new methodologies in key areas of cultural heritage. The organizers were conscious that there are other excellent examples in all of these key areas. For example, with respect to capturing material there are, of course, the VASARI scanner, the IBM scanner, originally used in the Brandywine project and now being used at the Vatican Library, and the suitcase camera developed by the Italian National Research Council (Consiglio Nazionale delle Ricerche), which is being used specifically for archaeological and museum projects in Naples. A greater spectrum of these projects was organized by the the Italian Ministry of Culture for the Mediartech Exhibition in Florence.

The G7 projects are intended to be complementary. The information from museums (project 5) has many links with the materials in libraries (project 4 directed by France) and relates closely to education (project 3). To make this work on a global basis requires awareness both of various hardware and software (project 1) and interoperability between/among different solutions (project 2 directed by Canada). Hence it is foreseen that the future will almost inevitably bring more co-operation between all five of these initial projects.

[6] It is foreseen that SUMMA will have four interdependent softwares:

SUMC (System for Universal Media Classing)	Classing
SUMS (System for Universal Media Searching)	Searching
SUMO (System for Universal Media Organizing)	Organizing
SUML (System for Universal Media Learning)	Learning

[7] J. Perrault, "Categories and Relators", *International Classification*, Frankfurt, vol. 21, no. 4, 1994, pp. 189-198, especially p. 191.