Documentation: A Reflective Practice Approach

by

Luke Odhiambo Ouko

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of

Master of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

[June 2004]

August 2004

Author ______________________________

Luke Odhiambo Ouko
Program in Media Arts and Sciences
August 06, 2004

Certified by ______________________________

Walter Bender
Senior Research Scientist
Electronic Publishing Group, MIT Media Lab
Thesis Supervisor

Accepted by ______________________________

Andrew B. Lippman
Chair, Departmental Committee on Graduate Students
Program in Media Arts and Sciences

Documentation: A Reflective Practice Approach

by

Luke Odhiambo Ouko

Submitted to the Program in Media Arts and Sciences
on August 06, 2004, in partial fulfillment of the requirements
for the degree of Master of Science

Abstract

The Center for Reflective Community Practice in MIT's Department of Urban studies is involved in projects helping community organizers working on social change. In order to foster reflection, they are currently utilizing what they refer to as "Critical Moments Reflection". This method entails the identification and naming of key shifts or turning points (critical moments) in the process of the community organizers' work. To drive learning through reflection, they use stories relevant to the turning points, they then analyze those moments using a pre-specified genre of poignant questions.

I have created an application, the CMReflector, that aids in the process of Critical Moments Reflection. It will facilitate the process of documentation by utilizing some of the rich computational tools that we now have access to. Since the learning that people acquire through their work stays largely under the surface, there is need to systematically examine the lessons learned and articulate the knowledge and questions that have come out of such work. The application provides an organizational structure and taxonomy around which to compile tacit knowledge and its representation, allowing for exploration of such knowledge in a richer fashion. In order to leverage the use of archived materials, tools such as TalkTV (an application that re-sequeces television content) have been used to augment my application allowing for a "low floor" entry to multiple media editing by the users. It is envisaged that these tools aid in bringing forth the intrinsic "ifs" and "thens," as well as generating the potential for serendipitous learning experiences. All this is very useful in bringing some form of rigor into the practice of reflective inquiry.

Thesis Supervisor: Walter Bender

Title: Senior Research Scientist, Electronic Publishing Group, MIT Media Lab

Thesis Committee

Thesis Supervisor ______________________________

Walter Bender
Senior Research Scientist
Director, Electronic Publishing Group, MIT Media Lab

Thesis Reader ______________________________

David Cavallo
Research Scientist
Director, Learning in the Future Group, MIT Media Lab

Thesis Reader ______________________________

Ceasar McDowell
Professor of Practice of Community Development
Director Center for Reflective Community Practice, MIT Department of Urban Studies and Planning

Acknowledgements

To my thesis readers, my sincere gratitude. I would never have made anyway headway if it were not for your vigilant guidance. Ceasar, thanks for showing me how much there is to knowing. I now appreciate that I have barely scratched the surface. David, you introduced me to the lab, I have learned a lot from you, you epitomize the name "Learning in the Future" . Walter, you have allowed me to have a great time at the lab, I have tried more things over the last two years than I would ever have imagined. Your steadying hand has been invaluable.

Thanks to all the members of the Electronic Publishing Group, Cameron, Joanie, Sunil, Vadim you guys take that collaboration thing pretty seriously. Erik, thanks for TalkTV, it rocks. Nexist thanks for building a really cool engine. Thanks Joy for letting me join your group and share in your reflection sessions. Felice, John, Jacqueline, Pat and Linda, you kept me on the straight and narrow. I am at this point because I did not have to look over my back. The faculty and students at the lab, you make it a joy to be around here, even despite the absence of windows.)-: Cornelle, Kevin and Isis you guys were the icing. NECSYS, and all the medical personnel, you kept me going, especially when I was ready call it quits.

Carla and Larissa, you have been my ángeles. I have no words.

The boys, Chris, Jason, Roger, Paulo, Hugo, Tristan, Victor, James, Davis, David, Martin, Vikas, Bernardo and Ali, did we or did we have a good time during those conveniently extended four day weekends ? You made Cambridge livable.

My brothers (Fred, Evans and Gabriel) and sisters (Jesh and Letty) you guys give me chuny. I wish I could only be a fraction as good to you guys as you have been to me. MekMek, what is there to say? Whatever! My parents, you have given me more than life, you have given me the life. All my successes have been due to your encouragement. The failures have been entirely of my own making (-:

Table of Contents

List of Tables

List of Figures

List of Abbreviations Used

3D	Three Dimensional
ANSI	American National Standards Institute
API	Application Program Interface
ASCII	American Standard Code for Information Interchange
BMU	Best Matching Unit
BNF	Backus-Naur Form
CAD	Computer Aided Design
CADD	Computer-Aided Design and Drafting
CALS	Continuous Acquisition and Life-Cycle Support
CASE	Computer Aided Software Engineering
CMR	Critical Moments Reflection
CRCP	Centre for Reflective Community Practice
DOM	Document Object Model
DTD	Document Type Definition
DTP	desktop publishing
DUSP	Department of Urban Studies and Planning
DVD	Digital Versatile Disc
ETS	Educational Testing Service
GPS	Global Positioning System
GUI	Graphical User Interface
HDL	Hewey Dewey Lewey
HTML	Hypertext Markup Language
ISO	International Organization for Standardization
IEC	International Electrotechnical Commission
JRE	Java Run-Time Environment
LOD	Level of Detail
LVQ	Learning Vector Quantization
MIT	Massachusetts Institute of Technology
SDK	Software Development Kit
SGML	Standard Generalized Markup Language

SOM	Self Organized Map
SVG	Scalable Vector Graphics
TM	Topic Map
URI	Universal Resource Identifier
VE	Virtual Environment
VR	Virtual Reality
VRML	Virtual Reality Modeling Language
W3C	World Wide Web Consortium
WWW	World Wide Web
XHTML	Extensible Hyper-Text Markup Language
XML	Extensible Markup Language

Chapter 1

1.1 Introduction

My hypothesis is that there are ways in which individuals can be aided in formalizing their thinking and learning by the ways in which they document their work, and if the documentation process is aptly structured, it will be compelling enough to foster reflection. Therefore by creating an application that helps practitioners in documenting their Critical Moments Reflection, we can foster experiential learning that is germane to the documentation itself. The application that I built creates a potent taxonomy for structured documentation.

Most of the ways in which we currently document do not allow for a richer expression in the process. A multi-modal approach to the process of documenting may allow for the better utilization of the computational resources available to us, serving to expand the limits for the capture of richer knowledge, especially tacit knowledge. We are able to allow for documentation to be rich in video, audio, photographs, illustrations or all other medium that one may have at their disposal and wish to express oneself in.

The nature of community social justice is that it is characterized by urgency, complexity, constant change, unpredictability and a high degree of personal commitment. People are learning through their experience of the work, but this learning stays largely under the surface, below their awareness. There is usually little time and space for systematically examining what is being learned or articulating the knowledge or questions that have come out of their work. Yet the stakes are high-awareness of the learning that comes through the experience of the community building work is important to responding effectively to the changing conditions of work. [1]

Very little is known about the ways in which individuals develop the feel for media, language and repertoire that shapes their reflection-in-action. This is an intriguing and promising topic for future research. [61]

I have tried to create an application that allows for better documentation using a multi-modal approach to create content that allows for reflection by community organizers.

1.2 Purpose of the Investigation

At the Center for Reflective Community Practice, they organize efforts around surfacing and sharing the knowledge that resides in communities and facilitating the awareness of the learning that emerges from community building practice. Their belief is that greater consciousness and visibility of the learning generated from social change work will allow practitioners to keep their work aligned with deeper democratic and justice principles, and to innovate in the face of enormous complexity and unpredictability. [1]

When a practitioner sees a new situation as some element of his repertoire, he gets a new way of seeing it, and a new possibility for action in it, but the adequacy and utility of his new view must still be discovered in action. Reflection thus involves experiment. [61] The tool I created will allow community organizers to create better artifacts that allow them to reflect both in action and on action. Their end artifacts may eventually be used to enable training of other organizers, motivate others, solicit funding from donors, or for media and public relations activities as well.

1.3 My Thesis and General Approach

I created the CMReflector [Figure 1] as an attempt to meet all the demands that were made on such a tool through a needs assessment with the community organizers. The tool has all the pre-stated requirements as well as a multi-modal approach to using available content. It enables a push on the computational capabilities available on computers, a seamless and intuitive user interface, as well as the ability to create multiple end products (print, video, audio, or mixtures thereof) and still conform to the semantic net to information management and organization.

I explore the usefulness of using rich media to create even richer end artifacts. I delve into pushing the limits of how we document and from what resources we may be able to do so. My application is intended to provide a divergent approach to what we finally come up with us our end products from the Critical Moments Reflection sessions.

My general approach adopts the multiple views to the design of systems. I seek to be first, an expert in the domain for which I seek to create the system, in this case a reflective practitioner, then get more comfortable and competent to iteratively design the system required to resolve the needs of the users.

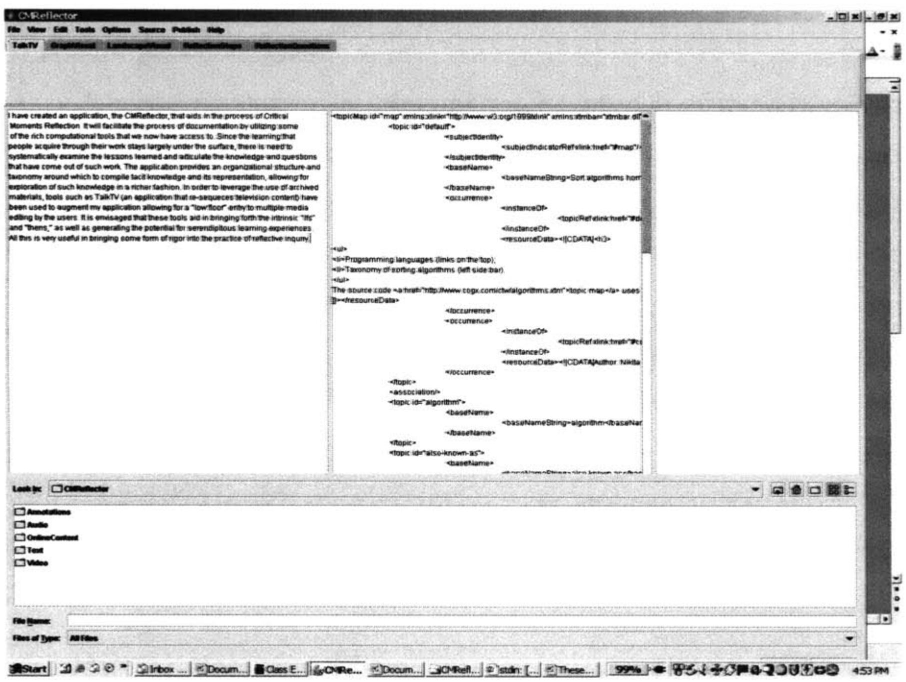

Figure 1: The GUI of the CMReflector

1.3.1 Critical Moments Reflection

A flexible method that has been developed for reflective learning is based on identifying where the key shifts in the work have taken place, and to use the stories of those significant moments to drive the learning in the reflection process. It is referred to as "Critical Moments Reflection", and it has been adopted to fit a number of different kinds of reflective learning processes for people doing social change work. It can be done by individuals, groups and groups of groups. It can be done retrospectively (about past experiences) or in real time (as something is being experienced), in more or less depth, and focused on a single critical moment or a whole set of moments relating to an area of work. [1]

1.3.2 Reflection and Tacit Knowledge

Critical Moments Reflection is geared towards surfacing and sharing knowledge that resides in communities and facilitating awareness of the learning that emerges from community building practice leading to collective transformative learning. [1] A lot of the knowledge that a practitioner possesses sometimes manifests itself tacitly. It is therefore imperative to bring such knowledge to the fore. It is hoped that this not only aids the practitioners in appreciating their hither to latent knowledge but also allows them to exploit it further.

So, reflective practice, though recognizing the value of private reflection, opens up for public scrutiny, our interpretations and evaluations of our plans and actions. We subject our assumptions, be they personal or professional, to the review of others. We do this not only before or after an event, but learn to inquire even in the heat of the moment. [54]

1.3.3 Good documentation

In order to create good documentation there are certain basic characteristics it should have:

Accuracy
Limit your documentation to particular details and make your information about them as accurate as you can.

Completeness
Complete documentation adequately answers each of the six fundamental questions: Who, What, When, Where, Why and How. If a topic is outside the scope of the documentation you are writing, or if someone else has already described that aspect somewhere else, provide an annotated link to the additional information.

Good Organization

Well organized documentation is presented in the order a typical new user will need it as he/she begins to use it. In addition to following a logical progression, well organized documentation adheres to a hierarchical structure that groups related concepts and provides an accessible table of contents.

Good Style -- Clear, Concise, Consistent

Clear writing is the art of creating and arranging distinct thought chunks. Be sure to illustrate important or tricky topics via detailed examples that show the readers exactly how the topic works.

The tool that I created attempted to fulfill these conditions in order to ensure that the end product was of a nature that conformed to the practices of good documentation. [8]

1.3.4 Multi-modal Approach to Documenting

Documents have often been regarded only in the form of “text only” however this has changed and documents are now regarded more as “containers of information or knowledge” irrespective of modality, be it text, graphics, video, sound, etc. The widening concept of document makes it necessary to process multi-modal documents with a unified perspective. [2]

1.3.5 TalkTV

Searching through video for a specific clip isn’t an easy task, and consumer devices such as VCRs only offer rewinding, pausing, and fast forwarding as their main search features. Perpetuating television as a browse only medium, DVDs structure their contents in the form of video chapters within a menu structure. Personal Video Recorders (PVR) such as TiVo also list recorded shows in menus.

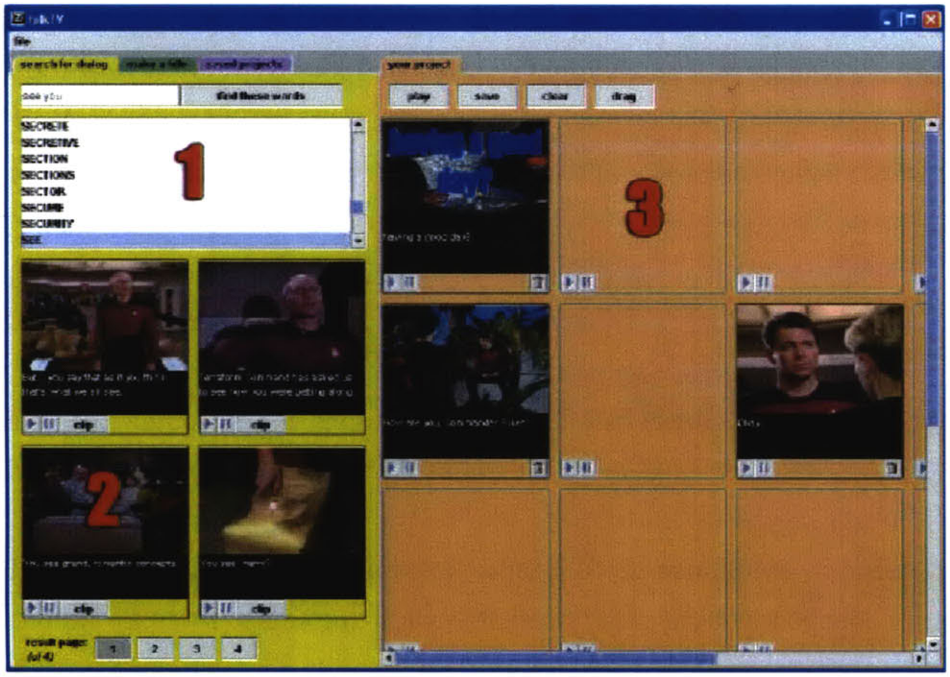

Figure 2: TalkTV User Interface

TalkTV is a graphic editor designed to assist television fans re-sequence television dialog.

With TalkTV one can:

1. Search for dialog
2. Storyboard and allow for re-sequencing of the video content
3. Add titles to the videos
4. Share the end product that is created.

1.3.6 TalkTV as a Documentation Tool

TalkTV may be considered an approachable and compelling means of constructing documents with the basic media elements captured in the previous meetings or during an ongoing meeting. It allows users to appropriate video for personal expression, it thus allows for "view source" of the video content.

TalkTV, video editing system uses closed-caption tracks from television programs to be searched and manipulated into novel presentations. It is hoped that by providing better ways to organize television for creative projects, passive television viewing will be supplemented by constructive television production.

In order to manipulate and leverage the presence of archived video content, TalkTV provides a "low floor" (an approachable tool for novice users of video editing software), easily appropriating the basic media elements and constructing more meaningful content. [7]

1.3.7 Topic Mapping

The ISO committee JTC 1/SC 34/WG 3 Information Technology – Document Description and Processing Languages – Information Association standardized ISO/IEC 13250 Topic Maps (ISO13250) in the autumn of 1999.

The ISO standard defines a model and interchange syntax for Topic Maps. The initial ideas – which date back to the early 1990's – related to the desire to model intelligent electronic indexes in order to be able to merge them automatically. But during several years of gestation, the topic map model has developed into something much more powerful that is no longer restricted to simply modeling indexes.

A topic map annotates and provides organizing principles for large sets of information resources. It builds a structured semantic link network above those resources. The network allows easy and selective navigation to the requested information. Topic maps are the "GPS of the information universe". Searching in a topic map can be compared to searching in knowledge structures. In fact, topic maps are a base technology for knowledge representation and knowledge management.[55]

SGML and XML have DTD defining classes of instances, but topic maps as currently specified do not have an equivalent construct. The standards working group has recognized this need and coined the term topic map template for the "declarative part" of a map.

To illustrate these concepts, let's look at something familiar to all of us: a traditional back-of-the-book index, in this case a review of the book "Practical XML for the Web". Under W3C we find (slightly modified to better illustrate the concepts):

```
W3C, see World Wide Web Consortium
World Wide Web Consortium,
  see also DOM; SVG; XLink.
  DOM specification, 97-98, 232
    Mozilla-based browsers and, 106
```

1.3.7.1 What's in an Index?

Various features can be identified in this example:

Typographical conventions are used to distinguish between different types of topics: XML standards are underlined. Similarly, typographical conventions are used to distinguish between different types of occurrences: references to the term's definition are shown in bold.

Use of *see* references handles synonyms by allowing multiple entry points with different names to the same topic, in our case the acronym W3C.

See also references point to associated topics.

Sub-entries provide an alternative mechanism for pointing out associations between different topics, such as special cases, e.g. browser-specifics.

There could also be different indexes for different concepts, such as XML standards, Web browser software and their versions and the like. Furthermore, occurrences could be distinguished as various types such as definition, quote, reference, etc.

The key features of a typical index are: topics, identified by their names, of which there may be more than one, associations between topics ("see also"), and occurrences of topics, pointed to via locators, such as a page number.

1.3.7.2 Glossary and Thesauri

Other types of indexes are glossary and thesauri. A glossary is an index with only one type of occurrence, namely the definition, which is given inline, not with a locator. A glossary entry could look like:

XML: A format for structured data.

A thesaurus, on the other hand, emphasizes other aspects of an index. The key features of a thesaurus are the relationships or associations between terms. Given a particular term, a thesaurus will indicate which terms have the same meaning, which terms denote a broader or narrower category of the same kind and which are related in some other way. To continue our example, a thesaurus entry might appear as follows:

XML

Definition
    A format for structured data
Broader terms
    Markup Language, structured data format
Narrower terms
    XSLT, XLink, XHTML, ...
Related terms
    SGML, HTML

The special thing about associations in a thesaurus, as compared to associations found in a typical index or glossary, is that they are typed. This is important because it means that the two terms are related and also illustrates how and why they are related. It also means that terms with the same association can be grouped together, making navigation much easier. Commonly used association types like "broader term", "narrower term", "used for" and "related term" are defined in standards for thesauri such as Z39.19, ISO 5964 and ISO 2788.

A related approach to representing the structure of data is semantic networks, usually depicted as conceptual graphs. This thesis could be captured this way:

[Ouko] <- (author) <- [wrote] -> (thesis) -> [MIT-Lib-Ref-MAS2004]
(Square brackets denote concepts, and parentheses describe relations.)

By adding the topic/occurrence axis to the topic/association model, topic maps provide a means of unifying knowledge representation and information management.

1.3.7.3 What is in a topic map?

1.3.7.3.1 Topics

A topic can be any "thing," a person, entity, concept, etc. Topics can be categorized according to their kind. In a topic map, any given topic is an instance of zero or more topic types. This corresponds to the categorization inherent in the use of multiple indexes in a book (index of XML standards, index of Web browser software, etc.). For instance, XML *is a* W3C standard, W3C *is a* committee.

For convenient reference a topic can have one or more names, such as "XML" and "Extensible Markup Language", although unnamed topics are theoretically possible. Sources of different names for the same topic can be the use of acronyms, different languages, or synonyms, which usually come from foreign language influences, whether recently or anciently.

1.3.7.3.2 Occurrences

A topic may be linked to one or more information resources that are deemed to be relevant to the topic in some way. Such resources are called occurrences of the topic. An occurrence could be an article about the topic on a Web site, a picture or video depicting the topic, a simple mention of the topic in the context of something else, or a commentary on the topic, e.g. an XML standard

Such occurrences are generally external to the topic map document itself, and they are "pointed at" using whatever locator mechanisms the system supports, for instance URI in (XTM) XML Topic Maps. Today, most systems for creating hand-crafted indexes (as opposed to full text indexes) use some form of embedded markup in the document to be indexed. One of the advantages to using topic maps is that the documents themselves do not have to be altered.

Occurrences have roles attached to them, like "article", "illustration", "mention", and "commentary" outlined above. Just like topics, roles can have types that codify their nature. For instance the role "article" could be of type "document". In an environment where "document" is well-defined, further information could be derived from this fact, such as asking for a creation or last-modify date.

1.3.7.3.3 Associations

A topic association describes a relationship between two or more topics. The obvious example is the "MIT-Lib-Ref-MAS2004 was written by Ouko" association between a document and its author. Associations, like topics and occurrences, have types, such as "produced by" and "included in". Every topic participating in an association plays a certain role in that association. Naturally, association roles can also be typed, and are also a topic. [32]

1.3.7.4 From Print to Web

Up until now there has been no equivalent of the traditional back-of-book index in the world of electronic information. People have marked up keywords in their word processing documents and used these to generate indexes "automatically", but the resulting indexes have remained as single documents. The World Wide Web removes the distinction between individual documents and now, indexes have to span multiple documents. Indexes have to cover vast pools of information, calling for the ability to merge indexes and to create user-defined views of information. In this situation, old-fashioned indexing techniques are clearly inadequate.

The problem has been recognized for several decades in the realm of document processing, but the methodology used to address it - full text indexing - has only solved part of the problem, as anyone who has used search engines on the Internet knows only too well. Mechanical indexing cannot cope with the fact that the same subject may be referred to by multiple names ("synonyms"), nor that the same name may refer to multiple subjects ("homonyms"). Yet, this is basically how a web search engine works, so it is no surprise when you get thousands of irrelevant hits and still manage to miss the thing you are looking for!

Topic maps provide an approach that marries the best of several worlds, including those of traditional indexing, library science and knowledge representation, with advanced techniques of linking and addressing.

1.4 Criteria for My Studies' Success

In order to check how effectively I will have met the user's needs I will check that all the requirements below have been met:

[1] Provision of a tool that augments existing tools and enhances their nature
[2] Provision of multi-modal capabilities to documenting
[3] Provision of a rigorous taxonomy for the tagging
[4] Provision of content that conforms to the semantic web
[5] Provision of multiple ways to view the information
[6] Provision of possible ways to decipher tacit knowledge from it
[7] Provision of a tool that fosters reflection while being used

1.4.1 The Feature Matrix

To help in the comparison of related tools I constructed a Taxonomic Feature Comparison Matrix, with the following general structure:

Table 1: The Feature Comparison Matrix

Features:	**Tool 1**	**...**	**Tool *N***
Feature Category 1			
Feature Category 1.1			
Feature 1.1.1	yes/no		yes/no
...			
Feature 1.1.*m*	yes/no		yes/no
Feature 1.2	yes/no		yes/no
...			
Feature Category 2:			
...			
Feature Category *n*			
...			

The left column is the taxonomic feature list. The top row lists each tool being compared. The entry for each tool feature is a "yes" or "no". These entries indicate whether or not a particular tool has a particular feature.

1.4.2 Taxonomy

In a software tool comparison, taxonomy can be used to organize the functionality of the tools. For example, we can consider the function categories found typically in the top-level menu bar to be primary candidates for the top-level categories of functionality. Each item in a menu is a subcategory, and items in submenus or dialogs are subsubcategories. It is likely that most tools will have at most four or five levels of command hierarchy, just by the nature of the user interfaces that modern GUI-based tools use.

The focus of the categorization is on the functions that are accessible anywhere in the tool's user interface, whether through menus, buttons, or typing. We specifically do not care about features that are not directly accessible to the user.

Chapter 2

2.1 Theoretical Use Scenario

2.1.1 Use of Topic

While considering Critical Moments Reflection, a topic might represent subjects such as "Race", "Conflict", "Education", "Poverty" or the civil rights activist "Martin Luther King". Anything that might be of interest to reflection can thus be a topic – but also much else besides. Any individual topic is an instance of zero or more topic types. Thus, "Education" would be a topic of type "Issue", Poverty a topic of type "Consequence", while "Martin Luther King" would be of a topic type "Leader", etc. In other words, topic types represent a typical *class-instance* relationship.

Exactly what one chooses to regard as topics in any particular application will vary according to the needs of the application, the nature of the information, and the uses to which the topic map will be put: In, *software documentation* topics would represent functions, variables, objects and methods while in *technical documentation*, components, suppliers, procedures, error conditions, etc.

Topic types are themselves defined as topics by the standard. You must explicitly declare "Issue" or "Leaders", as topics in the topic in order to use them as types (in which case you will be able to say more about them using the topic map model itself).

Key to Figures 2 to 10

The circles represent topics
Difference shapes represent topic types
Different text colors represent different names
Green lines represent occurrences
Dotted green lines represent occurrence roles
Grey lines represent associations

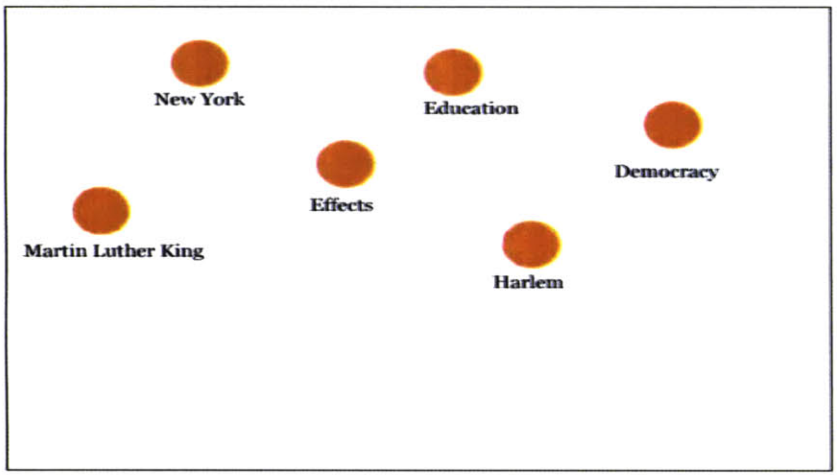

Figure 3: Topics

Normally topics have explicit names, since that makes them easier to talk about. However, topics don't *always* have names: A simple cross reference, such as "see page 97", in an index would be considered to be a link to a topic that has no (explicit) name. Names exist in all shapes and forms: as formal names, symbolic names, nicknames, pet names, everyday names, login names, etc. The topic map standard doesn't attempt to enumerate and cover them all. Instead, it recognizes the need for some forms of names, that have particularly important and universally understood semantics, to be defined in a standardized way (in order for applications to be able to do something meaningful with them), and at the same time the need for complete freedom and extensibility to be able to define application-specific name types.

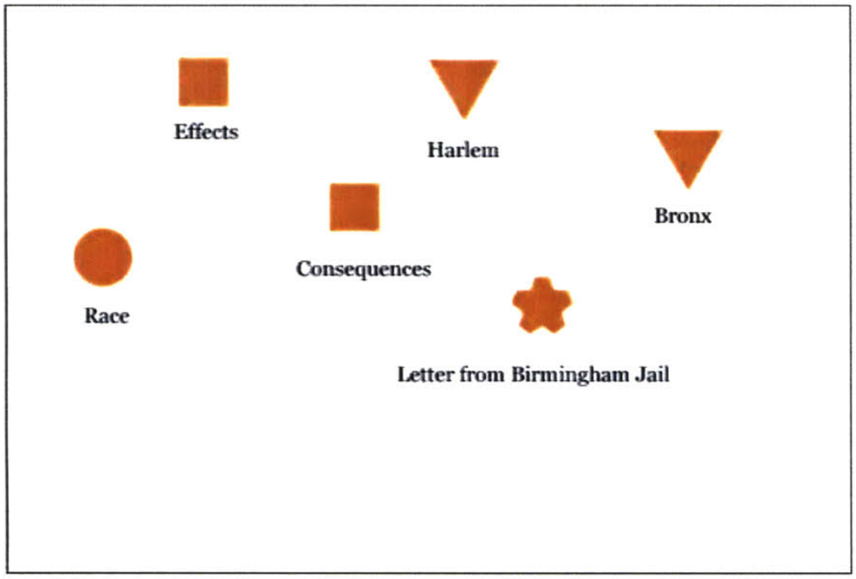

Figure 4: Topic Types

The standard therefore provides an element form for topic name, which it allows to occur zero or more times for any given topic, and to consist of one or more of the following types of names:

• *base name* (required)
• *display name* (optional)
• *sort name* (optional)

The ability to be able to specify more than one topic name can be used to indicate the use of different names in different contexts or *scopes* (about which more later), such as language, style, domain, geographical area, historical period, etc. A corollary of this feature is the *topic naming constraint*, which states that no two subjects can have exactly the same name in the same scope.

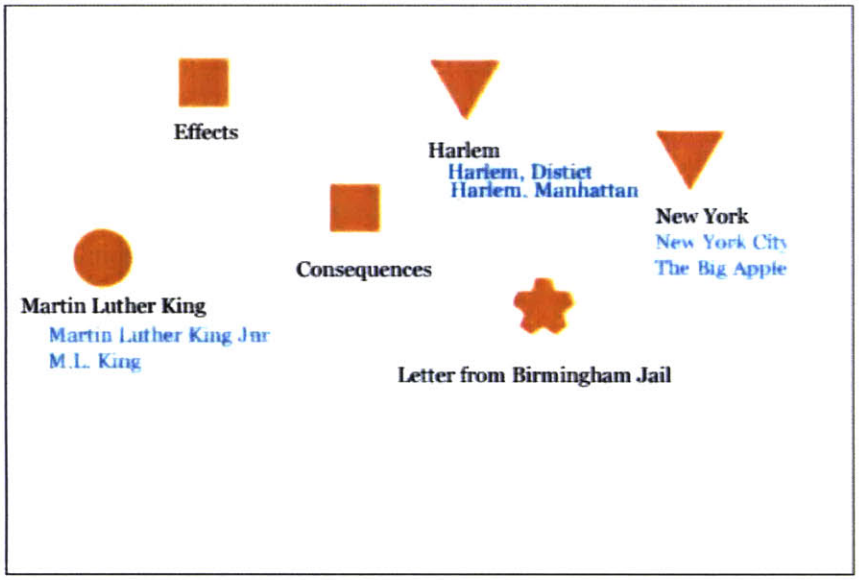

Figure 5: Topic Names

2.1.2 Use of Occurrence and Occurrence Role

A topic may be linked to one or more information resources that are deemed to be relevant to the topic in some way. These resources are the occurrences of the topic. An occurrence could be a monograph devoted to a particular topic, for example, or an article about the topic in a web site; it could be a picture or video depicting the topic, a simple mention of the topic in the context of something else, a commentary on the topic (if the topic were a law, say), or any of a host of other forms in which an information resource might have some relevance to the subject in question. Such occurrences are generally outside the topic map document itself (although some of them could be inside it).

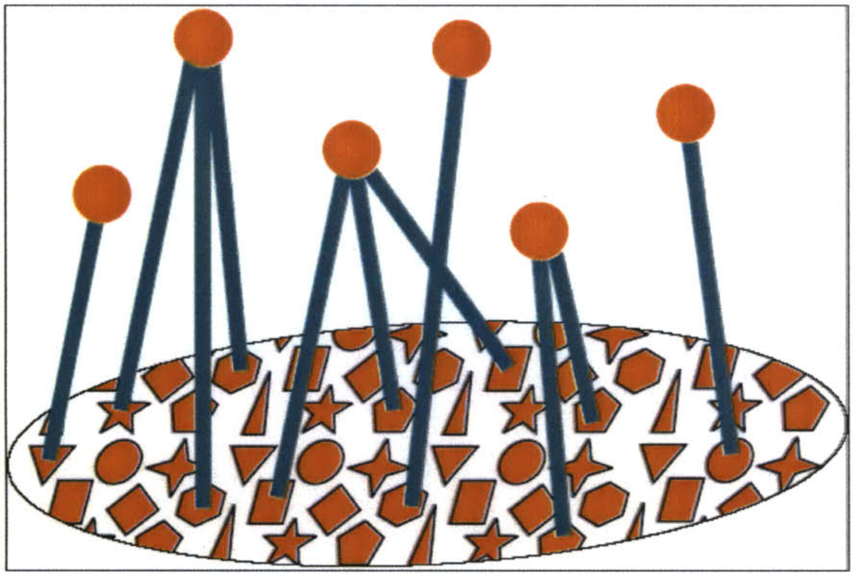

Figure 6: Occurrences

An important point to note here is the *separation into two layers* of the topics and their occurrences. This separation is one of the clues to the power of topic mapping. Occurrences, as we have already seen, a topic may be of any number of different types. Such distinctions are supported in the standard by the concept of the occurrence role.

As with topic types, occurrence roles are really topics and you can therefore use the facilities of topic maps to say useful things about them (such as their names, and the relationships they partake in).

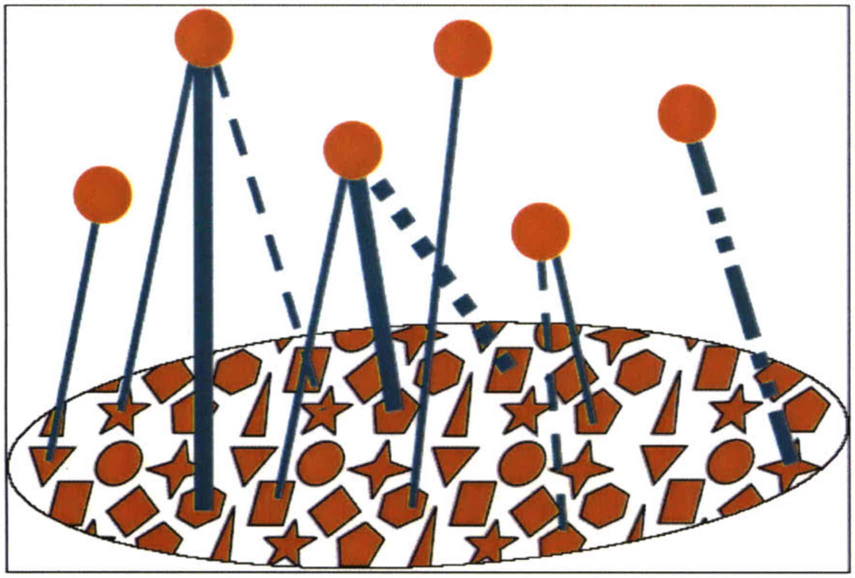

Figure 7: Occurrence Roles

Topic types provide the same facility, but extend it in several directions to enable the creation of multiple, dynamic, user-controlled indexes organized as taxonomic hierarchies. With a topic map it is easy to create and maintain much more complex glossaries than this; for example, ones that use different kinds of definitions (perhaps suited to different kinds of users).

2.1.3 Use of Topic associations and Types

Describing the relationships between topics, through the construct called topic association is the most powerful concept in topic mapping. A topic association is (formally) a link element that asserts a relationship between two or more topics. Examples might be as follows:

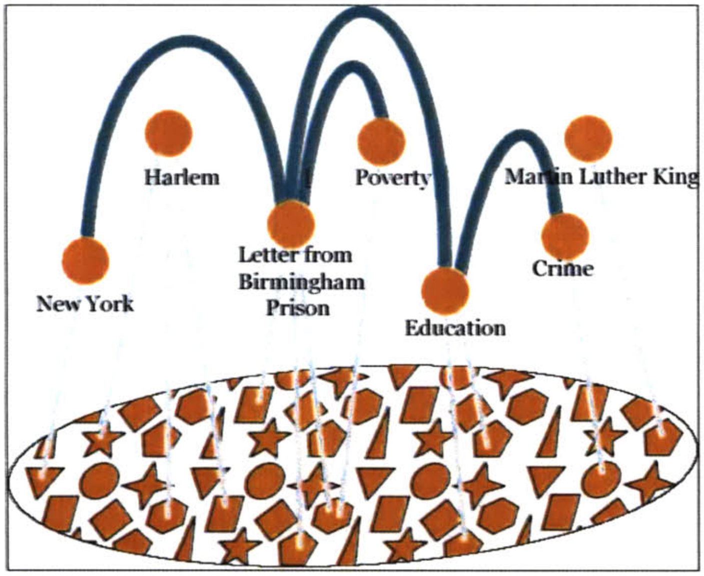

Figure 8: Topic Associations

- "Harlem *is in* New York"
- "Poverty *is a consequence of* Limited Education"
- "Crime is *consequence of* Poverty"
- "Martin Luther King is *author of* "Letter from Birmingham Jail"

Just as topics can be grouped according to type (country, city, poet, etc.) and occurrences according to role (mention, article, commentary, etc.), so too can associations between topics be grouped according to their type. The *association types* for the relationships mentioned above are *is-in* (or geographical containment), *is-a-consequence-of* and is *author-of*. As with most other constructs in the topic map standard, association types are themselves defined in terms of topics.

The ability to do typing of topic associations greatly increases the expressive power of the topic map, making it possible to group together the set of topics that have the same relationship to any given topic. This is of great importance in providing intuitive and user-friendly interfaces for navigating large pools of information.

It should be noted that topic types are regarded as a special (i.e. syntactically privileged) kind of association type; the semantics of a topic having a type (for example, of Education being an Issue) could quite easily be expressed through an association (*of type* *instance-of*) between the topic "Education" and the topic "Issue". The reason for having a special construct for this kind of association is the same as the reason for having special constructs for certain kinds of names.

While both topic associations and normal cross references are hyperlinks, they are very different creatures: In a cross reference, the anchors (or end points) of the hyperlink occur *within the information resources* (although the link itself might be outside them); with topic associations, we are talking about links (between topics) that are *completely independent* of whatever information resources may or may not exist or be considered as occurrences of those topics.

2.1.3.1 Why are Hyperlinks so Important?

Because it means that topic maps are information assets in their own right, irrespective of whether they are actually connected to any information resources or not. The knowledge that "Crime is a consequence of Poverty", that "Letter from Birmingham Jail was written by Martin Luther King" is useful and valuable, whether or not we have information resources that actually pertain to any of these topics. Also, because of the separation between the information resources and the topic map, the same topic map can be overlaid on different pools of information, just as different topic maps can be overlaid on the same pool of information to provide different "views" to different users.

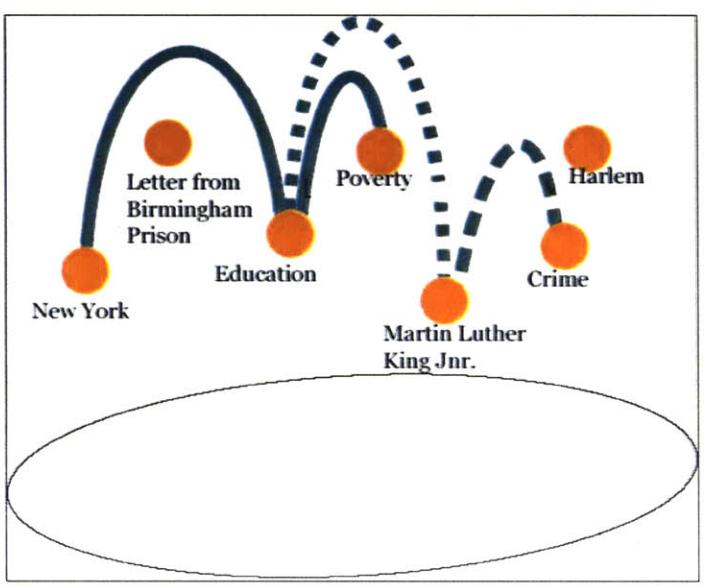

Figure 9: Topic Maps as Portable Semantic Networks

Furthermore, this separation provides the potential to be able to interchange topic maps among users and thus to merge one or more topic maps. Other association types, such as those that express superclass/subclass and some part/whole (meronymy/holonymy) relationships, are transitive: If we say that poverty is a consequence of limited education, and that crime is a consequence of poverty, we have implicitly said that crime is a consequence of poor education. Adding the topic/occurrence axis provides a means for "bridging the gap" between knowledge representation and the field of information management. Adopted from [52].

2.2 Reflection-On-Action Scenario

Dr. Nancy Stewart is a community organizer and a reflective practitioner. She is the Executive Director of a small non-profit that is involved in advancing the education needs of at risk youth in inner city communities around the Boston area. She has recently participated in the "Critical Moments Reflection" sessions at MIT Center for Reflective Community Practice to aid in her in (1) identifying what she has learned about how to improve education by working outside the system, and (2) how to apply that learning to resolve important organizational, logistic and managerial issues in her organization.

2.2.1 Availability of Content

Dr. Stewart has access to all the resources that were availed during the reflection sessions. She wants to revisit the material and create content that is pertinent to her specific needs at her organization. Since she has access to the organizations intranet as well as her notes from the sessions she begins the process of creating her content. She thus has an opportunity to document in multiple fashions utilizing the rich media that is available to her.

She has access to:

1. Transcribed Audio and Video content
2. Transcriptions in text format
3. Related web sites and URLs
4. Personal notes

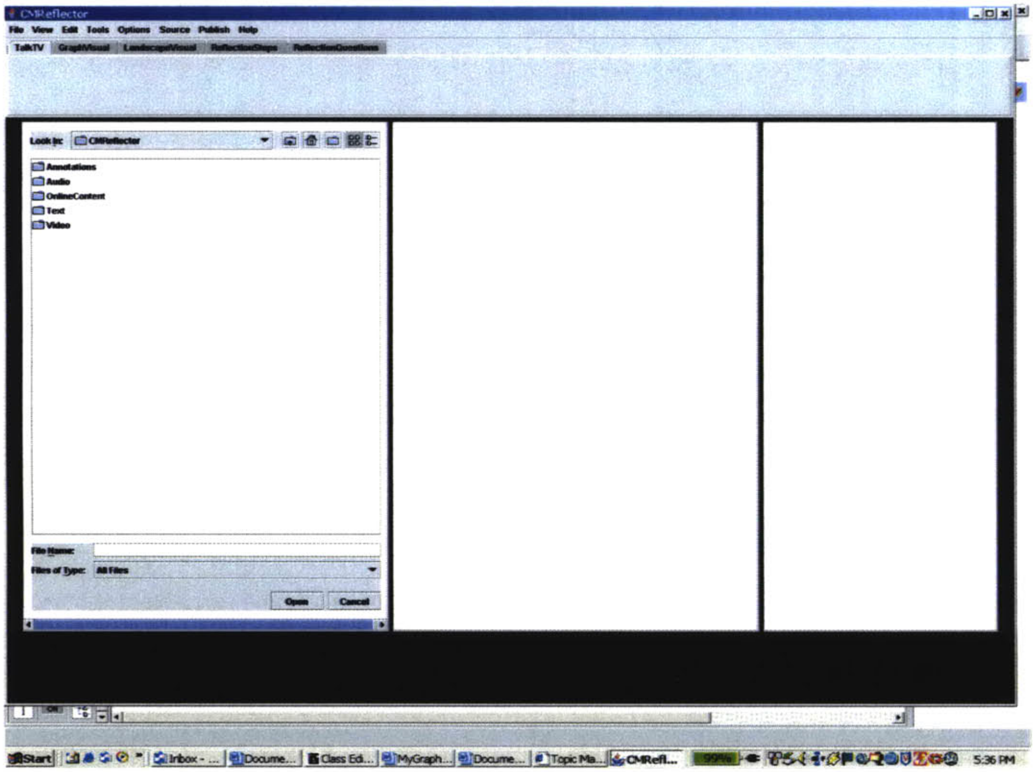

Figure 10: Resources for Content Generation

The content is in multiple medium. She can thus use audio, video, text, pictures, or even personal handwritten notes.

2.2.2 What are the Critical Questions?

She goes through the critical questions [Figure 11] that are asked during reflection. The application aids her in doing this. Among the questions that participants asked themselves are:

1. What were some of the really significant moments around a particular event?
2. What was the most significant moment of the event?
3. What are the different perspectives as regards this moment?
4. What is the analysis obtained from this moment?

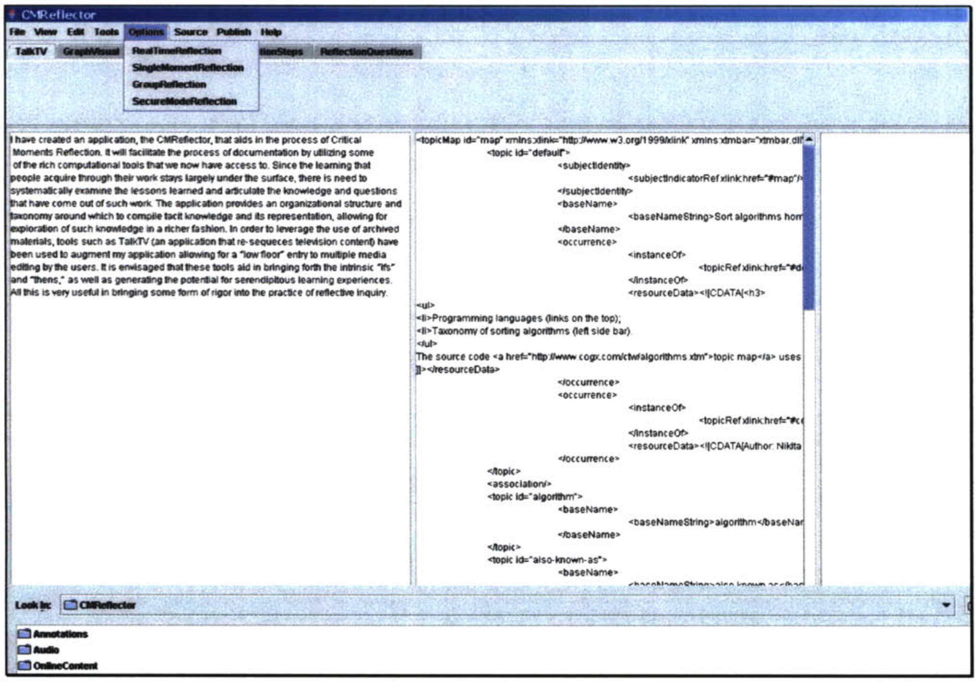

Figure 11: Menu of General Critical Moments Questions

In further analyzing the critical moment further questions are asked. These may include questions such as:

- Why are particular moments so significant?
- What is the nature of their significance?
- What challenges are better understood?
- What are the important deeper questions?

All these questions [Figure11 & Figure 12] have been previously pre-tagged and thus conform to the topic map paradigm. Dr. Stewart need not learn any form of syntax in order to tag appropriately. She can also opt to tag the content in a manner that is suitable or most comfortable for her.

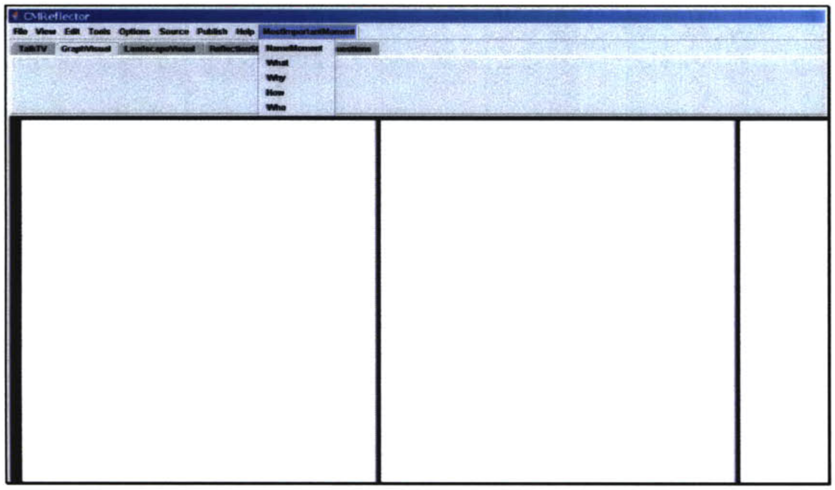

Figure 12: Menu of Detailed Reflection Questions

2.2.3 Tagging Content on the Different Resources

She accesses a video of one of the participants talking about his problems with dealing with the education board in his district. In this video he describes how he went about convincing them to set up a committee within the board to look into drop-out rates for teenage girls as a result of pregnancy.

She also takes a look at the dialogue that occurred earlier in the reflection session between a parent and a high school student that managed to complete school though she was pregnant.

She then takes a better look at something she thought she heard someone say but is not sure who. She therefore tries to search the content to find out who said it and specifics of what was said.

At this point in time she is creating her own tags (personal tagging) to the content. She can then share with others who have used topic mapping. The tags that are created will help her in finding content as well as in sharing her approach or view with others that use the shared system or repository.

All the content that exists in the video is annotated. [Figure 13] This content can be searched from the index or through a "word search" and the content (clip with the audio) will be displayed on the bottom left as "search results" The user can then collect all the clips that they need and upload them on the right hand side of the screen ready to create a personalized movie out of it.

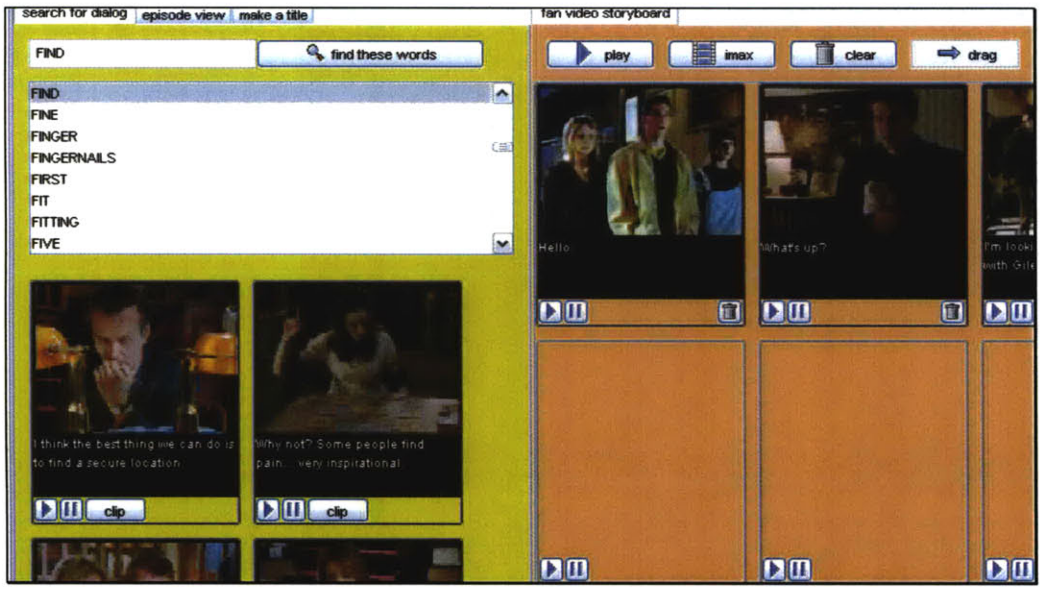

Figure 13: Searching for Specific Content on Archived Video Resources

A further iteration of TalkTV allows for one to create their personal tags for the content in whatever way they choose. A user could for instance tag the content in terms "very important", "issues to follow up on" etc. They will then be able to search the content using their "Self-defined tags".

2.2.4 Topic Mapping in the Background

All the content conforms to the topic mapping ISO standard. And there exists a direct mapping of all the content that is tagged.

Table 2: Direct Mapping for Topic Mapping

Item	Direct Mapping To
Issue	Topic
Relevance	Association
Location	Occurrence
Perspective	Scope
Description	Association
Contributor	Association

We also notice that in as much as items or elements may be associations or scopes they are still endowed with a topic name.

In Figure 14, we see that the user of the system has the option to choose how they view the content that they are manipulating. They may choose to see the XTM code and tinker with it. They could view the code (in the middle), the content in a database format (on the extreme right in blue) or as they have typed it in the text entry area (on the left).

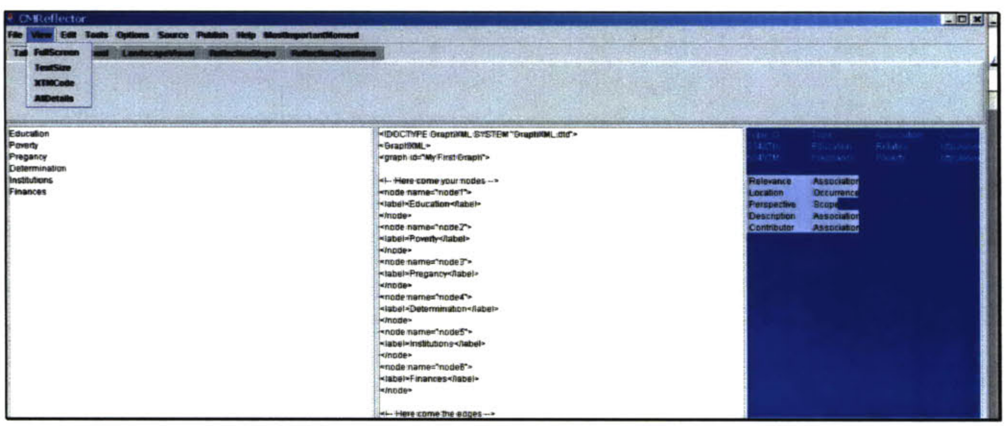

Figure 14: Knowledge Representation Topic Map Style. Key words on the left, XTM tagging at center and database table on the right.

2.2.5 Viewing the Content

Having input and tagged all the content Dr. Stewart now wants to see where the relationships or inferences that she may not immediately have known are. These are those created by her as well as other participants of the reflection sessions.

Something that she may never have picked up can become obvious when she looks at the landscape visualization of the content with the Self organizing Map algorithm applied to it.

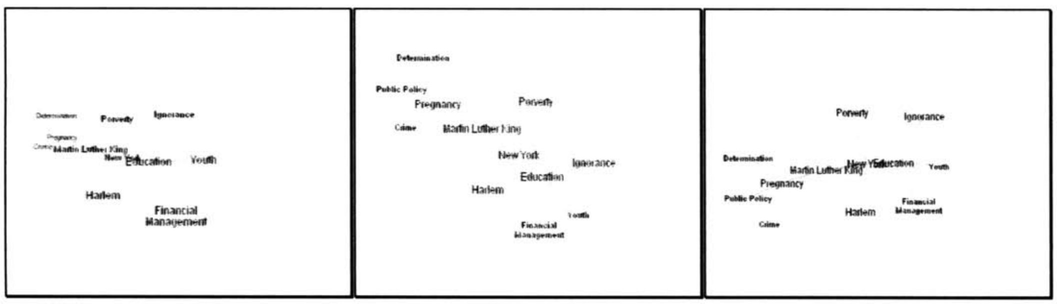

Figure 15: Series of Graphs Generated from the tagged content

Depending on what she wants to concentrate on she can decide which part of the graph to look at. By clicking on an element it becomes the center of the graph and thus illuminating the relationships with other cells in the graph. By double clicking on an element it opens the resource linked to it. Content could thus be linked to web sites, video or audio content or paths to a server.

She realizes that she needs to have a much larger perspective of what is going on. She thus uses the bird's eye view to visualize the content that she has tagged as well as that tagged by other participants. She thus opts for the landscape visualization. [Figure 16] This shows her from an even larger perspective the inter-relatedness of her content given the specific tagging applied to it.

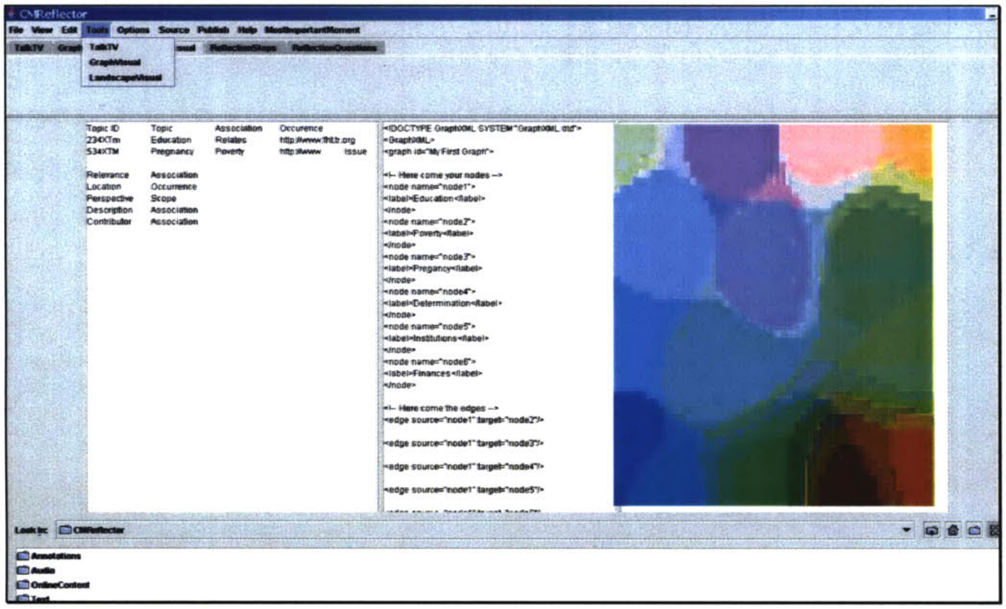

Figure 16: Landscape Visualization of the Tagged Content. On the left is the user's notes, in the middle is the XTM code and on the right is the landscape visualization.

2.2.6 Final Artifact

Since she needs to create the content for different users she may call up the scope element and finds out the needs of the divergent users that she needs to address. She compiles all these into one file and readies it for the final content artifact [Figure 17].

This content may end up in the form of:

- Print
- Create XTM
- Create Final Video
- Create Text Document
- Create Audio Content

And this content may be used for

- Training
- Public Relations
- Motivation

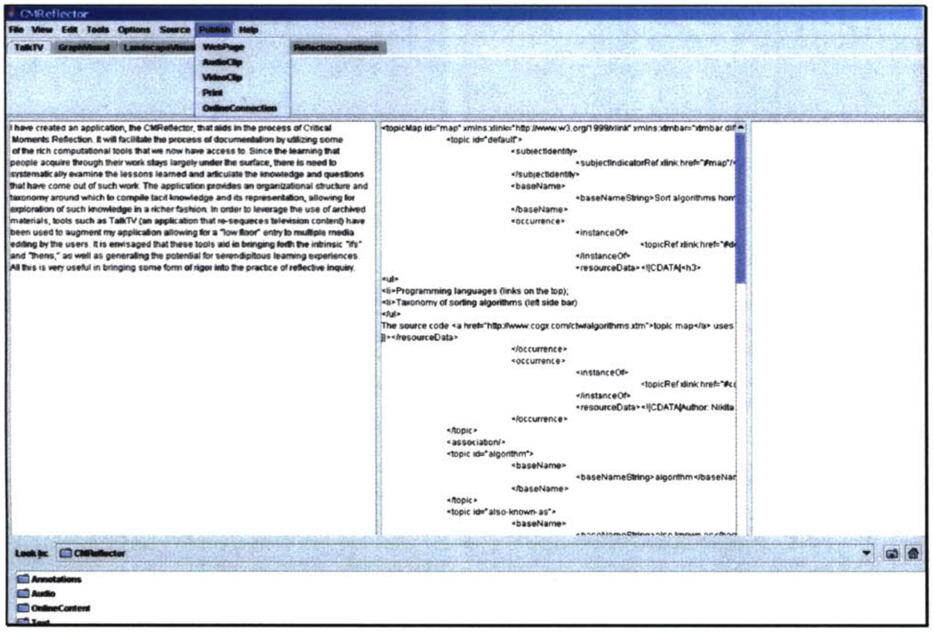

Figure 17: Final Artifacts with all the Publishing Options.

Chapter 3

3.1 Documentation Tools

Different documentation tools are developed to meet the very different needs of their potential users as is illustrated in [Table 3]

Table 3: Documentation Tools

Documentation Category	Application Examples
Color Publishing	Alias, Full Color Publisher
Database Publishing	BASISplus
Desktop Publisher (DTP)	FrameMaker, Interleaf
Document Management	Metamorph, RDM
Editor	Qedit, Microstar
Electronic Distribution	Acrobat, Replica
Extract from Source Code	DocGen
Filter/Translator	Filtrix, SGML Translation Products
Flowcharting	EasyFlow, DiagramMaker, Visio
Graphics/Presentation	CorelDRAW, Power Point
Forms	PerForm, WordPerfect InForms
Hypermedia	HyperCard, AnyImage
Integrated Office Automation	Microsoft Works, Framework, Enable
Page Layout	Adobe Illustrator, Archetype Designer
Scanning	Omnipage
SGML-Based/CALS	FastTag, Omnimark
Technical Drawing	Generic CADD, Illustrator, MapCon
Text Processor/Batch Compiler	, Documenter's Workbench
Word Processor	Word, WordPerfect, Ami Pro

While some elements are similar in terms of what documentation tools do in general, for a lot of them their orientation is one of meeting highly customized requirements. [68]

3.2 Why Use Multi-modal Media?

Among the reasons given for using multi-modal media include:

- Convenience (mobility and speed)
  The scanned documents can thus be used with ease. Users do not have to learn the ontology in order to use the system.
- Minimize errors
  Since all the content is referred to directly there is never need to make assumptions or interpret because everything is available on record. By aggregating the content into a shared repository mistakes can be noticed easily.
- More expressive powers
  There are more avenues in which one can choose to express themselves. One chooses the one that appeals most to them. Since there are visualizations, users are able to view the content from a unique perspective.
- More flexibility
  The possibility to tag as pre-set or create new tags, the use of multiple forms of media, divergent views of the content, all these serve to allow for multiple intelligences to be used. [24]

3.3 Why Use Topic Mapping?

A topic map annotates and provides organizing principles for large sets of information resources. It builds a structured semantic link network above those resources. The network allows easy and selective navigation to the requested information.

3.3.1 How are Topic Maps Relevant to XML?

XML is making semantic markup the lingua franca of the Web. HTML has been the first choice language for the Web as far as presentation and one-way hypertext links are concerned, but XML is now meeting requirements that HTML could never meet. Unlike

HTML, XML provides a way to use markup to associate arbitrary semantics with arbitrary chunks of information. By using appropriate algorithms and a modicum of good sense, creators of topic maps can significantly improve their productivity by leveraging such semantic markup. They can more readily incorporate additional resources into the base of resources for which a topic map provides the basis of a set of finding aids, and they can more easily identify topics, topic types, topic relationships, topic relationship types, occurrences, occurrence types, names and scopes that should be added to their topic maps.

3.3.1.1 Topic Maps extend the power of XML to make information self-describing

Even in the coming XML-dominated Web, topic maps are important because persons other than the authors or owners of arbitrary chunks of XML information often have need to associate their own arbitrary semantics with such chunks, even though they do not have the authority to write on them or to contribute to their inherent semantic markup. For a variety of technical, economic and political reasons, standard topic map documents are an ideal way to allow anyone to contribute what, in effect, amounts to added or alternative semantic markup to any component(s) of any set of information resources, without having to change those resources in any way. With pure XML, we can have only a resource author's views (and only the author's purely hierarchical views) as to the semantics of each information component. With additional topic maps, we can take advantage of different perspectives on the same information. The topic maps paradigm dramatizes the fact that the distinction between data and metadata (data about the data) is purely a matter of perspective. Topic map documents reflect the perspectives of their authors on all the information components they address; they define semantically customized views. [46]

3.3.1.2 Topic maps are XML documents

Topic map documents can already be expressed as XML documents, in complete conformance with the W3C XML 1.0 Recommendation and with the syntax specified by the ISO Topic Maps standard. To express a topic map document in XML, one need only respect the syntactic constraints imposed by XML, in addition to those imposed by SGML. (For example, unlike XML, SGML permits the omission of redundant markup, such as implicitly unnecessary end-tags for elements.) None of the additional syntactic constraints imposed by the syntax of XML interferes with the expressibility of topic maps; it is straightforward to express topic maps in XML.

3.3.2 Topic Maps exploitation in Mass Markets

The ISO Topic Maps standard, like the other members of the ISO SGML family of standards, is designed for extreme generality and flexibility, and for comprehensive utility in all systems contexts. This level of generality is highly desirable in an international standard for system-independent, vendor-independent, and application-independent information interchange, but it introduces complexities that may be irrelevant to the majority of applications.

In recent years, considerable marketing effort has been directed at making SGML exploitable in mass markets. These marketing efforts have given SGML a new name ("XML"), a simplified syntax, and an expectation that XML software and systems will cost much less (while being almost as powerful) than comparable SGML software and systems. Similarly, the draft W3C XLink Recommendation embodies a subset of the concepts and syntax of HyTime's linking and addressing facilities. XML and XLink can afford to be much more limited in their scope and flexibility partly because they have been designed for exactly one information delivery system: the World Wide Web. It seems reasonable to assume that, if XML and XLink are viable mass-market standards, then some similar syntactic and/or functional subset of the Topic Maps standard should also be a viable mass-market standard. Only the worldwide information management industry can determine exactly how to do that, by means of some consensus-seeking process. The process is intended to culminate in the publication of a standard called the "XTM Specification."

The general mission of TopicMaps.Org is to "engage in technical, educational and marketing activities to facilitate the use of topic maps based on XML, including but not limited to application on the Web." The XTM Specification will be "an application of the ISO/IEC 13250:2000 Topic Maps standard that trades some of the generality of the international standard in exchange for a simpler, more predefined and more immediate way to exploit topic maps in mass market applications. One aspect of the task is to develop industry-wide consensus about how topic maps should be expressed in XML, and how the Topic Maps paradigm can be most readily exploited in Web-based applications.

XTM topic maps should retain the ability to be merged with other kinds of topic maps, in accordance with ISO/IEC 13250:2000. [46]

3.3.3 Types of Maps Possible

I use "Template Maps" and "Self-Defined Maps" that adapt to the Topic Map principle to aid in this endeavor. Template maps are pre-determined structures for how the map will look like while the "Self Defined" ones are created by the user of application. They make the decision over how all the elements of the map are related.

I took the topic map standard and created an engine that allows for it's manipulation by simply requiring users learning the most elementary concepts of the paradigm. Once an appreciation of the underlying concepts of topic, association, occurrence and role is gained, the user can create sophisticated topic maps. The XTM code is available for those that would like to further "tinker" with it.

3.4 Reflection "In action" and "On action"

Schön's great contribution has been to bring the notion into the centre of any understanding of what professionals through the ideas of reflection in and on action. In the case of the former, the practitioner allows himself to experience surprise, puzzlement, or confusion in a situation which he finds uncertain or unique. He reflects on the phenomenon before him, and on the prior understandings which have been implicit in his behavior. He carries out an experiment which serves to generate both a new understanding of the phenomenon and a change in the situation. [61]

To do this we do not closely follow established ideas and techniques - textbook schemes. We have to think things through, for every case is unique. However, we can draw on what has gone before.

We can link this with reflection-on-action. This is done later. Workers may write up recordings, talk things through with a supervisor and so on. The act of reflecting-on-action enables us to spend time exploring why we acted as we did, what was happening in a group and so on. In so doing we develop sets of questions and ideas about our activities and practice.

This distinction has been the subject of some debate (see [22] and [69]). Indeed he may well fail to clarify what is involved in the reflective process - and there is a problem, according to Eraut, around time - 'when time is extremely short, decisions have to be rapid and the scope for reflection is extremely limited' [22]. There have been no psychological elaborations of the psychological realities of reflection in action. [57]

3.5 The Steps in Critical Moments Reflection

It starts by asking the individuals to name those moments that they experienced as significant within a particular area of their work. If these individuals are doing critical moments reflection as members of an organization, the contrast in the set of moments that are experienced as most significant from different perspectives in the work is often very interesting and can provoke a powerful learning discussion

The next step is to narrow down to a subset of critical moments, to choose one or more to get into in more depth. This decision is usually based on which moment(s) will likely produce the learning that is most important to those involved (for example, they may have thought beforehand about questions or issues they hope to learn about and choose the moments that hold learning for those questions).

For the selected critical moment(s), the story of the moment is told (sometimes from more than one perspective if done in a group), and then there is a collective analysis of the moment- what shifted, why, what led to the moment, what happened as a result. It is through the analysis that the meaning of the moment to the person(s) who experienced it is brought out and explored. The goal of the analysis is to become aware of why particular moments are experienced as significant and to examine the nature of the significance, allowing the challenges in the work to be better understood, and often revealing important deeper questions. [1]

3.6 Ontology

Ontology is an explicit specification of a conceptualization. The term is borrowed from philosophy, where ontology is a systematic account of Existence. For AI systems, what "exists" is that which can be represented. When the knowledge of a domain is represented in a declarative formalism, the set of objects that can be represented is called the universe of discourse. This set of objects, and the describable relationships among them, are reflected in the representational vocabulary with which a knowledge-based program represents knowledge. Thus, in the context of AI, we can describe the ontology of a program by defining a set of representational terms. In such ontology, definitions associate the names of entities in the universe of discourse (e.g., classes, relations, functions, or other objects) with human-readable text describing what the names mean, and formal axioms that constrain the interpretation and well-formed use of these terms. Formally, ontology is the statement of a logical theory.

The word "ontology" seems to generate a lot of controversy in discussions about AI. It has a long history in philosophy, in which it refers to the subject of existence. It is also often confused with epistemology, which is about knowledge and knowing.

Ontology is a description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or a community of agents. This definition is consistent with the usage of ontology as set-of-concept-definitions, but more general. And it is certainly a different sense of the word than its use in philosophy. [26]

I have tried to formalize the reflection process by adopting the sometimes formal and rigorous and at times loose and fluid form into a pre-formatted tagging system. While it forces or fosters some formalization, it also allows for creativity and innovation on the part of the user. I have attempted to create a focused ontology solely for Critical Moments Reflection.

3.7 Semantic Web

The Semantic Web is a mesh of information linked up in such a way as to be easily processable by machines, on a global scale. You can think of it as being an efficient way of representing data on the World Wide Web, or as a globally linked database.

The Semantic Web was thought up by Tim Berners-Lee, inventor of the WWW, URI, HTTP, and HTML. There is a dedicated team of people at the World Wide Web consortium (W3C) at MIT working to improve, extend and standardize the system, and many languages, publications, tools and so on have already been developed. However, Semantic Web technologies are still very much in their infancies, and although the future of the project in general appears to be bright, there seems to be little consensus about the likely direction and characteristics of the early Semantic Web. [48]

3.7.1 What's the Rationale for Such a System?

Data that is generally hidden away in HTML files is often useful in some contexts, but not in others. The problem with the majority of data on the Web that is in this form at the moment is that it is difficult to use on a large scale, because there is no global system for publishing data in such a way as it can be easily processed by anyone. For example, just think of information about local sports events, weather information, plane times, Major League Baseball statistics, and television guides... all of this information is presented by numerous sites, but all in HTML. The problem with that is that, in some contexts, it is difficult to use this data in the ways that one might want to do so.

So the Semantic Web can be seen as a huge engineering solution... but it is more than that. We will find that as it becomes easier to publish data in a re-purposable form, so more people will want to publish data, and there will be a knock-on or domino effect. We may find that a large number of Semantic Web applications can be used for a variety of different tasks, increasing the modularity of applications on the Web. [13]

By the creation of the semantic markers during the documentation process, the end documents become more amenable to be used in a repository in a knowledge base. They could for instance be used to derive common-sense knowledge for applications relating to the domain area.

Since all the content ends up being tagged in XTM and by extension XML, it lends itself to conforming to the semantic web and therefore optimizing the searching (information retrieval) and or browsing capabilities of content in the future.

3.8 What is Knowledge Representation?

Randall Davis of the AI Lab at MIT argues in [20], that the notion can best be understood in terms of five distinct roles it plays, each crucial to the task at hand: Knowledge representation (KR) is most fundamentally a surrogate, a substitute for the thing itself, used to enable an entity to determine consequences by thinking rather than acting, i.e., by reasoning about the world rather than taking action in it.

It is a set of ontological commitments, i.e., an answer to the question: In what terms should I think about the world?

It is a fragmentary theory of intelligent reasoning, expressed in terms of three components:

(i) the representation's fundamental conception of intelligent reasoning;

(ii) the set of inferences the representation sanctions; and

(iii) the set of inferences it recommends.

It is a medium for pragmatically efficient computation, i.e., the computational environment in which thinking is accomplished. One contribution to this pragmatic efficiency is supplied by the guidance a representation provides for organizing information so as to facilitate making the recommended inferences.

It is a medium of human expression, i.e., a language in which we say things about the world.

3.9 Information and Knowledge Management

Knowledge Management (KM) according to Gregory Wenig is described: As the word implies, the ability to manage "knowledge". We are all familiar with the term Information Management. This term came about when people realized that information is a resource that can and needs to be managed to be useful in an organization. From this, the ideas of Information Analysis and Information Planning came about. Organizations are now starting to look at "knowledge" as a resource as well. This means that we need ways for managing the knowledge in an organization. We can use techniques and methods that were developed as part of Knowledge Technology to analyze the knowledge sources in an organization. Using these techniques we can perform Knowledge Analysis and Knowledge Planning.

3.10 Tacit knowledge and Implicit Learning

Tacit (silent) knowledge [53] and implicit learning [6] have in common the idea of not knowing what you do know or have learned. "Tacit knowledge" has been all but hi-jacked by management gurus, who use it to refer to the stock of expertise within an organization which is not written down or even formally expressed, but may nevertheless be essential to its effective operation.

All that one can do to "teach" such knowledge is to provide opportunities for people to learn it, perhaps through exposure to examples: there is nevertheless a qualitative leap between individual examples of either propositional knowledge or practical skill, and the ability to integrate them into something which is "more than the sum of its parts". The distinction is similar to that between Bateson's Learning I and Learning II. Ideas of situational learning similarly suggest that it is most effectively "picked up" in real world situations, and that attempts to reduce it to standard teachable forms are in danger of distorting or destroying it. [3]

There is a need to create annotation systems that capture the various metaphors that are unique to each domain knowledge area. There are obviously many ways in which certain concepts are understood in different cultures, locations, languages, professions or ages.

It would therefore be very useful to create a mechanism that would put into a repository the nuanced metaphors and match them against each other in a database.

3.11 Searching and Browsing Information Repositories

What is it that makes searching and browsing repositories easy or hard, and what is it that I am doing in this new application that makes this process easy? How do search engines work and how do topic maps allow for better searching?

The growth of the World Wide Web has turned the Internet into a huge repository of information with an estimate of two to ten billions of Web pages, depending on the source. However, efficient access to this information becomes more and more difficult. Obviously, this immense information space becomes virtually worthless unless the information on it can be efficiently located and retrieved. This is why the understanding of information searching processes is a highly relevant research issue.

The need for a systematic study about searching information arises from the discrepancy between two views. The first of those is based on the perception of the Internet as an ultimate source of information for learning and for every-day purposes. This perception assumes that everyone can easily and naturally access specific information on the Web. In the second view, effective information searching, being a complex skill, is far from easily accomplished. This view is supported by current research that reports on the "Web rage" phenomenon. According to this survey, 71% of Internet users get frustrated in the process of searching the Internet for specific information. [56]

Clearly, we must improve our understanding of the skills required for information searching on the Web, and of the processes involved. Beyond affecting our behavior as individual Web users, such understanding might affect pedagogical decisions regarding the usage of information in education systems.

3.12 Information Visualization

Information visualization is an emerging means to make sense of vast amount of data. It is the use of computer-supported interactive visual representations of abstract data to aid understanding and analysis. Many works on visualization refer to the representation being graphical, i.e. visual and non-textual. The graphical presentation may entail manipulation of graphical entities (points, lines, shapes) and attributes (color, size, position); Text strings are often used in visualizations as annotations, or labeling.

Information visualization is distinguished from scientific visualization in the type of data sets the fields refer to, and the graphical form derived. Scientific visualization deals with data collected or measured from physical phenomena, or from the simulations of physical phenomena. For example, data from weather forecasts, mechanical structures, magnetic resonance scans of the human body and cosmology. As such, in scientific visualization the resultant visual representations are usually based on the geometric component inherent within the data set. In contrast, information visualization generally deals with abstract data that does not have intentional spatial components, such as data contained within a vast majority of databases: financial data, documents, hypermedia, the World Wide Web (WWW) and file directories. Such data presents additional freedom to visualization researchers on the choice of effective visual forms. However, it is also an additional challenge with no obvious starting point.

Research has shown that visualization has two major benefits:

1) It provides visual abstraction to speed pattern detection; and
2) It communicates large amount of information effectively. [43]

3.13 Abilities Relevant to Navigation

In information structures, in particular, it includes spatial ability, associative memory, and visual memory. Spatial ability in information exploration has received significant research attention. [16] cites several studies confirming the importance of this ability; one such study was that by ([12]. In these cited studies, spatial ability has generally been found to correlate positively with task performance in information exploration tasks.

[17] and [16] conducted empirical studies of the relationship between individual differences and information exploration in a virtual world organized on the basis of visualized semantic relationships. As a general goal, the studies sought to assess the usability of a prototype VE, and to examine methodological issues for development. In so doing, the studies examined the search strategies and general preferences of users. Three cognitive abilities mentioned above – spatial ability, as well as associative memory and visual memory – were central to the studies. For assessing individual ability, the researchers used the ETS kit.

Results showed that spatial ability was positively correlated with recall (number of relevant abstracts found, as a share of the total number of relevant abstracts available), but negatively correlated with precision (number of relevant abstracts found, as a share of the total number of abstracts found). Spatial ability was also positively correlated with the number of abstracts judged relevant by each participant. These results suggest a connection between spatial ability and ability to utilize the structure of visualization. Participants' world sketches varied substantially, whose accuracy apparently reflected individual differences in spatial ability.

Results also showed that subjects retrieved more abstracts that were located near to structural joints in the visualization, relative to abstracts located farther from such joints. Moreover, subjects tended to navigate slowly outwards from the center in their search patterns. In general, participants liked the virtual environment, but identified usability problems such as unfamiliarity with VRML viewers and clustering models for visualization. In addition, the researchers noted the need for a larger number of explicit navigational cues.

In summary we would say that research on physical and information exploration has considered a variety of issues with a range of approaches. A few themes and generalizations stand out among the variety:

- Navigational expertise consists of an expanded knowledge of domain structure and environmental perceptual cues. Well-designed environments can support this expertise.

- Hierarchy is fundamental to psychological and information structures.
- The distinctions between landmark, route, and survey knowledge are fundamental. They appear in the physical world, as well as in both hypermedia and virtual worlds. Accordingly, environment and tool design should consider these distinctions.
- Way finding and design principles from the real world seem to apply to electronic worlds, with certain qualifications and adaptations.
- A successful study of navigation should use several signatures (sets of cognitive encoding and decoding operations) and data collection mechanisms.
- Environment, tools, and individual differences (e.g., in spatial and memory abilities) affect user strategies and performance, to varying degrees.

The information to be assimilated falls broadly into two categories, identified by [5] as navigational ("information that serves to organize us and the world in spatiotemporal terms") and destinational ("information judged to be of intrinsic value").

As globally networked information spaces develop, information visualization research has begun to consider navigational issues as well. (Besides, as mentioned above, multi-purpose elements can be designed to serve as both navigational and destinational data.) A general definition of information visualization is the following: "The use of computer-supported, interactive, visual representations of abstract data to amplify cognition. The fundamental strategy of visualization is to convert data to a visual form that exploits human skills in perception and interactive manipulation." [13]

Chalmers investigated user exploration using a map or landscape metaphor. His work derives in part from wayfinding research by [41] and [51]. The prototype Bead system had several design goals, including a legible information space [41] reliance on sensori-motor skills [14] and good information design. The resulting system builds and displays a spatial model, which shows both detail and overview information. The construction uses patterns of document similarity. Visual proximity thus represents an abstract information dimension, and the spatial metaphor supports visualization of overall relationships. Also, visualized proximity often reveals useful connections.

A key design goal of information visualization is to support the user in the processes of pattern recognition and aggregation/abstraction. Related system problems include the display of large quantities of information while retaining legibility. For navigation, physical realism through UI metaphor and computer graphics is essential. In addition, users require system support for visualizing tradeoffs between navigational cost and informational benefit.

There exists some tension between the predominant approach that seeks to off-load cognition onto perception [14] and the opposite approach that seeks to generate strong cognition through basic perceptual cues [21] For the future, many of the reviewed techniques would benefit from validation and/or refinement by behavioral testing. Extrapolation of these techniques to immersive VR would also be instructive. Research opportunities exist for integrating different combinations of the reviewed techniques, particularly those that can be paired as environment and tool. Ultimately, effective computational and interaction techniques are both required for effective navigation in electronic worlds.

3.13.1 General Principles for Landscape Visualizations

Preparers and presenters of landscape visualizations should adhere to the following general principles:

> Accuracy: realistic visualizations should simulate the actual or expected appearance of the landscape as closely as possible (at least for those aspects of the landscape being considered); visualizations should be truthful to the data available at the time.
>
> Representativeness: visualizations should represent the typical or important range of views, conditions, and time-frames in the landscape which would be experienced with the actual project, and provide viewers with a range of viewing conditions (including typical worst-case conditions at a minimum).

Visual clarity: the details, components, and overall content of the visualization should be clearly communicated.

Interest: the visualization should engage and hold the interest of the audience, without seeking to entertain or "dazzle" the audience.

Legitimacy: the visualization should be defensible by following a consistent and documented procedure, by making the simulation process and assumptions transparent to the viewer, by clearly describing the expected level of accuracy and uncertainty, and by avoiding obvious errors and omissions in the imagery.

Access to visual information: visualizations (and associated information) which are consistent with the above principles should be made readily accessible to the public via a variety of formats and communication channels. [63]

Different ways of visualization of the acquired information would provide a valuable resource. By using two and three dimensions to observe the information new insights may be acquired.

3.13.2 Information Landscapes

David Small's pioneering Information Landscapes at the Visible Language Workshop is a source of inspiration for this work. [64]

Similarly a three-dimensional interactive topic map visualization tool, UNIVIT, was implemented by Le Grand and Soto [37] which uses virtual reality techniques such as three-dimensional, interaction and provided differing levels of detail.

Initially, new media emulate the media that they replace, before creating any new paradigms. This can certainly be said of the way in which electronic media have handled text. The glowing glass screen of the computer is seen as a flat surface on which is pasted images that resemble sheets of paper. Window-like systems have advanced this emulation only to the extent that they allow for many rectangular planes of infinitely thin virtual paper to be stacked haphazardly around the glass surface of the computer display. The graphic power of computer workstations has now advanced to the point where we can begin to explore new ways of treating text and the computer display. By allowing the computer to do what it can do well, such as compute three-dimensional graphics and display moving images, we can develop a truly new design language for the medium.

By escaping the confines of the flat sheet of paper, we can arrange information into meaningful landscapes that exhibit qualities of mystery, continuity, and visual delight.

3.13.3 Themescape

ThemeScape software by Cartia Incorporated [67] provides different types of topographical maps. They explored the use of landscape visualizations with varying levels of details in the maps generated.

It is a suite of four software products that automatically organizes text-based material into interactive landscapes of information, helping users quickly understand what is inside a document collection, find information of interest, and discover new insights.

They look like topographical maps with mountains and valleys, as shown in Figure 18.

Figure 18: Themescape Landscape Visualization

This visualization is very interesting since it combines different representations in several windows. Users may choose one of them according to the selected type of information. Text is available to ensure that topics are easily discerned.

> The concept behind a ThemeScape map is simple: the greater the similarity between any two documents, the closer together they appear on the map. Concentrations of documents about a similar topic literally "pile up" to form peaks, and the distance between peaks shows how closely the topics are related. ThemeScape gives you a clear picture of the information inside.
>
> Because documents are simply points on the map, it is possible for ThemeScape to show thousands of documents at once without overwhelming the user. Zooming into the map reveals greater detail. For any region on the map, a click of the mouse pops up a list of documents with related content. Pointing to any document title displays a short text summary. A mouse click links the user directly to the original document.

Visual tools may utilize animated 3D landscapes which take advantage of human beings' ability to navigate in three-dimensional spaces, recognize patterns, track movement, and compare objects of different sizes and colors. Users may have complete control over the appearance of data. Virtual reality techniques include interactivity and the use of different levels of detail (LOD). Immersion in virtual worlds makes users feel more involved in the visualization. [37]

I use two forms of visualization, initially one gets a graph and from it one is able to see the tree like structure of the documentation that they are working on. The peripheral subjects being furthest from the base or beginning point on the tree. Alternatively one may look at the landscape visualization and by simply clicking on certain regions they can zero in and obtain the details of the relationships between the concepts (topics)

3.14 Self-Organizing Maps

The Self-Organizing Map represents the result of a vector quantization algorithm that places a number of reference or codebook vectors into a high-dimensional input data space to approximate to its data sets in an ordered fashion. When local-order relations are defined between the reference vectors, the relative values of the latter are made to depend on each other as if their neighboring values would lie along an "elastic surface". By means of the self-organizing algorithm, this "surface" becomes defined as a kind of nonlinear regression of the reference vectors through the data points. A mapping from a high-dimensional data space greater than n onto, say, a two-dimensional lattice of points is thereby also defined. Such a mapping can effectively be used to visualize metric ordering relations of input samples. In practice, the mapping is obtained as an asymptotic state in a learning process. A typical application of this kind of SOM is in the analysis of complex experimental vectorial data such as process states, where the data elements may even be related to each other in a highly nonlinear fashion.

The process in which the SOM is formed is an unsupervised learning process. Like any unsupervised classification method, it may also be used to find clusters in the input data, and to identify an unknown data vector with one of the clusters. On the other hand, if the data are a priori known to fall in a finite number of classes, identification of an unknown data vector would optimally be done by some supervised learning algorithm, say, the Learning Vector Quantization (LVQ), which is related to the SOM. [34]

The SOM [35] is a means for automatically arranging high-dimensional statistical data so that alike inputs are in general mapped close to each other. The resulting map avails itself readily to visualization, and thus the distance relationships between different data items (such as texts) can be illustrated in a familiar and intuitive manner. The SOM may be used to order document collections, but to form maps that display relations between document contents a suitable method must be devised for encoding the documents. The relations between the text contents need to be expressed explicitly. If the *words* are first organized into word categories on a *word category map*, then an encoding of the documents can be achieved that explicitly expresses the similarity of the word meanings.

The encoded *documents* may then be organized with the SOM to produce a *document map*. The visualized document map provides a general view to the information contained in the document landscape, where changes between topics are generally smooth and no strict borders exist. Easy exploration of the document landscape may then be provided. [35]

To provide a general understanding of what the SOM is and why it is an especially suitable method for ordering large collections of text documents, the following expedition of thought may be helpful: Consider an information processing system, such as the brain, which must learn to carry out very different tasks, each of them well. Let us assume that the system may assign different tasks to different sub-units that are able to learn from what they do. Each new task is given to the unit that can best complete the task. Since the units learn, and since they receive tasks that they can do well, they become even more competent in those tasks.

This is a model of specialization by competitive learning. Furthermore, if the units are interconnected in such a way that also the *neighbors* of the unit carrying out a task are allowed to learn some of the task, the system slowly becomes ordered so that units near each other have similar abilities, and the abilities change slowly and smoothly over the whole system. This is the general principle of the SOM. The system is called a *map* and the task is to imitate, i.e., *represent* the input as well as possible. The representations become ordered according to their similarity relationships in an unsupervised learning process. This property makes the SOM useful for organizing large collections of data in general, including document collections. [36]

3.14.1 The SOM Algorithm

3.14.1.1 Initialization

Before the training, initial values are given to the prototype vectors. The SOM is very robust with respect to the initialization, but properly accomplished it allows the algorithm to converge faster to a good solution. Typically one of the three following initialization procedures is used:

- Random initialization, where the weight vectors are initialized with small random values
- Sample initialization, where the weight vectors are initialized with random samples drawn from the input data set
- Linear initialization, where the weight vectors are initialized in an orderly fashion along the linear subspace spanned by the two principal eigenvectors of the input data set. The eigenvectors can be calculated using Gram-Schmidt procedure.

3.14.1.2 Learning process of the SOM goes as follows:

1. One sample vector x is randomly drawn from the input data set and its similarity (distance) to the codebook vectors is computed by using e.g. the common Euclidean distance measure:

   $||\mathbf{x} - \mathbf{m}_c|| = \min_i\{||\mathbf{x} - \mathbf{m}_i||\}$,

2. After the Best Matching Unit (BMU) has been found, the codebook vectors are updated. The BMU itself as well as its topological neighbors are moved closer to the input vector in the input space i.e. the input vector attracts them. The magnitude of the attraction is governed by the learning rate. As the learning proceeds and new input vectors are given to the map, the learning rate gradually decreases to zero according to the specified learning rate function type. Along

with the learning rate, the neighborhood radius decreases as well.

The update rule for the reference vector of unit i is the following:

$$m_i(t+1) = \begin{cases} m_i(t) + \alpha(t)[x(t) - m_i(t)], & i \in N_c(t) \\ m_i(t) & i \notin N_c(t) \end{cases}$$

3. The Steps 1 and 2 together constitute a single training step and they are repeated until the training ends. The number of training steps must be fixed prior to training the SOM because the rate of convergence in the neighborhood function and the learning rate are calculated accordingly.

After the training is over, the map should be topologically ordered. This means that n topologically close (using some distance measure e.g. Euclidean) input data vectors map to n adjacent map neurons or even to the same single neuron.

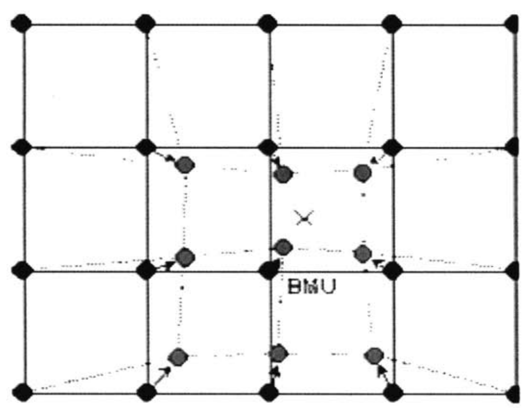

Figure 19: Updating the BMU and its neighbors. This is done towards the input sample marked with x. The solid and dashed lines correspond to situation before and after updating, respectively.

3.14.1.3 Map Quality Measures

After the SOM has been trained, it is important to know whether it has properly adapted itself to the training data. Because it is obvious that one optimal map for the given input data must exist, several map quality measures have been proposed. Usually, the quality of the SOM is evaluated based on the mapping precision and the topology preservation.

3.14.1.4 Mapping precision

The mapping precision measure describes how accurately the neurons 'respond' to the given data set. For example, if the reference vector of the BMU calculated for a given testing vector x_i is exactly the same x_i, the error in precision is then 0. Normally, the number of data vectors exceeds the number of neurons and the precision error is thus always different from 0.

A common measure that calculates the precision of the mapping is the average quantization error over the entire data set:

$$E_{q=\frac{1}{N}}\sum_{i=1}^{N}\| x_i + m_i \|$$

3.14.1.5 Topology Preservation

The topology preservation measure describes how well the SOM preserves the topology of the studied data set. Unlike the mapping precision measure, it considers the structure of the map. For a strangely twisted map, the topographic error is big even if the mapping precision error is small.

A simple method for calculating the topographic error:

$E_{t=\frac{1}{N}}\sum_{k=1}^{N}u(x_k)$, where $u(x_k)$ is 1 if the first and second BMUs of (x_k) are not next to each other. Otherwise (x_k) is 0.

3.14.2 SOM and Constructivism

The epistemological theories of knowledge have traditionally been based on predicate logic and related methodologies and frameworks. The basic assumption is that the world consists of objects, events and relationships. The language and the conceptual structures and then supposed to reflect rather straightforwardly this ontological structure. Learning has been seen as a means to memorize the mapping from the epistemological domain (to put it simply: words) into the ontological domain (objects, events and relationships). This view has been dominant at least partly because of the consistent formalization of the theory through the use of symbolic logic. Moreover, the use of the von Neumann computer as the model or metaphor of human learning and memory has had similar effects and has strengthened the idea of the memory as storage of separate compartments which are accessed processed separately and which are used in storing and retrieving information more or less as such. In the philosophical discourse, there have been several opposing views but there have been little means to model or formalize these ideas before connectionist models.

In cognitive science constructivists have questioned the idea that the origin of knowledge is outside human beings. Instead, they find that knowledge is a human construction. The so-called "flat ontology" implies that one's view of reality emerges directly from sensory data, without the need for any intervening cognitive mechanism. However, the flat ontology appears to have no empirical support [40].

SOMs can also be considered as a memory model. It is dynamic, associative and consists of elements that can also be called adaptive prototypes. Inputs are not stored as such but comparison is made between the input and the collection of prototypes. The closest prototype of the input is adapted towards the input. The same operation is also conducted for the neighboring prototypes, which gives rise to the topographical order on the map. Thus, the adaptation process in the SOM algorithm is based on the principle that what already exists in the system also influences the learning result.

Considered superficially, one could claim that modeling learning phenomena through the use of the SOM would be mentalistic. However, it is possible to construct a model in which a number of autonomous map-based agents interact in such a way that they perform social construction of knowledge and find inter-subjective epistemological agreements [27].

The SOM can be used in visualization of conceptually rich and complex phenomena. Through constructive theory it becomes apparent that there can be many and different interpretations and points of view towards an area of discussion. This kind of the subjectivity and inter-subjectivity of interpretation may be visualized by using SOMs [28].

3.14.3 WEBSOM

The WEBSOM method organizes a document collection on a map display that provides an overview of the collection and facilitates interactive browsing.

By virtue of the Self-Organizing Map algorithm, documents can be mapped onto a two-dimensional grid so that related documents appear close to each other. If the word forms are first organized into categories on a word category map, an encoding of the documents can be achieved that explicitly expresses the similarity of the word meanings. The encoded documents may then be organized by the SOM to produce a document map. The visualized document map provides a general view of the document collection. [33]

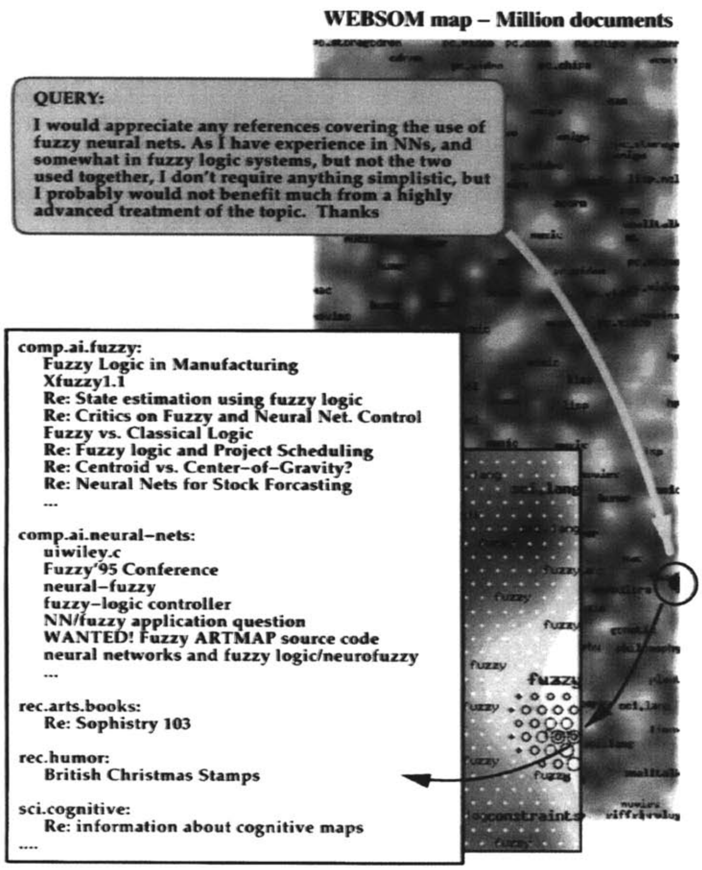

Figure 20: WEBSOM at Work. An example of a content-directed search on a map of over a million documents from various Usenet newsgroups. The labels have been derived from newsgroup names and selected using an automatic method. As an alternative to navigating based on the labels one may perform searches on the map. A natural language description of interest,

even a full document, can be used as a search key to find map units containing similar documents. The map units that are found can also serve as starting points for further exploration.

I have implemented the SOM in order to obviate the connections to hitherto unobserved relationships and thus tacit knowledge. The inter-relatedness of information that is not at a glimpse obvious is brought to the fore.

3.15 The Proposed Solution

Highly usable systems can be designed only after intrinsic knowledge of the intended user has been obtained. The application allows users to express their innovative minds in ways to that no tool currently allows them to.

3.15.1 Schon's Reflection Steps

The methodology that I use to develop the application will explore the use of Schon's theory of reflection and will thus have the following steps:

a) Framing of the problem
b) Making a move
c) Reflecting on the move
d) Repeating the process

In other words there is no standard approach to solving problems but rather an iterative build-up experimentation methodology is utilized. [66]

I use graphs and landscape as data-visualization techniques to display interconnected information.

The user gets an overview of the map created, which shows the main features of the structure in order to quickly explore the structure. The map generated allows the user to identify unexpected or peculiar information by zeroing into it. "The visual information-seeking mantra is overview first, zoom and filter, then details on demand". [59]

3.15.2 Critical Moments Reflection

People who have done the critical moments reflection process report that it has allowed them to gain visibility of their experience, the thinking they bring to it, their methods for doing their work, and the often hidden theories and assumptions that guide their work. We have also done analyses of transcripts from group reflection sessions that have let us observe the mechanics of collective transformative learning that take place during a Critical Moments Reflection session. From these analyses, they have identified the following outcomes of Critical Moments Reflection:

- At a minimum, the process generates basic awareness of the tacit learning that has taken place for people doing their work.
- The reflection process usually extends and deepens the learning from experience.
- People sharpen and clarify their questions about the work, or find new questions.
- When done in a group, the process leads to developing connections across individual experiences of learning through the work and integrating the meanings from different perspectives on those experiences.

People develop more complex understandings of what happened, why it happened, and what it led to. [1]

(Practitioners) feel profoundly uneasy because they cannot say what they know how to do, cannot justify its quality or rigor.

The dilemma of rigor or relevance may be dissolved if we can develop an epistemology of practice that: (1) places technical problem solving within a broader context of reflective inquiry; (2) shows how reflection-in-action may be rigorous in its own right; and (3) links the art of practice and uncertainty and uniqueness to the art of research.

The extent of our capacity for reciprocal reflection-in-action can be discovered only through an "action science" that seeks to make what some of us do on more rare occasions into a dominant pattern of practice. [60]

3.15.3 Reflection Modes by McDrury

I have found very few tools that aid in reflective practice. McDrury of University of Otago [44] explored the use of storytelling, journaling and spontaneous drawing to try and tease out reflective practice.

3.15.3.1 Storytelling:

This was perceived as a very useful way to understand an event from differing perspectives and to get constructive ideas about how to deal with similar situations in the future. It values and incorporates the traditional method of sharing information, enables clarification and debriefing while affirming practice.

The main perceived weaknesses of storytelling involved worries about the accuracy of the story. There was also uncertainty about how much of the story to tell - this particularly related to telling of the context. Participants noted the importance of the teller of the story being open to hearing their story from other perspectives. Participants felt that they needed to have good group skills and that the facilitator played a particularly important role within the reflective process.

3.15.3.2 Journal:

Having a written record of an event was seen as a key strength, especially when it enabled the writer to become aware of the progress they were making in their clinical performance. Participants had the feeling that the record of the story was an "accurate" recording of events and feelings which occurred.

The problems with using journal entries as a tool were consistent with the literature: not making entries because of lack of time or tiredness, and finding it easier to record diary events rather than feelings. In addition some participants wondered about the "accuracy" of the entries when they were examining them at a later point, and felt they recognized bias or areas of embellishment .

3.15.3.3 Drawing:

This tool was found to be particularly helpful. It valued the intuitive and feeling responses to a situation, and uncovered layers of meaning and content in clinical interaction which lead to a better understanding of a clinical incident. The creative expression which used color was a fun alternative way of exploring difficult or abstract events. Being able to get feedback from other participants who viewed the drawing was particularly valued. Working within the framework was essential to develop skills with a new technique.

The major perceived weakness of the drawing related to concerns about artistic ability, being unsure about how much the drawer would reveal, and needing more time. This time issue related both to the time needed to complete the drawing and the time needed to become more familiar with the technique.

The most popular tool among participants was storytelling. Drawing was the second choice and journaling was the third choice.

3.15.4 Progress Portfolio

Progress Portfolio was designed to be a catalyst for reflective inquiry activities. It enables students create artifacts that represent the evolving states of their investigation process.

First, our basic approach to promoting reflective inquiry revolves around the notion of creating concrete artifacts to anchor reflective processes. In order to reflect, one must have something to reflect on. Second, we designed the tools of the Progress Portfolio to serve as an implicit task model for reflective inquiry. The tasks of creating a record, monitoring, and communicating are directly supported by our tools. [39]

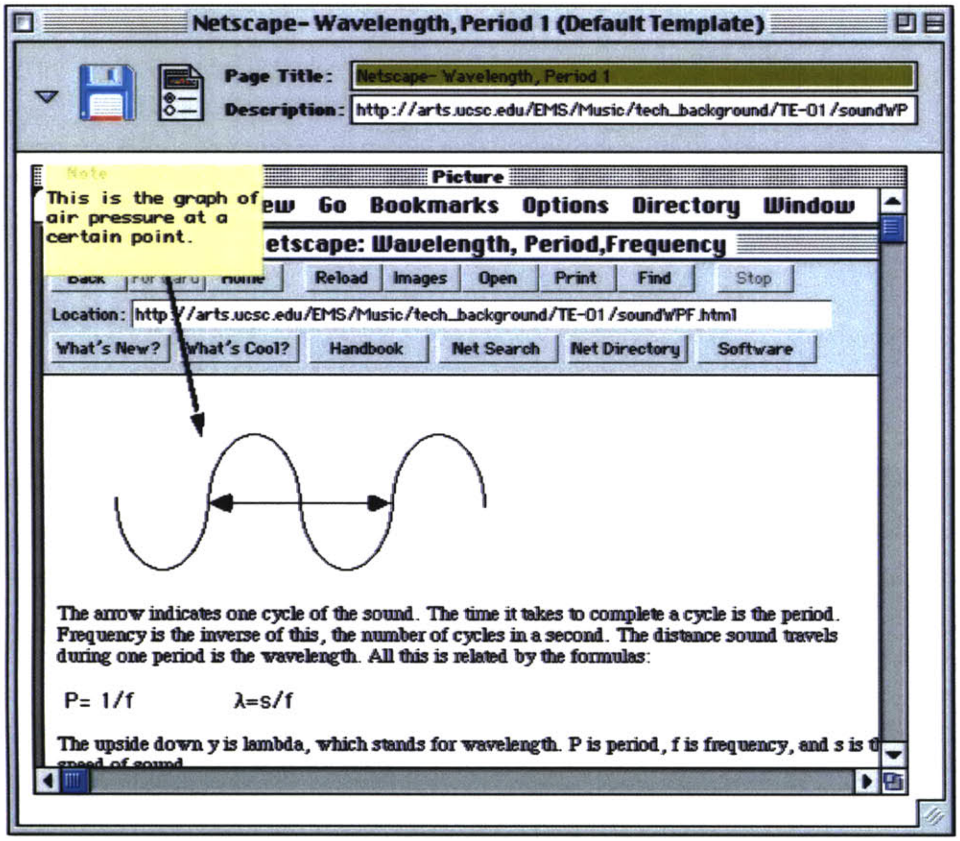

Figure 21 : Progress Portfolio in Use. Students capture graphics and text from a web site onto a Progress Portfolio page and annotate it with a sticky note to help them remember why they captured it.

3.15.5 TalkTV

Searching through digitized text is less computationally expensive than searching through image data, and has the additional benefit of being relatively easy to query. One of these early projects was *Network Plus* [4], which augmented televised news broadcasts with contextual information from the Associated Press news wire. Keywords extracted from the closed- captioned were used to retrieve relevant information from external database, such as the wire-services, and this information was presented as a synchronous overlay to the broadcast program using the closed captioning as a time code.

Television programs can be structured by their transcripts, which are embedded as closed captions in the signal of most television shows. With TalkTV video-editing software, rearranging lines of dialog automatically creates new video scenes, thereby enabling television viewers to become editors. The idea is to allow individuals, to create and investigate ways to index and deliver content which may then be reused.

Working with a computer database, [65] have sliced and diced several emblematic television shows into their component parts. A result of this archaeology of entertainment is the digital equivalent of thousands of carefully labeled pot shards.

However, televisions don't use closed captions to structure the display of television programs in the same way that a web browser uses HTML. But if television shows were organized this way, reordering the sequence of the closed captions would create new videos. With this in mind, a video editor was created to treat closed captioning as if it were a markup language to see if they might be more easily appropriated into creative projects.

Digitized television shows are placed into a folder where TalkTV loads in their closed-captioned transcript, which includes time code of when a line of dialog approximately begins and ends in a television show. The TalkTV editor segments television shows into small clips of dialog by marking *in points* at the beginning of a closed caption and marking *out points* on stop characters in the subtitles such as periods, question marks, and exclamation points. In this way, TalkTV processing roughly parses the content at the boundaries where dialog exchanges take place. [7]

Since all the content from the individuals can be aggregated, it is possible to share with the experiences of others in the rich repository that is created from the repository.

Chapter 4

4.1 Design and Implementation

In building this application I chose to use Java as my programming language of choice. The main reason that I chose Java after initially experimenting with Smalltalk and Jython was due to the wide availability of open source applications that could be easily added onto CMReflector without the additional trouble of having to master a new language. This became very useful when aggregating TalkTV and the WPISOM system into the final application.

I used Eclipse during most of the development phase. Eclipse is an open platform for tool integration built by an open community of tool providers. The Eclipse Platform is designed for building integrated development environments (IDE) that can be used to create applications as diverse as web sites, embedded Java TM programs, C++ programs, and Enterprise JavaBeans TM.

The Eclipse Platform's principal role is to provide tool providers with mechanisms to use, and rules to follow, that lead to seamlessly-integrated tools. These mechanisms are exposed via well-defined API interfaces, classes, and methods. The Platform also provides useful building blocks and frameworks that facilitate developing new tools.

I integrated JEasy into Eclipse in order to explore possible user interfaces. It presents you the most important swing components with the most usual properties. All the GUI-components of your application are stored in a project inside the JEasyRepository. The repository helps you to specify the properties of the Swing components and the relationship between them.

A JEasy object exists for almost each SWING component. The repository creates an xml file with all your objects, their properties and their relationships. The JEasy objects read these entries at start up of the program and create all of the origin SWING components. The repository uses forms and list boxes to help you to define your properties and the hierarchical order. JEasy is completely written in Java. It works with the runtime environment JRE from JAVA 2 or the JAVA 2 SDK 1.3 or higher.

Since I only had to make changes to make changes in the JEasyRepository it allowed me to make iterations to the GUI easily and I could thus experiment much more rapidly with the look of the interface.

4.2 Architecture

4.2.1 CMReflector Architecture

The architecture I have created is primarily a client-server (message-passing architecture). Thus the Client sends messages to Server, and Server responds with messages back to Client. Both the Client and Server can exist on one machine, or can be separated by a network. Client can be logged into a local Server and to networked Servers at the same time. The client-side is comprised of a number of Context elements. The tool, Nexist, provided an appropriate template from which to borrow from.

All *primary information* stored in the relational database is stored in the form of an XTM document, a Topic Map. *Secondary information* is comprised of the following kinds of information:

- Lexicon (all important words and word phrases)
- Concordance (word and phrase frequencies)

Secondary information is created and maintained during all transactions in support of background activities performed by the Knowledge services.

The application aids users in gathering the relevant information in a coherent fashion through Topic Maps. [50] Using simple graphs obtained from the Topic Maps of their areas of inquiry; "Landscape Maps" are generated that allow for further scrutiny. The Landscape Map acts as a data visualization technique to display interconnected information by assigning coordinates to topics according to their interconnections and height to coordinates according to the degree of relevance or the degree of convergence of multiple topics. [42] [18]

The XTM standard for XML Topic Maps should provide an adequate API for serialization of all information transferred between client and server. An underlying assumption in this hypothesis is that XTM documents can be robust enough to represent and index all information contained in the CMReflector *Universe of Discourse*.

The XTM standard provides a means by which object and attribute descriptors can be serialized. CMReflector provides a message-passing architecture that can carry XTM documents along with appropriate commands back and forth between client and server.

Topic occurrences create "sibling" relationships between repository objects. A single object may be an occurrence of one or more topics, each of which may have many other occurrences. When a user finds or browses to a given repository object, this sibling relationship enables inquiry into whether there are other objects related to the same topic.

Topics associations create "lateral" relationships between subjects, allowing a user to see what other concepts covered by the repository are related to the subject of current interest and to easily browse them.

Figure 22 illustrates the primary elements in this architecture and the sequence in which events occur.

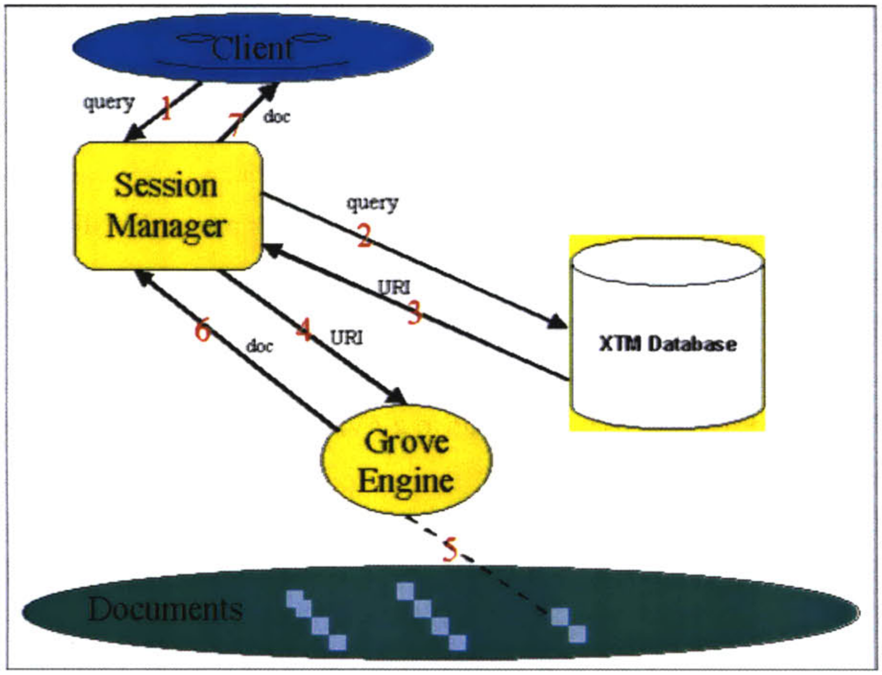

Figure 22: Primary Elements in Architecture and the Flow of Events

The primary objects in this universe are:

- Context

  Client

- Content

  Server (implied)

  Session manager

  XTM database (relational, in this case)

  Grove engine

  Documents (Legacy and CMReflector-specific) *The Universe of Discourse*

A user at a client application may make some gesture causing a request to server for some information, perhaps to display a requested document. That request is parsed by Session manager and turned into an XTM database query, which returns a URI for the requested document. This action illustrates a key point in this architecture.

All information in the Universe of Discourse is referenced in one or more Topic Maps

A returned URI could locate information in any of the following kinds of documents:

- Text files
- XML files
- Legacy relational databases
- Office documents
- Other documents

Because of the inherent heterogeneity of this universe of discourse, a single URI would otherwise not be useful to access such information. Therefore, I borrow from the SGML community and their concept of a grove engine.

A grove engine is capable of mapping an input URI into the location of whatever the URI represents, which could be whole documents, specific elements of a particular document, and so forth; a grove engine provides a single API for addressing heterogeneous information resources.

Most implementations of grove engines represent a document as a node in a graph, the node containing properties which describe the referenced document. Those properties make it possible to convert a request into a specific address. If an object is, indeed, addressable, a grove can find it, assuming that a property set exists for that document.

This architecture suggests that there will be a requirement for construction of a node for every document that enters the CMReflector Universe of Discourse. One cannot just "drop a document into the pool;" rather, one must submit it to an engine designed to read the document and construct a property set.

The Template Maps serve to act us a guide for new users, they exhibit the possible ways in which to formalize their reflective practice. They accord a skeleton from which to work. The inherent structures of the maps help to facilitate non-obvious connections between ideas. They declare base constructs that may then be used across multiple maps. Self-Defined Maps allow for a much more free form expression. The user defines each option at every step and is thus forced to be more articulate about their description.

I used HSQLDB as my database. It is a relational database engine written in Java, with a JDBC driver, supporting a rich subset of ANSI-92 SQL (BNF tree format). It offers a small (less than 160k), fast database engine which offers both in memory and disk based tables. Embedded and server modes are available. Additionally, it includes tools such as a minimal web server, in-memory query and management tools.

A parsed Topic Map can be graphed to create a simple graph. This provides the simple tree from which one can then drill down on, to find the details as described in Chapter 2. A more detailed view becomes available showing all relationships in a landscape view. Queries can be made on the Topic Map from a search window, allowing for varied search definitions.

4.2.2 SOM Architecture

4.2.2.1 The Document Map

The documents are automatically arranged into a visual map using the SOM method. Prior to be arranged, the textual content of the documents is converted into numerical form. Corresponding to each document are numeric values. This type of document encoding is based on a system whereby a certain combination of numeric values corresponds to each word. The share of this combination is the larger in the code the more often the word occurs in the document.

The labels of the document map, i.e. the words written on the map, are selected with an automatic method to best describe the documents of the area in question.

4.2.2.2 The search

The search text is coded numerically in the same way as in forming the document map. The closest corresponding area to the code thus obtained is sought on the map by comparing the code to the models matching the units of the map. The closest map units retrieved through the search are shown as circles on the map.

XTM files using XSLT sheets are generated for the topics, these act us the portable transportation and storage layer that then populates a relational database. The metadata is to be captured by the overt description of the user as well as by derivation from the structures build into the map. Afterwards an analysis of the relations is made to weight the relationships between them using the SOM.

I have organized topics onto a two-dimensional grid so that related topics appear close to each other. This principle was proposed by Kaski et al. [33] to find optimal coordinates for topics using the SOM algorithm.

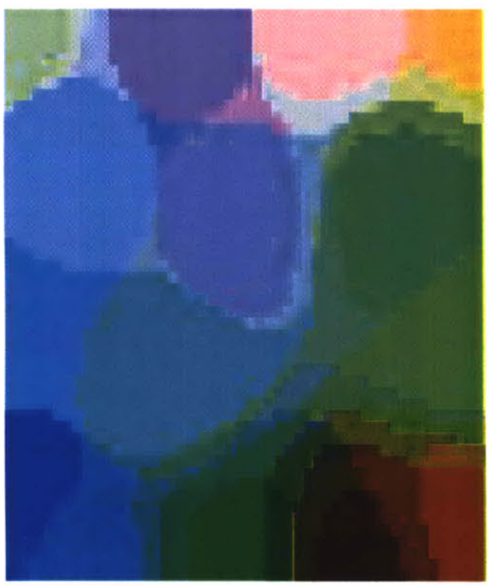

Figure 23: Landscape View of documentation

The layout that I have used is that of a Landscape Map [See Figure 23]. It organizes documents such that those with similar content are placed close to each other and protrusions appear where there is a concentration of closely related topics. Thus if most of the articles related to a practitioner's questions are from the same source, it would appear as a peak. If there are valleys that are evident in such a map they could be explored since they may harbor interesting hitherto unknown circumstances, events or content. [10] [25]

The different topics in this case are represented by different colors. A user will have the option of making the selection of the color in which they want to portray any of the aspects of their map. There is the option of allowing the system to choose colors by default.

Seymour Papert once observed that the difficulty is not in the inherent complexity of the material brought to consciousness but in our ways of representing the complexity. [49]

Chapter 5

5.1 Evaluation

Does CMReflector enhance reflection?

I used CMReflector over the cause of two months. During this time I went through several iterations to try and meet the requested requirements. I also used the tool many times in trying to test both the effectiveness and efficiency of the features as I developed them.

I participated in three reflection sessions at the CRCP. During these sessions I was both an observer as well as a participant. I thus tried to be a reflective practitioner first, in order to get a good impression of what the process entailed. During these sessions I compiled notes from the comments made by the participants, these were both my personal observations as well as those of the facilitators of the sessions. It is these notes (my reflection) of the sessions that I relied on mainly in coming up with the design that CMReflector would have.

During one of the sessions, participants were asked to brainstorm on what would be their requirements as far as getting a tool that aided them in reflection would be. I compiled a list of these requirements. At the subsequent meeting I got a non scientific criteria of what order of importance those features ranked. This order was agreed upon by consensus and was thus not amenable to much study. It is from this list, described in Chapter 1.4 that I obtained what would be my user requirements.

5.2 In Built Evaluation Tools

I have built consistency tests to check on how many constrains will have been fulfilled. Statistical tests can thus be easily performed to check on the number of topics, associations and occurrences in each or specific classes. This can also be used to check on the lengths of the association chains within certain classes or class combinations. These tests would be carried out on the maps created by the test users.

5.3 Difficulty in Evaluation

CMReflector suffered from the lack of user testing. Participants at these reflection sessions were from all over the country and thus the possibility of recruiting them for a such a study was greatly diminished.

Since the tool would be used by very disparate practitioners, the criteria for evaluation would be quite complex in order to be accurate. Users could be experts at reflection or completely new to the practice, they may be avid computer users or novices.

It is difficult to accomplish reflection or to clearly observe its occurrence. [38] Creation of a tool that allows for or encourages reflection was a challenge. This is because it is fairly difficult to pin down when reflection has taken place or what part the mediation tool (CMReflector) has played in facilitating the process.

Since the tool only begins to show it's effectiveness after the compilation of content for a while, it would be necessary to carry out longitudinal studies. These would not only be expensive but would also be time consuming. The duration of this thesis did not allow for engaging in such an exercise.

Another potential difficulty in evaluation is the need for many users in order to aggregate the content in a repository. Unless a critical mass of users is gained and therefore a large amount of content, the possibility of making serendipitous discoveries off the graph or off the landscape visualization is severely limited.

It was also very difficult to find appropriate comparisons especially considering the specificity of CMReflector. Most tools are not created to cater for a such a narrow user base.

Table 4: Documentation Tools Use to be used for Comparison

Documentation Category	Specific Example
Blogging Tool	Moveable Type
Community Information Sharing	HDL
Desktop Publisher (DTP)	PageMaker
Document Management	Metamorph
Editor	Notepad
Electronic Distribution	Acrobat
Filter/Translator	Filtrix
Flowcharting	Visio
Graphics/Presentation	Power Point
Hypermedia	HyperCard
Integrated Office Automation	Microsoft Office
Page Layout	Adobe Illustrator
Scanning	Omnipage
SGML-Based/CALS	Omnimark
Technical Drawing	MapCon
Text Processor/Batch Compiler	Eroff
Video Blogging	Visual Communicator
Web Authoring/HTML Editor	FrontPage
Word Processor	Microsoft Word

I used Table 4 to compare the features in these tools to those in CMReflector.

5.4 Functional Comparison Matrix

In order to fully appreciate the differences that CMReflector has in comparison to the other tools it would require comparing it across many dimensions as is evident from Appendix B. The list below merely suggests the really important ones. However for purposes of comparison in the matrix on Table 7, only ten features are considered.

Does it allow for personalization?
Is it engaging to the user?
Is there an opportunity for fun?
Are serendipitous experiences possible?
Can annotations that capture metaphor be made?
Can semantic markers be more creatively generated?
Does it allow for merging knowledge bases?
Do they have visualization systems?
Is it possible to seed content into the tool?
Provision of a tool that can be used to enhance existing tools
Provision of multi-modal capabilities to documenting
Provision of a rigorous taxonomy for the tagging
Provision of multiple ways to view the information
Provision of a way to decipher tacit knowledge from it
Provision of a tool that fosters reflection while being used
Does it allow for multiple operating systems?
Is the source code available?

From Table 5, it is obvious that most tools have a very specific user base and they cater to those users very effectively. HypeCard is the only tool that compares to CMReflector.

Documentation Category	Application Example	Collaboration	Support Mechanism	Capture Tacit Knowledge	Aids in Reinforcement	Capture Context	Multimodal	Repository Manipulation	Divergent End Artifacts	Personalization Capable	Semantic Web Enabled
Customized Application	CMReflector	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Graphics/Presentation	PowerPoint	N	N	N	N	N	N	N	Y	Y	N
Video-Blogging	Visual Communicator	N	N	N	N	N	N	N	N	Y	Y
Blogging	Moveable Type	N	N	N	N	N	N	N	N	Y	Y
Word Processor	Word	N	N	N	N	N	N	N	Y	Y	N
HTML Web Authoring	FrontPage	N	N	N	N	N	N	N	Y	Y	Y
Community Information Sharing	HDL	Y	N	N	N	N	Y	N	N	N	Y
Desktop Publishing	PageMaker	N	N	N	N	N	N	N	N	N	N
Text Editor	Notepad	N	N	N	N	N	N	N	N	N	N
Flowcharting	Visio	N	N	Y	N	N	N	N	N	N	N
Hyper-Media	HyperCard	N	N	N	N	N	Y	Y	Y	Y	N
Page-Layout	Adobe Illustrator	N	N	N	N	N	N	N	N	Y	N
SGML Based/CALS	Omnimark	N	N	N	N	N	N	N	N	N	Y
Technical Drawing	MapCon	N	N	N	N	N	N	N	N	N	N

Table 5: Functional Matrix of Documentation Tools

Chapter 6

6.1 Conclusion

My thesis presents an application that seeks to allow for multi-modal documentation in a manner that conforms to the semantic net and that caters for multiple representations or views.

CMReflector focuses on utilizing the computational capabilities of information-technology tools to create much better documentation. As it is, current tools mimic the limited capabilities that existed with pen and paper or, in certain instances, modestly enhances them. [11] My application explores the use of more computational capabilities and hopefully revealing some of the linearity prevalent in documentation.

6.2 Difficulties Experienced

Mastering the discipline of reflective practice is difficult. This is probably due to the fact that few of us have learned how to build the skills of inquiry and reflection into our thoughts, emotions and everyday behavior. [62] Appropriate tools that help facilitate this skill could be invaluable. And since reflection in itself is such a hard problem, how is one then to tell when it has been achieved and whether the mediator (in this case CMReflector) is actually a catalyst or a hindrance ? It is thus hard to establish the choke point in such an undertaking.

How do you formalize something that is inherently fluid and very intangible? Whatever choices were made in such a formalization, some form of compromise was always going to be made. It is thus very difficult to establish how rigorous one has to be in order to ensure that there is some formalization in accomplishing the task while at the same time allowing for spontaneity and ingenuity.

Given that the tool would use XTM as the code for the content and considering that XTM is a very new concept, adoption is likely to take a very long time. While the ISO is working on enhancing this adoption it remains unknown as to what extent the topic mapping paradigm will become adopted.

6.3 Future Work

There is a need to create annotation systems that capture the various metaphors that are unique to each domain or subject area. There are obviously many ways in which certain concepts are understood in different cultures, locations, languages, professions or ages. It would therefore be very useful to create a mechanism that would put into a repository the nuanced metaphors and match them against each other in a database.

By the creation of semantic markers during the documentation process, the end documents become more amenable to be used in a repository in a knowledge base. They could for instance be used to derive common-sense knowledge for applications relating to the domain area.

Different ways of visualization of the acquired information would provide a valuable resource. By using two or even three dimensions to observe the information new insights may be acquired.

The application could be extended to allow instructors or teachers the opportunity to more creatively utilize them to enhance the learning experience of their students by seeding the maps or critiquing them. [8] [71]

6.4 Raelin's Rationale for Reflective Practice

By using Raelin's [54] rational for reflection we are able to asses what the importance of this reflection practice is. Among the reasons are:

- "At times, we are, unfortunately, unaware of our behavior and its consequences. To complicate matters, our unawareness occasionally does not allow us to be open new data or information that would help us to learn from our actions." The CMReflector provides avenues in which to observe or project consequences when one assesses the aggregate opinion on the group reflection landscape visualization.
- "There is an unfortunate gap between what many of us say we will do and what we actually do. "Opportunities potentially exist to capture tacit knowledge both at the individual and group levels.
- "Most of us are biased in how we obtain information that, in turn, produces cognitive errors in our perceptions of reality." By having a collective aggregate to compare with, we are able to fully appreciate our biases or lack thereof. The self-defined templates and pre-defined templates collate the documentation accordingly.
- "Although intuition and past practices can give us very cogent clues in deciphering future situations, often the new situation presents itself in a different context." By virtue of having a repository of information that is aptly tagged, reflective practitioners may start to have some rigor into how they approach problems and obtain solutions.

6.5 Summary

I have used examples, both theoretical and real, to demonstrate that actual use can be made of CMReflector. The theoretical aspect focuses on the appreciation of the literature and body of knowledge in Chapter 3. This implementation was informed by the literature from this past research. I have explored the possibilities of the requirements as expressed by actual practitioners. I believe that the tool offers a different and interesting approach to aiding reflection. While a lot of work remains to be done in terms of testing the tool, a starting point from which to make such studies has been made available.

Although Critical Moments Reflection ends with summarizing the learning that comes from the analysis of one or more critical moments, only some people stop there. More often, the process leads to an effort to re-integrate the learning back into the work, or to identify an inquiry question, a conscious learning or research process going forward. [1]

Appendix A

Dominant Documentation Tools

Color Publishing

Color publishing products provide color separations for high quality color hardcopy.

Examples: Alias, Full Color Publisher

Applications: Color publishing products may be useful in producing color portions of technical orders.

Database Publishing

Database publishing refers to the production of documentation using automated methods to extract information from a database for inclusion in that documentation.

Examples: BASISplus

Applications: Database Publishing products may be used in a software development process where text is retrieved from a database to produce portions of a DoD-STD-2167A documents

Desktop Publisher (DTP)

Desktop publishers provide full screen editing (often with a what-you-see-is-what-you-get interface), graphics manipulation and editing, graphics anchoring to text, multiple text columns, font variety, structuring of complex documents, hanging indents, and precise control over kerning, and may provide close coupling with one or more Computer Aided Software Engineering (CASE) products.

Examples: FrameMaker, Interleaf

Applications: Use desktop publishers for software development projects with resources and a life cycle that justify a higher end desktop publishing product; also for projects that use other CASE products and for production of structured technical documents that include graphics.

Document Management

Document Management products focus on the manipulation and organization of large collections of documents. Document Management products may include indexing, version control, and search and retrieval capabilities.

Examples: Metamorph, RDM.

Applications: Used in the implementation of libraries or repositories where paper is being replaced by electronic media. Document Management products may also be used in the configuration management and archiving of software documentation associated with software maintenance projects that have a long life cycle with frequent and/or significant updates.

Editor

Editors may be line oriented or may be full screen editors. They tend to require less hard-disk space than word processors and may be faster than word processors for searching large files. Some editors allow execution of compilers, linkers, and debuggers without exiting the editor. Specialized language sensitive editors enforce the syntax of specific programming languages.

Examples: EDT, VI, Qedit, Microstar

Applications: Use editors for entering program code, for writing letters and short documents.

Electronic Distribution

Electronic distribution products allow documents to be transmitted electronically complete with graphics and formatting; the recipient's display is independent of application or platform. This means that the receiver can view a WordPerfect document, for example, without having WordPerfect on his/her machine.

Examples: Acrobat, Replica

Applications: Use these to distribute information that does not need to be edited at the receiving end. An electronic distribution process may be implemented using these products to distribute documents for review or printing where there is no need to edit the documents at the receiving end of the distribution.

Extract from Source Code

Extract from Source Code products provide textual or graphical documentation based on information in the programming language source code. An example is a product that extracts comments from source code or captures design information (reverse engineering).

Example: DocGen

Application: These products may be useful if source code documentation is non-existent or of poor quality.

Filter/Translator

Filters/translators are products that modify the format of an input file (text or graphics) and produce an output file in another desired format, e.g. a document in WordPerfect format may be converted to a FrameMaker format. No attempt is made to differentiate between filter and translator in the Documentation Domain. Many word processors and DTPs include filtering functions.

Examples: Filtrix, SGML Translation Products, Hijaak

Applications: Filters are used when data from a product in format x is needed as input to another product in format y. A filter may be used to convert a file to a Continuous Acquisition and Life-Cycle Support (CALS) compliant format.

Flowcharting

Flowcharting products are designed specifically for making flowcharts.

Examples: EasyFlow, allClear, DiagramMaker, Visio

Applications: Use flowcharting products where flowcharts are currently being generated manually.

Graphics/Presentation

The Graphics/Presentation Subdomain includes paint, draw, and presentation packages. Some word processors and DTPs include paint and draw functions. Draw function allows a user to select an object from a menu (e.g., a circle or square) and manipulate it to alter size shape or placement. Paint functions work with images as bit-maps, allowing free-hand sketching and pixel-editing. Presentation products represent numerical information as pie charts, bar charts, and histograms.

Examples: CorelDRAW, Harvard Graphics, Power Point

Applications: Graphics/Presentation products may be useful for flowcharts, data flow diagrams, and images of the computer hardware. Presentation products are used to communicate project status to decision makers.

Forms

Forms products provide framework forms with the facility to create, edit, and modify the forms and the entry of data on the forms.

Examples: PerForm, WordPerfect InForms

Applications: To generate and track software problem reports and change requests.

Hypermedia

Hypermedia as a documentation subdomain refers to products that allow the integration of several forms of information including text, graphics, video, animation, music, voice and sound. The information is linked non-sequentially to provide flexibility to the user receiving the information. Hypertext is a related term applied to textual information that is linked, usually by pointers.

Examples: HyperCard, AnyImage

Applications: Hypermedia products could be used in users' manuals, requirements documents, design documents, test documents, and maintenance documents.

Integrated Office Automation

Integrated office automation products often provide an integrated product with two or more of the following applications: word processing, spreadsheet, database, draw, paint, and mail. Because the applications are integrated, data may be linked between applications allowing updates in one application to be reflected in another application. A spreadsheet may have data that needs to appear in various documents. The spreadsheet updates will appear in all the documents that are linked to the spreadsheet.

Examples: Microsoft Works, PFS: Choice, Framework, Enable

Applications: Use integrated office automation products for projects where the same data may need to be used in spreadsheets, briefings and structured documents.

Miscellaneous

This subdomain includes any product not found in the other subdomains that may be used to aid the documentation process. Examples of the types of products included are acquisition documentation products, on-line documentation generators, documentation generators, document comparers, and modules of CASE products whose functionality is specific to documentation.

Page Layout

Page layout products emphasize features that allow flexibility in the placement of headers, footers, multiple columns, and graphics on a page.

Examples: Adobe Illustrator, Archetype Designer

Applications: Page layout products may provide the DTP functions that a specific word processor lacks without incurring the expense of a high-end DTP. They are also good for newsletters and short documents.

Scanning

Scanning products are software used to convert hardcopy images to an electronic form. They require specialized scanning devices. The images may be converted into text or graphics. The result of the scanning process may be graphics stored in a raster or vector format, text treated as graphics (stored in a raster format), or text stored in an editable ASCII format.

Example: Omnipage

Applications: Projects with hardcopy that should be converted to electronic media for best usage.

SGML-Based/CALS

Includes products whose functionality hinges on SGML or other CALS standards. Product functionality may include editing, viewing, conversion and validation of SGML/CALS compliant data. Standard Generalized Markup Language (SGML) and Computer-aided Acquisition and Logistics Support (CALS) are defined in 2.7.

Examples: FastTag, Omnimark.

Applications: Projects that have a documentation requirement to be compatible with CALS standards; also projects that involve several organizations that need to exchange data and would benefit from having access to data in a common database.

Technical Drawing

The Technical Drawing subdomain includes computer-aided drawing products that are used for drawing straight lines and Bezier curves and for dimensioning. Some DTPs include these functions.

Examples: Generic CADD, Illustrator, MapCon

Applications: Technical drawing products may be used to produce drawings of hardware associated with the software being developed.

Text Processor/Batch Compiler

Text processors and batch compilers accept text files containing embedded codes that are interpreted as formatting information; the text file which results after the formatting information has been interpreted is in a format that can be sent to a printer or display device. A user may generate the input file using an editor, a word processor, or the editing function of the text processor or batch compiler. Text processors and batch compilers are capable of producing high-quality, complex, structured documentation. They can be less expensive than high-end DTPs, but they are also more difficult to use since the user needs to understand the formatting codes to generate or modify a document. In the past, the portability of documents produced by this subdomain was an advantage since it permitted a document to be produced compatibly on different machines, but that advantage is now available via the Standard Generalized Markup Language (SGML).

Examples: Eroff, Documenter's Workbench

Applications: If a project lacks the funds for a high-end DTP but has a requirement for high quality complex documentation involving multiple authors, a text processor or batch compiler may be useful. The most likely project would already have personnel with expertise in this subdomain or would allow personnel plenty of time to climb the learning curve. The trend is away from this subdomain and toward the DTPs.

Word Processor

Word processors provide easy text editing and are typified by full screen editing, cut and paste, word wrap, justification, search/replace, and spell checking.

Examples: Microsoft Word, WordPerfect, Ami Pro

Applications: Use word processors for creation of input files to DTP and text processors and for creation of structured documents of less than a few hundred pages.

Appendix B

System-Related Non-Functional Requirements

Non-functional system requirements include some or all of the following:

a. performance
    i. time
    ii. space
b. operational environment
    i. hardware platform
    ii. software platform
    iii. external software interoperability
c. standards conformance
d. general characteristics
    i. reliability
    ii. robustness
    iii. accuracy of data
    iv. correctness
    v. security
    vi. privacy
    vii. safety
    viii. portability
    ix. modifiability and extensibility
    x. simplicity versus power

Process-Related Non-Functional Requirements

Non-functional process requirements include some or all of the following:

a. development time
b. development cost
c. software life cycle constraints
d. system delivery
    i. extent of deliverables
    ii. deliverable formats
e. installation
    i. developer access to installed environment
    ii. phase-in procedures to replace existing system
f. standards conformance
g. reporting
h. marketing
    i. pricing
    ii. target customer base
i. contractual requirements and other legal issues

Personnel-Related Non-Functional Requirements

Non-functional personnel requirements include some or all of the following:

a. for developers:
    i. credentials
    ii. applicable licensing, certification
b. for users:
    i. skill levels
    ii. special accessibility needs
    iii. training

Black Box - The person adapting the component should only need to understand the interface to the component.

Flexible - it should be possible to induce a wide range of adaptations (functional as well as behavioral).
Reusable - one should be able to reuse the code written to adapt a component.
Architecturally Aware - A component-based application should have some global concept of architecture, and the specification and/or implementation of the adaptation should be visible at this architectural level.
Configurable - An adaptation mechanism should be capable of applying the same particular adaptation (a generic part) to a particular set of target characteristics (the specific parts). Synonymous with *Repeatable* or *Template-drive*.

Present results with accompanying tables and graphs. Characterize the patterns and quality of the results and estimate their accuracy and precision. Use analytical graphics

References

[1] Amulya, J., “Critical Moments Reflection.” Unpublished Publication for the Center for Reflective Community Practice, MIT Department of Urban Studies and Planning, April 2004.

[2] Arne, J., "Multimodal Interaction for Information Appliances.", Linköpings Universitet, Department of Computer and Information Systems, Available at http://www.ida.liu.se/~nlplab/miina/ir-processing.shtml, 2004.

[3] Becker, H., “Why School Is a Lousy Place to Learn Anything in.” Reprinted in R J Burgess." (Ed.) Howard Becker on Education Buckingham; Open University Press, 1963.

[4] Bender, W., & Chesnais, P., "Network Plus." In Proceedings of the SPIE Electronic Image Devices and Systems Symposium, Los Angeles, CA, January 1988.

[5] Benedikt, M., “Cyberspace: Some Proposals.” In Michael Benedikt, (Ed.), Cyberspace: First Steps. Cambridge, Massachusetts: MIT Press, 119-224, 1991.

[6] Berry, D. C., (Ed.) "How Implicit is Implicit Learning?" Oxford; Oxford University Press, 1997.

[7] Blankinship, E., Smith, B., Bender, W. & Holtzman, H., “Closed Caption, Open Source.” Accepted for Upcoming BT Technology Journal: London 2004

[8] Bolton, G., "Reflective Practice: Written and Professional Development." London: Sage Publications, Inc., 2001.

[9] Brittle, J., Boldyreff, C., "Self Organizing Maps in Visualizing Large Software Collections." VISSOFT, Visualizing Software for Understanding and Analysis, 2003.

[10] Brown, J., et al., "Visualization: Using Computer Graphics to Explore Data and Present Information." John Wiley & Sons: New York, 1995.

[11] Bruggemann, B., & Holtz, K., "Dynamic Interactive Documents in Hydro-engineering Projects." In Proceedings of 4th International Conference on Hydroinformatics, Iowa City, IA, USA, 2000.

[12] Campagnoni, F.R., & Ehrlich, K., "Information Retrieval Using a Hypertext-Based Help System." ACM Transactions on Information Systems, 7(3), 271-291, 1989.

[13] Card, S.K., Mackinlay, J.D., & Shneiderman, B., (Eds.), "Readings in Information Visualization: Using Vision to Think." San Francisco: Morgan Kaufmann, 1999.

[14] Card, S.K., Robertson G.G., & Mackinlay, J.D., "The Information Visualizer, an Information Workspace." In Proceedings of the CHI '91 Conference, ACM, 181-188, 1991.

[15] Chalmers, M., "Using a Landscape Metaphor to Represent a Corpus of Documents." In A.U. Frank and I. Campari, (Eds.), Spatial Information Theory: A Theoretical Basis for GIS: European Conference, In Proceedings of COSIT '93. London: Springer, 377-390, 1993.

[16] Chen, C., "Information Visualisation and Virtual Environments." London: Springer, 1999.

[17] Chen, C., & Czerwinski, M., "Spatial ability and Visual Navigation: An Empirical Study." The New Review of Hypermedia and Multimedia 3, 67-69, 1997.

[18] Cooper, M., et al., "Information Landscapes." MIT Technical Note, April, 1994.

[19] Modjeska D.K., "Hierarchical Data Visualization in Desktop Virtual Reality." Ph.D. Thesis, University of Toronto, 2000.

[20] Davis, R., Shrobe, H., & Szolovits, P., AI Magazine, 14(1):17-33, 1993.

[21] Dieberger, A., "Browsing the WWW by Interacting with a Textual Virtual Environment -- A Framework for Experimenting with Navigational Metaphors." In Proceedings of ACM Hypertext '96. ACM, 170-179, 1996.

[22] Eraut, M. "Developing Professional Knowledge and Competence." London: Falmer, 1994.

[23] Foley, J., & Ribarsky, W. "Next-Generation Data Visualization Tools. In Scientific Visualization.", (Eds.)Rosenblum, Earnshaw, Encarnacao, Hagen, Kaufman, Klimenko, Neilson, Post, and Thalmann, Academic Press, pp. 103-127, 1994.

[24] Foulger, D., "Bridging Media: Computers and Human Communication.", Available at http://evolutionarymedia.com/papers/Fut-cmc1.htm, 1992.

[25] Gershon, N., Eick, S. G., "Visualization's New Tack: Making sense of information." IEEE Spectrum 32, 38-56, 1995.

[26] Gruber., T. R., "A Translation Approach to Portable Ontologies. Knowledge Acquisition." 5(2):199-220, 1993.

[27] Honkela, T., "Neural Nets that Discuss: a General Model of Communication based on Self-Organizing Maps." In Proceedings of ICANN'93, S. Gielen & B. Kappen (eds.), London: Springer, pp. 408-411, 1993.

[28] Honkela, T., "Learning to Understand: General Aspects of Using Self-Organizing Maps in Natural Language Processing." D. Dubois (Ed.), In Proceedings of CASYS'97, Woodbury, New York: Am. Institute of Physics, pp. 563-576, 1997.

[29] Honkela, T., Leinonen, T., Lonka, K., & Raike, A., "Self-Organizing Maps and Constructive Learning.", In Proceedings of ICEUT'2000, IFIP, Beijing, August 21-25, pp. 339-343, 2000.

[30] ISO13250, "Maps Paradigm, ISO/IEC 13250:2000 'Topic Maps' Standard" Available at http://www.topicmaps.org, 2000.

[31] ISO13250, International Organization for Standardization: ISO/IEC 13250:2000, Information Technology - "SGML Applications - Topic Maps." Geneva: ISO, 2000.

[32] Jupiter Media, "What's In a Topic Map?" Available at http://webreference.com/xml/column77/2.html, June 2004.

[33] Kaski, S., Honkela, T., Lagus, K., & Kohonen, T., "WEBSOM - Self Organizing Maps for document collections." Neurocomputing 21, 101-117, 1998.

[34] Kohonen, T., "Self-Organizing Maps." Berlin; Heidelberg; New-York: Springer; 1995

[35] Kohonen, T., "Self-Organized Formation of Topologically Correct Feature Maps." Biological Cybernetics, 43:59 – 69, 1982.

[36] Krista, L., Kaski, S., Honkela, T., & Kohonen, T., "Browsing Digital Libraries with the Aid of Self-Organizing Maps.", In Proceedings of the Fifth International World Wide Web Conference WWW5, May 6-10, Paris, France, volume Poster Proceedings, pages 71-79. EPGL, 1998.

[37] Le Grand, B., et al., "XML Topic Maps and Semantic Web Mining. Semantic Web Mining Workshop." In Proceedings of ECML/PKDD Conference, Freiburg, Germany, 2001.

[38] Lin, X., Hmelo, C., Kinzer, C. K., & Secules, T. J., "Designing Technology to Support Reflection." Educational Technology Research & Development, 43-62, 1999.

[39] Loh, B., et al "Developing Reflective Inquiry Practices: A Case Study of Software, the Teacher, and Students" In K. Crowley, C. Schunn, & T. Okada (Eds.), Designing for Science: Implications from Everyday, Classroom, and Professional Settings. Mahwah, NJ: Erlbaum, 2001

[40] Lonka, K. "Explorations of Constructive Processes in Student Learning." Doctoral Dissertation. Department of Psychology, University of Helsinki, 1997.

[41] Lynch, K., "Image of the City." Cambridge, Massachusetts: The MIT Press, 1960.

[42] Matthew C., "Using a Landscape Metaphor to Represent a Corpus of Documents." In Proceedings of COSIT 93, Conference on Spatial Information Theory, pp. 377-390, 1993.

[43] McCormick, B., Defanti, T., & Brown, M. "Visualization in Scientific Computing." Computer Graphics 21, 6, November 1987.

[44] McDrury J., "Developing Tools for Reflective Practice. Different Approaches: Theory and Practice in Higher Education." In Proceedings of HERDSA Conference, Perth, Western Australia, 8-12 July 1996.

[45] Nachmias, F., & Gilad, A. "Needle in a Haystack: Searching for information on the World Wide Web." Journal of Research on Technology in Education, 34(4), 475-487, 2002.

[46] Newcomb, S., & Biezunski M., "Topic Maps go XML." Working Group 3 of Sub-Committee 34 of ISO/IEC's Joint Technical Committee (JTC 1), 1999

[47] Nisbett R. E., & Wilson, T. D., "Telling More Than We can Know : Verbal Reports on Mental Processes." Psychological Review 84, 231-59, 1977.

[48] Palmer, S.B., "The Semantic Web: An Introduction.", Available at http://infomesh.net/2001/swintro/, May 2004.

[49] Papert, S., "Mindstorms: Children, Computers and Powerful Ideas." Basic Books, New York, 1980.

[50] Park, J., & Hunting, S., (Eds) "XML Topic Maps.", Addison Wesley, 2003.

[51] Passini, R., "Wayfinding in Architecture." New York: Van Nostrand Reinhold, 1984.

[52] Pepper, S., "Navigating Haystacks and Discovering Needles." Available at http://www.ontopia.net/topicmaps/materials/mlangart.pdf, May 2004

[53] Polanyi, M. (1958/1962), "Personal Knowledge.", Corrected Edition, Routledge, London, 1962

[54] Raelin, J, "I Don't Have Time to Think! Versus The Art of Reflective Practice." Reflections: The SoL Journal on Knowledge, Learning, and Change, 4 (1): 66-70, 2002

[55] Rath, H., & Pepper, S. "Topic maps: Knowledge Navigation Aids.", In: Goldfarb, C.F., Prescod, P. (Eds): XML Handbook, 2nd edition, Prentice Hall, Available at http://www.infoloom.com/gcaconfs/WEB/paris2000/S29-01.HTM, 1999

[56] Roper Stratch (2000). "Survey Highlights New 'Web Rage' Phenomenon." A Roper Stratch World wide Report. Available: Available at http://www.roper.com/news/content/news230.htm, 2000.

[57] Russell, T. & Munby, H., Reframing. "The Role of Experience in Developing Teachers Professional Knowledge." In D. A. Schön (ed.) The Reflective Turn. Case Studies In and On Educational Practice, New York: Teachers Press, Columbia University, 1991.

[58] Sanders, B., "Essentials of Good Technical Documentation.", University of Rochester Computer Science Department, Available at http://www.cs.rochester.edu/research/vision/doc/writing-good-docs.shtml, 2003.

[59] Schneiderman, B., "The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations." In Proceedings of the 1996 IEEE Visual Languages, Boulder, CO, pp. 336-343, 1996.

[60] Schön, D. "Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions." San Francisco: Josey-Bass Publishers, 1987.

[61] Schön, D. "The Reflective Practitioner: How Professionals Think in Action." New York: Basic Books, 1983.

[62] Senge, P., Kleiner, A., Roberts, C., Ross, R., & Smith, B., "The Fifth Discipline: Fieldbook, Strategies and Tools for Building a Learning Organization." Doubleday, New York, 1994.

[63] Sheppard, S.R.J. "Guidance for Crystal Ball Gazers: Developing a Code of Ethics for Landscape Visualization." Landscape & Urban Planning (Special Issue), Volume 54, 25, Pages 183-199 May 2001

[64] Small, D., "Navigating Large Bodies of Text." IBM Systems Journal 35, Nos. 3&4, 514-25, 1996.

[65] Smith, R. "Art In Review; Jennifer and Kevin McCoy.", The New York Times. New York: 36, 2002.

[66] Terry, M. & Mynatt, E.D., "Recognizing Creative Needs in User Interface Design." In Proceedings of the Fourth Conference on Creativity and Cognition. October 2002.

[67] Themescape Product Suite, Available at http://www.cartia.com/products/index.html , June 2004.

[68] United States Air Force, "Software Metrics Capability Evaluation Guide." The Software Technology Support Center (STSC) Available at http://www.stsc.hill.af.mil/resources/tech_docs/?doc_tech/Doc_Rpt.doc, 2004

[69] Usher, R. et al "Adult Education and the Postmodern Challenge.", London: Routledge, 1997.

[70] Wenig, J., & Silverman, B.G., "The Development and Maintenance of An Organization's Most Important Asset: Knowledge.", Atlantic Resources Corporation, for National Center for Education Statistics, U. S. Department of Education, February, 1993.

[71] Westberg, J. & Jason, H. “Fostering learners’ Reflection and Self-Assessment.” Family Medicine 26, 278-82, 1994.