SE-3D: A CONTROLLED COMPARATIVE USABILITY STUDY OF A VIRTUAL REALITY SEMANTIC HIERARCHY EXPLORER

Charles-Antoine Julien

School of Information Studies

McGill University, Montreal

January 2011

A thesis submitted to McGill University
in partial fulfillment of the requirements of the
degree of Ph. D. in Information Studies

Abstract

Keyword searching (e.g., Google or Yahoo!) is based on uncontrolled vocabulary matching which often produces large and noisy result sets. This can waste the time of the searcher who has to sift through long lists of often irrelevant information. The Semantic Web initiative aims to address this issue and includes the description of content using controlled ontologies (i.e., sets of descriptive terms and their relations). Ontologies are partly hierarchical structures too large to display on a single computer screen and thus difficult for searchers to explore efficiently.

In an attempt to address these issues, this research has developed and tested Subject Explorer 3D (SE-3D): an information visualization (IV) virtual reality (VR) information retrieval (IR) application based on the metaphor of exploring a physical space. SE-3D aimed to facilitate the visual exploration of information by offering searchers an interactive representation of the subject structure found in the Library of Congress Subject Headings (LCSH). SE-3D is a visual subject ontology navigation tool integrated with keyword searching and relevance ranking of a realworld information collection.

SE-3D was tested by 24 undergraduate students during a repeated measures within-subject experiment. As compared with a text-only baseline, SE-3D produced an advantage in accuracy. Participants were more patient with SE-3D, they preferred it and perceived it as more useful. The application used a new technique to manage hundreds of overlapping textual labels in virtual reality, and offered a novel integration of explorative and specific keyword searching. The analysis of the collection revealed that subject assignments followed a power law; the top 1% most assigned subjects contained over 58% of the collection and 65% of non-empty subjects contained a single document.

The findings suggest it is possible to extract additional value from organized collections by offering untrained users a reconstructed subject structure integrated with keyword searching. This research is significant for the development and testing of improved bridges between information organization and IR, and interactive information visualization.

Résumé

La recherche d'information textuelle à l'aide de mots clés (e.g., Google, Yahoo!) est basée sur la correspondance du vocabulaire utilisé et produit souvent de nombreux résultats à faible pertinence. Le chercheur perd son temps à examiner de longues listes de résultats contenant beaucoup d'information sans rapport avec le sujet. L'initiative du Web sémantique a pour objectif de pallier à ce problème en décrivant l'information à l'aide d'ontologies contrôlées (c.-à-d. des ensembles de termes descriptifs et leurs relations). Toutes les ontologies sont en partie de vastes structures hiérarchiques trop grandes pour être affichées à l'écran, elles sont donc difficiles à explorer de façon efficiente.

Développé dans le cadre de cette recherche, le Subject Explorer 3D (SE-3D) est un logiciel de visualisation de l'information (VI) en réalité virtuelle (RV) basé sur une métaphore d'exploration de l'espace physique. SE-3D tente de faciliter l'exploration visuelle de l'information en offrant aux chercheurs une représentation visuelle de la structure des vedettes-matières de la bibliothèque du Congrès (LCSH). SE-3D est un outil visuel de navigation d'ontologies sémantiques fortement intégrés avec la recherche par mots clés et le classement par pertinence d'une réelle collection d'information.

SE-3D a été testé par 24 étudiants au baccalauréat durant une expérience à mesures répétées intra-sujet. En comparaison avec un système purement textuel équivalent, SE-3D a produit une plus grande précision. Les participants se sont révélés plus patients avec SE-3D, ils l'ont préféré et perçu comme étant plus utile. SE-3D utilise une nouvelle technique de gestion des chevauchements entre étiquettes textuelles en RV, et offre une intégration novatrice de la recherche par mots clés permettant de filtrer la structure sémantique. L'analyse de la collection a démontrée que les affectations des termes LCSH au sein de la collection suivent une loi de puissance. Par exemple, 1% des sujets les plus affectés contiennent plus de 58% de la collection, et 65% des sujets non vides contiennent un seul document.

Ces résultats suggèrent qu'il est possible d'extraire une valeur ajouté provenant des collections organisées en offrant aux utilisateurs novices une structure sémantique

reconstruite intégrée avec la recherche par mot clé. Cette recherche est significative pour le développement de ponts entre les structures d'organisation de l'information et la recherche d'information, ainsi que la visualisation de l'information interactive.

Table of Contents

List of Figures

List of tables

List of Acronyms and Terms

3D: 3-dimentional, having width, height and depth

CV: Controlled Vocabulary

DDC: Dewey Decimal Classification

Graph: also called network plot, general study of visually representing objects and their relations using lines between them. These relations can be undirected or directed meaning they can either go both ways or only from one object to the other.

HCI: Human-Computer Interface.

IR: Information Retrieval

IV: Information Visualization

LCSH: Library of Congress Subject Headings

LIS: Library and Information Science

MARC: MAchine-Readable Cataloguing, format standard for the storage and exchange of bibliographic records.

MeSH: Medical Subject Headings

Node: single object part of a hierarchy.

OPAC: Online Public Access Catalogues

Precision: ratio of the number of relevant items in a result set divided by the total number of items in the result set

SE-3D: Semantic Explorer in Three-Dimensions

Semantic Hierarchy: hierarchical part of an ontology, also called subject hierarchy, also called subject tree

SKOS: Simple Knowledge Organization System

Subject Heading (Term): word(s) which describes a semantic content or "what an intellectual work is about".

SW: Semantic Web (initiative)

TAM: Technology Acceptance Model (Davis, 1989; Davis, Bagozzi & Warshaw, 1989)

TC: (automatic) Text Classification

TREC: Text REtrieval Conference

VR: Virtual Reality

Acknowledgements

This research has brought me in contact with thought-provoking people of excellence. I am truly thankful for their time. Specifically, McGill professors Jamshid Beheshti, Andrew Large, John, E. Leide, France Bouthillier, and Catherine Guastavino. I thank them for their repeated gifts of patience, guidance and knowledge. The patience and knowledge of the School of Information Studies administrative assistants (Kathryn Hubbard, Ancy Joseph, Susann Allnutt) was often critical and always appreciated.

Financial support was generously provided by the "Fond Québécois pour la Recherche sur la Société et la Culture" (FQRSC). Participation in research projects led by France Bouthillier and Andrew Large/Jamshid Beheshti also provided valuable experience and remunerated work.

Jacques Julien and France Turcotte have generously provided critical support. Normand Hénault, Heather Brydon, Stella Drury and Paul Julien agreed to test and pilot the application. Heather Brydon also performed the experiment data validation. They all deserve my most sincere appreciation.

This has been a six year journey for which I am thankful. I enjoyed the diversity of experiences in research, teaching and service. I have obviously aged, I like to think I've also matured and improved. Time will tell.

Chapter 1: PROBLEM STATEMENT

Web search engines such as Google have difficulty ranking non-fiction books the way they rank Web pages or scholarly articles. Google Ranking draws much of its competitive edge from ordering of query results based on inbound-outbound link/citation analysis[1]. This data are often not available for books because many knowledge domains do not often cite books such as introductory manuals and knowledge domain overviews (Larivière, Archambault, Gingras & Vignola-Gagné, 2006). Searching for books using a search box amounts to little more than blunt word matching which often produces large result sets difficult to rank (Fox, Das Neves, Yu, Shen, Kim & Fan, 2006). This wastes the time of the searcher who has to sift through too much irrelevant information; as well, for engines that rely on keywords, bibliographic records typically provide few keywords beyond the title and subject headings when available.

The Semantic Web (SW) initiative defines a road map for the future of digital information retrieval based on the meaning of information items instead of strictly matching words (Berners-Lee, Hendler & Lassila, May 2001). SW hopes to produce online search tools that can understand the meaning of information since current "user interfaces often take little or no advantage of (…) semantics" (Fluit, Sabou & van Harmelen, 2003, p. 36).

SW includes the description of content using recognized ontologies which are sets of descriptive terms and their relations. Ontologies are partly hierarchical (Noy & McGuiness, March 2001) meaning that they can be conceptualized as broad parent concepts being subdivided into more and more specific child concepts. These semantic hierarchies are often too large to display on a single computer screen and thus difficult for searchers to explore efficiently (Akrivi, Elena, Constantin, Georgios & Costas, 2006; Parsia, Wang & Goldbeck, 2005; Plaisant, Grosjean & Bederson, 2002).

---

[1] see http://www.google.com/corporate/tech.html

The SKOS[1] initiative aims to translate existing ontology knowledge into a vocabulary compatible with the Semantic Web. Library of Congress Subject Headings (LCSH) make up one of the largest and most widely used (Chan & Hodges, 2007; Taylor & Miller, 2006) list of subject terms and their relations. This large[2] list of terms acts as controlled vocabulary (CV) used to describe the semantic content of information items.

Assuming knowledge of how information is organized helps users become better searchers, the semantic Web has a need for a large semantic hierarchy exploration interface (Fox, Urs & Cronin, 2002). With the transfer of existing library ontologies to the semantic Web, an LCSH exploration application becomes a representative design case for future SW development. This research is about facilitating access to organized collections by using existing information organization structures.

As described in the following sections, hierarchy navigation and exploration has been a problem of interest for ontology visualization research. LIS professionals and researchers have also accumulated knowledge concerning people's interactions with semantic structures that organize library collections.

1.1 Ontology Visualization Tools

Most ontology visualization systems are 2-dimentional and offer a Windows Explorer (Akrivi, Constantin, Georgios, Costas & Eugenia, 2007, p. 6) tree or outline view (see Figure 1.1)

[1] see http://www.w3.org/2004/02/skos/
[2] The set used in this research contained over 200,000 concepts and relations between them.

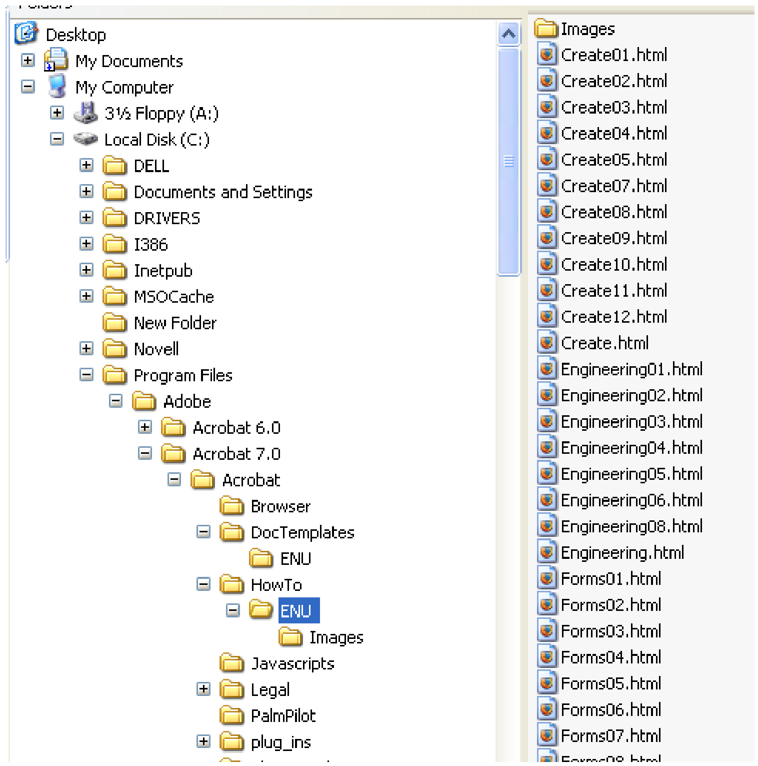


Figure 1.1: Hierarchy Visualization using Windows Explorer Tree or Outline View

Although ubiquitous and well understood by users, outline views suffer from known problems of inefficient screen usage (Akrivi, *et al.*, 2006; Parsia, *et al.*, 2005; Plaisant, *et al.*, 2002; Robertson, Mackinlay & Card, 1991) especially troublesome for large semantic hierarchies such as LCSH. Outline views hide branches to fit the screen forcing users to repeatedly expand, scroll and collapse branches causing "disorientation produced by the change in the overall layout" (Freitas, Luzzaerdi, Cava, Winckler, Pimenta & Nedel, 2002, p. 10). This is a problem if the hierarchy contains as little as a few hundred nodes (Bladh, Carr & Scholl, 2004; Bruls, Huising & van Wijk, 2000; van Ham & van Wijk, 2003).

Additionally, because they force a strictly piecemeal exploration of the hierarchy (Turo, 2003), outline views do not easily allow users to gain a global overview of the subject structure (Kumar, Plaisant & Shneiderman, 1997, p. 104). Finally, the very act of representing a tree in two dimensions effectively *flattens* the structure causing a loss of information concerning semantic relations between items; in other words, within each hierarchy level, 2D tree representations arbitrarily list items sequentially (e.g., alphabetical order) as opposed to representing more complex links between items such as semantic similarity.

1.2 Library Subject Hierarchies

Since the early 1980s library online public access catalogues (OPACs) have offered access to their collections using keyword matching and CV subject headings (e.g. LCSH) browsing (Calhoun, 2006). Semantic hierarchy browsing interfaces have long been offered by OPACs; thus their development and usage history is a promising source of information concerning the problem of large semantic hierarchy navigation by the general public.

Library and information science (LIS) research has shown that OPACs users, especially novices, make little or no explicit use of semantic CV (e.g., LCSH) when searching for information (Borgman, 1996; Drabenstott & Weller, 1996b; Iglesias & Stringer Hye, 2008; Larson, 1991). Research has shown that when using natural language searching (i.e., keyword in all fields), subject or topical search tasks (e.g., coffee production) will often produce large numbers of retrieved bibliographic records with which users have very little patience (Van Pulis & Ludy, 1988; Wiberley & Daugherty, 1988). Lack of user knowledge of CV is a problem for searchers who may miss valuable information buried amongst too many often irrelevant items. Reference librarians explicitly use controlled vocabulary because they know it addresses this issue by yielding more relevant (i.e., less noisy, more precise) subject searches (Chen, Houston, Sewell & Schatz, 1998, p. 597).

Lack of explicit public use and knowledge about CV is also a problem for libraries that have difficulty demonstrating the value created by an expensive information resource not explicitly used or understood by its patrons (Iglesias & Stringer Hye, 2008, p. 13). This is especially critical since “as a means of retrieving information libraries have to compete with the internet” (Grun, Gerken, Jetter, Konig & Reitener, 2005, p. 174), and although quality CV is offered almost exclusively by libraries, unfortunately this competitive edge is not adequately communicated by the traditional online library catalogue.

1.2.1 Problems with Library Subject Browsing

The inadequacy of CV subject browsing features offered by OPACs has been reported since the early 1980s (Markey, 1985; Matthews, Lawrence & Ferguson, 1983) and this problem is still prevalent today (Mercun & Zumer, 2008; Papadakis, Stefanidakis & Tzali, 2008). Traditional OPACs offer a text-only list of alphabetically ordered subject headings which “puts a heavy burden on users to formulate the right query” (Driessen, Jacobs & Huijsen, 2006, p. 217). Research findings from two separate studies demonstrate that some users do not understand what they are looking at (Janosky, Smith & Hildreth, 1986; Markey, 1989), and they simply “do not browse subdivided subject headings in alphabetical browsing displays” (Drabenstott & Weller, 1996a, p. 721). These types of human-computer interfaces (HCI) are problematic because as the number of list entries increases, scrolling and scanning can become tedious, unproductive and error-prone (Leuski & Allen, 2000; Sebrechts, 2005).

Alphabetical ordering of results is also problematic since "there are nearly always better sequences than alphabetical" (Tufte, 1983, p. 178). Alphabetical order is useful for looking up individual items, but "not for seeing patterns across items according to adjacency” (Hearst, 1999, p. 303). Alphabetical order assumes the searcher can recall the beginning of the spelling of the target subject which "is analogous to one of the most familiar complaints about dictionaries: sometimes you need to know how to spell a word in order to look up its correct spelling in the dictionary" (Dushay, 2004, p. 2).

Assuming the searcher can provide a valid description of the topic, a relevance ranked list could replace the alphabetical listing; however, this option is not without its own set of known issues. Research in Web searching behavior has shown sequential lists often come with no indication of the system’s evaluation of each document and no clear indication of the relationships that may exist among the retrieved documents (Iwayama & Tokunaga, 1995). Ranked lists can be confusing because the results on different topics or different aspects of the same topic are sometimes mixed together (Becks, Seeling & Minkenberg, 2002). Internet search engine users can miss relevant information because they usually inspect at most 20 to 30 documents before quitting the search (Allan, Leuski, Swan & Byrd, 2001; Chen, 2000; Silverstein, Marais, Henzinger & Moricz, 1999). Long lists of results do not provide indications of the scope of the collection: the quantity of information, the density in each subject area or the general audience for the material (Dushay, 2004, p. 1).

How should interfaces support searchers who are unfamiliar with a specific subject area and its terminology? The design of online searching interfaces that support domain novices to express their information need using the "vocabularies" (Ding, Chowdhury, Foo & Qian, 2000, p. 1190) from the subject area has become one of the most pressing questions in the information retrieval field (Chen, Yim, Fye & Schatz, 1995).

1.3 Problem Statement Summary

The Semantic Web relies on large semantic CV hierarchies difficult to explore on a computer screen. Subject CV hierarchies are a part of LIS and a defining characteristic of library information organization, yet searchers are not aware of their existence. The traditional textual list of alphabetized subject terms has not adequately supported OPAC searchers to learn and explicitly use controlled vocabulary, and ranked lists are not a promising alternative.

Existing semantic hierarchy browsing interfaces can be improved to better communicate the organization of the information to searchers. There are indications that library catalogue users respond better to graphic than text and expect highly intuitive interfaces and convenience (Mercun & Zumer, 2008, p. 246); therefore, this research aims to facilitate the exploration of organized collections by providing a novel interface for semantic hierarchy navigation based on advances in the field of information visualization (IV).

1.4 Anticipated Contributions

The anticipated contributions of this research to scientific knowledge concern the domains of information organization, information visualization, and usability testing. The corpus analysis aims to provide an automatic algorithm to reveal patterns in the distribution of an information collection within its organization structure. These distribution patterns might facilitate the browsing of complex information structures. The development of a novel information visualization tool will likely produce novel human-information interaction techniques which will add to the body of knowledge in interactive visual interfaces for information retrieval. Formative and controlled testing of this novel software will become one of few reports of systematic usability testing of information visualization for information retrieval. These anticipated contributions aim to improve the design of interactive visual tools for information exploration and searching.

Chapter 2: THESIS

> *"...libraries offer the perfect submarket to slowly establish alternative forms (of online search tools, and) thus could lead to a paradigm shift." (Grun, et al., 2005, p. 184)*

Like Grun (2005), this research assumes the library is a promising information submarket that can provide innovative access to its unique and valuable semantic data; furthermore, it is a promising design case study for future applications of Semantic Web ontologies. A real-world complex example of semantic ontology data are the LCSH organization of library collections which places each information item within a semantic hierarchy of broad to narrow subject terms. Hierarchy visualization is "one of the most mature and active branches in information visualization" (Chen, 2004, p. 90), and aims "to provide users (with) semantic views to help them understand the semantic relations" (Nguyen & Zhang, 2006, p. 981).

There are many hierarchy visualization interfaces but most are 2-dimentional (2D) even though a 3-dimentional (3D) hierarchy representation may be more appropriate. 3D hierarchies can make more efficient use of screen space (Robertson, *et al.*, 1991) and are better suited to convey hierarchical relationships between levels (Bladh, et al., 2004). Using the 3rd dimension to convey hierarchy depth (van Ham & van Wijk, 2003) allows designers to make use of 2D IV techniques (Chen, 2000; Lin, 1997) for each individual hierarchy level. 2D IV techniques can convey more complex semantic relationships between items as opposed to alphabetical ordering.

This leads to the thesis of this research:

> *An interactive and searchable 3D visualization of a large semantic hierarchy may provide a performance advantage and be preferred by searchers as compared with an equivalent text-only search box and ranked result list.*

This research will attempt to answer the following specific question:

> *Are there differences between a 3D IV system and a text-only subject browser in terms of user performance and experience for undergraduate students performing IR tasks? Performance is measured by completion time and accuracy. Experience is measured by perceived speed, usefulness, ease of use and preference of the system.*

Chapter 3: THEORETICAL FRAMEWORK

As shown in Figure 3.1, the theoretical framework situates this research in the context of interactive information retrieval (section 3.1) which includes the well established LIS practice of subject indexing and the current development of the semantic Web and its ontologies.

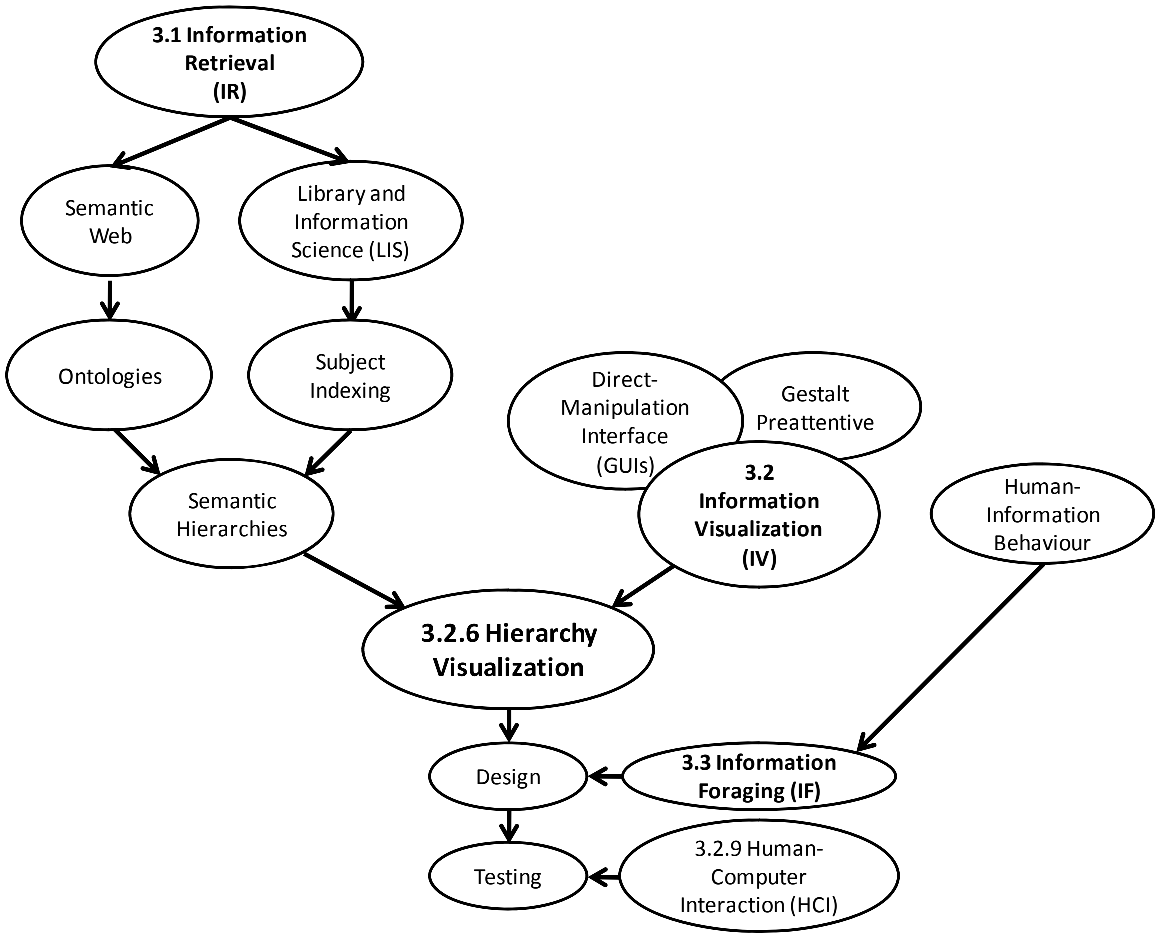


Figure 3.1. Theoretical Framework Overview

SW ontologies and subject indexing contain hierarchical structures that have been extensively studied by information visualization researchers (section 3.2). IV has foundations in empirical evidence from the mass adoption of direct-manipulation visual interfaces, the Gestalt descriptive model of visual perception and preattentive cognitive processing.

The design of an IV interface for semantic hierarchies is guided by the human-information behaviour model of information foraging (section 3.3). Finally, user testing of the innovative IV application is supported by an extensive review of IV usability studies (section 3.2.8).

3.1 Information Retrieval

Information retrieval (IR) is the interdisciplinary science of humans searching for information (Chowdhury, 1999; Chu, 2003; Meadow, Boyce & Kraft, 2007). The domain is multi-disciplinary and based on computer science, mathematics, library science, information science, information behaviour, cognitive psychology, linguistics, statistics and physics. IR is pertinent for this research because it provides definitions of known-item searching vs. browsing or exploratory searching.

HCI and human-information behaviour knowledge suggests it is useful to establish a set of usage scenarios which a tool design should support (Winckler, Palanque & Freitas, 2004). Borner *et al.* (2002) identified usage scenarios of IR interfaces and Shneiderman *et al.* (2005) describe similar phases of the search process. Together they form the following list of IR usage scenarios:

1. Formulate query
2. Identify the composition of search results
3. Understand the interrelations between retrieved documents
4. Refine the search
5. Gain an overview of the coverage
6. Browse

These usage scenarios are performed within the context of IR tasks most often categorized in a dichotomy between specific (also called *known item*) versus exploratory or browse search (Marchionini, 1995). A known item search is performed for clearly identifiable items (e.g., the capital of Laos, the book titled “War and Peace”) with a clearly identifiable success or stop criterion. A browse search is often associated with subject or topical searching performed when the user’s existing knowledge does not provide a clear definition of the information need (e.g., coffee production, dogs), and the success criteria is not apparent before beginning the search. In layman's terms:

> *(known item) searching implies that you have a good-to-perfect idea of what you want. Browsing implies that you will be able to recognize what you want when you see it. (Buzydlowski, White & Lin, 2002, p. 134)*

The result of browsing is often multiple facts and/or documents which taken as a whole provide information on the desired subject.

The known item vs. browse search dichotomy may not provide sufficient granularity to usefully classify all IR tasks. Shneiderman *et al.* (2000) describe a human-information interaction process where users begin the search process by considering their information needs and clarifying a task which can be classified into the following four general categories (Shneiderman & Plaisant, 2005, p. 562):

1. **Known-item search**: searching directly for a readily identifiable outcome. E.g., Find the capital of Canada.
2. **Extended fact-finding**: searching indirectly for relatively uncertain but replicable outcomes. E.g., Are there other books by the author of War and Peace?
3. **Open-ended browsing and problem analysis**: gaining an understanding of a general subject area. E.g., How can I fix my leaky faucet?
4. **Exploration of availability**: knowing what information is available where. E.g., What genealogical information does the library offer?

Specific search from Marchionini (1995) could be equivalent to the combination of known-item and extended fact-finding from Shneiderman & Plaisant (2005) because they both ask for clearly identifiable items and provide clearly identifiable success criteria. Exploratory search from Marchionini (1995) might arguably cover open-ended browsing/problem analysis and exploration of availability as described by Shneiderman & Plaisant (2005). Table 3.1 illustrates this correspondence.

Marchionini (1995)	Shneiderman & Plaisant (2005)
Specific	Known-item
	Extended fact-finding
Exploratory	Open-ended browsing + problem analysis
	Exploration of availability

Table 3.1: Search task types

Defining the elusive concept of browsing or exploratory searching is not a simple task and requires further description presented in the following section.

3.1.1 Browsing

> *"Browsing is natural because it coordinates human physical, emotive, and cognitive resources in the same way that humans monitor the physical world and search for physical objects." (Marchionini, 1995, p. 100)*

Although, as Marchionini states above, it may be "natural", there are many definitions of browsing and a unique standard does not exist. An integrative content analysis of various definitions suggests that browsing is:

- an **explorative** process (Chen, *et al.*, 1998; Heo, 2000; Lin, 1997; Marchionini, 1987; McAleese, 1989)
- where searchers **scan** information (Heo, 2000; Lin, 1997; McAleese, 1989)
- in order to **focus or refine** (Heo, 2000; Lin, 1997)
- a **broad or ill-defined** search need (Baeza-Yates & Rivabeiro-Neto, 1999; Lin, 1997)
- by inspecting the **structure or links** between information items as well as individual items (Chen, *et al.*, 1998; Heo, 2000; Lin, 1997; McAleese, 1989)
- with **no planning** (Marchionini, 1987, p. 69) or in a "vague and non-specific manner" (McDonald & Stevenson, 1996, p. 62).

Perhaps the most exhaustive definition of the concept is provided by Chang's Browsing model (Chang, 2005) which states that browsing is

> *"an examination of unknown items of potential interest by scanning or moving through an information space in order to judge the utility of the items, to learn about something of interest in the item, or to satisfy curiosity". This is a "fundamentally evaluative and inclusive" process which may "locate information not considered beforehand". Browsing can be goal driven or simply a "recreational activity" which often produces learning effects simply due to the "opportunity of encountering the unknown" (Chang, 2005, p. 73).*

Browsing is indeed associated with serendipity (Marchionini, 1987, p. 69) which refers to making fortunate information discoveries not actively sought at the time they are encountered (Reitz, 2004). Merton (1957) defined serendipity as an observation of a surprising fact followed by a correct 'abduction'. This explains the finding that "novelty (…) was found to be a key motivation in browsing" (Toms, Dufour & Hesemeier, 2004, p. 52). This concept is echoed by Simon (1996) who "suggests than information obtained along any particular branch of a search tree may be used in many contexts besides the one in which it was generated" (p. 127). Finally, a pure search for interesting or novel information is possible and "provides the mechanism for scientific discovery" (p.162).

Browsing is said to be appropriate when

- Users look for information that is easier to recognize than to describe i.e., recognition over recall (Bates, 1986a; Olston & Chi, 2003; Shneiderman, 1998)
- In cases where a great deal of information and context is obtained along the browsing path itself, not just at the final information item (Olston & Chi, 2003); in other words, the enquiry demands an exhaustive search at the expense of assessing irrelevant information (Lin, 1997, p. 41)
- There is a good organizational structure and related information items are often located near each other (Thompson & Croft, 1989)
- Users have difficulties in articulating their information needs (Belkin, Oddy & Brooks, 1982) or the problem is ill-defined (Marchionini & Shneiderman, 1988, p. 71) because they are not familiar with the content of the collection and thus they need to explore the collection (Buzydlowski, *et al.*, 2002; Marchionini, 1987; Motro, 1986)

This last point would make browsing ubiquitous since it has been argued that most everyday tasks can be characterized as ill-defined problems (Reitman, 1964, 1965). Some have suggested that “once search results are retrieved, the exploration of this collection is a browsing task” (Turetken & Sharda, 2005, p. 274).

This research uses the following definitions of IR task types:

- Known-item: search for a single identifiable fact with clear success criteria
- Browsing: any search which is not a known-item search. Synonymous with exploratory searching as described by (Marchionini, 1995)

The boundary between known-item searching and browsing is not clear; indeed it should NOT be since, as the following section describes, these two types of IR tasks are often used in conjunction during a single search session.

3.1.2 Integrate Search and Browsing

> *…to blur the unnecessary line between query and results (Shneiderman, et al., 2000, p. 62)*

IR systems should integrate both types of searching and allow the user to easily switch from one to the other. Findings from Sutcliffe *et al.* (2000b) show that one of the major determinants of search success was the user's persistence in iterative cycles of search and relevance evaluation of results. This suggests that “information exploration, or browsing, should be implemented not as a necessary but an integral part of the retrieval process” (Lin, 1997, p. 41) because of their complementary advantages (Jul & Furnas, 1997; Manber, Smith & Gopal, 1997).

Known-item and browse searching are tactics not performed independently since people may apply different mixes of tactics during a single search session (Marchionini, 1995; Pejtersen, 1988). According to Belkin *et al.* (1993), “information seeking behavior is characterized by movement from one strategy to another within the course of a single search episode” (p. 257). This behavior may be a natural result of human non-linear cognitive process where "interruptions, breaks, and pauses are characteristic of human mental activity" (Hornbaek & Frokjaer, 2003, p. 494).

Current IR systems are efficient word matching engines but they do not understand the meaning of the information they are searching and retrieving. This is one of the issues the Semantic Web hopes to address.

3.1.3 Semantic Web

The Semantic Web is a future generation of search tools that are able to understand the meaning of information items as opposed to strictly matching keywords (Berners-Lee, et al., May 2001). Several constituent parts are being developed to make this project a reality. One of these parts is a standard content description schema using predefined sets of terms (e.g., "War and Peace" is a title of the current document). The terms can also be used to describe content subject concepts ("War and Peace" is about the "Russian Napoleonic Campaign").

One of the objectives of the SW is to design search tools able to comprehend the semantic content of information items and make inferences using common relations between concepts stored in authoritative ontologies. Noy (March 2001) defines an ontology as a formal and explicit description of a domain, consisting of classes, which are the concepts found in the domain. Each class may have one or more parent classes, creating a broad to narrow concept hierarchy (e.g., "City" is a part of "Province or State").

The Simple Knowledge Organization System (SKOS) is a Semantic Web initiative which aims to transfer the concepts and relations stored in existing knowledge organization systems such a MeSH[1] and LCSH (Harper, 2006; Summers, Isaac, Redding & Krech, 2008). SKOS hopes to create large computer readable ontologies of concepts and their relationships (e.g., broad, narrow, related, used for). The ongoing maintenance and usage of subject ontologies are a long standing part of the LIS practice of subject indexing.

[1] See MeSH conversion case at http://thesauri.cs.vu.nl/eswc06/

3.1.4 Subject Indexing

The library science practice of subject indexing (Svenonius, 1986) describes the semantic contents of information items by assigning subject headings from a restricted and controlled list of vocabulary (CV). This laborious practice places each information item (e.g., book, article, etc.) within a hierarchical structure of broad to narrow subjects which provides valuable contextual information not offered by keyword search tools. Shiri, Revie and Chowdhury (2002) distinguish between two kinds of subject indexing term lists: a “standard” and a “search or end-user” thesauri (p. 115). The former is a strict list of terms assigned to individual works and the latter is enhanced with synonyms to assist end-users in finding alternative or additional search vocabulary.

The utility of subject indexing is illustrated by a search for the keyword “chicken” which would, at the very least, yield both works on the study of birds and recipes. The searcher must then weed through a long list of results to find relevant items, which contribute to the problem of information overload (Blair & Maron, 1985). In a library collection organized using LCSH, works about chickens can be found both under 'Ornithology' and 'Cookery'. Specifying the desired context using the appropriate subject term will likely produce a more relevant set of results meaning that it contains a higher proportion of useful information items about the desired context.

Subject indexing is a two-part process consisting of CV list maintenance and CV assignment to individual information items such as books or articles. Maintenance pertains to the continuous updating of the CV list in order to reflect the evolution of knowledge and the vocabulary used to describe it (e.g., new subjects are added, existing subject vocabulary may be changed). CV assignment is usually referred to as subject indexing by LIS professionals assessing individual information items and describing their semantic content using terms from a common CV list. For example, millions of items in thousands of LCSH organized collections have received one or multiple LCSH terms to describe their semantic contents. These LCSH terms and their relations form a large semantic ontology: the words used by most libraries in the world to describe knowledge.

Since 1898, LIS professionals have been maintaining LCSH which has become the standard list of CV subject terms used by most large libraries in the United States, and is used throughout the world (Chan & Hodges, 2007, p. 213). Despite "perennial criticisms" (Taylor & Miller, 2006, p. 350) centered around cost, scalability and consistency, LCSH is simply the "most widely accepted CV list in use in English-language libraries today" (Taylor & Miller, 2006, p. 350).

3.1.4.1 Advantages of Subject Indexing for Searchers

When searching for information users need to

> *"have a working knowledge of the system where the information is stored, in particular how to navigate through the information system (Chen & Dhar, 1990) and of how the information is organized or categorized. Second, they must have a knowledge of the subject of interest, in particular the vocabulary of the subject domain."(Chen, et al., 1998, p. 583)*

Controlled vocabulary and its hierarchical structure specifically address both these needs when it is communicated through adequate subject browsing interface features. First, CV ontologies provide an explicit organization of information which "disambiguates terms by placing them within a hierarchy, eliminating ambiguity" (Dushay, 2004, p. 4). Second, CV subject terms are useful for search tasks in unfamiliar domains since “the subject content structure itself gives the domain novice a key to subject entry” (Large, Beheshti & Cole, 2002, p. 833). Fox *et al.* (2006) describes this process:

> *As users are aided in breaking down complex information needs into parts, as they see what terminology is used for each of the subtopics that emerge during the work on partial solutions, and as they see how subtopics are related, they learn more about the collection and area of interest...(p. 56)*

These valid subject terms suggested by the CV structure mitigate "one of the major causes of failure in IR systems" (Ding, et al., 2000, p. 1191): vocabulary mismatch between the words known by the searchers and those recognized by the search tool. This issue stems from the fact that individual searchers are usually unaware of the many terms that might be used by different authors to describe the same subject (Bates, 1986b). Original search statements "typically consists of just a few terms germane to the topic, and it is often necessary to add variety to achieve an effective search" (Ding, et al., 2000, p. 1191). Lack of vocabulary control for title, abstract or full-text keywords “places onto the user the burden of finding any synonyms to the search terms chosen” (Larson, 1991, p. 210).

There are other reasons for users to employ CV subject terms during their searches. CV reduces the variety in natural language stemming from “variety in word forms (e.g., singular/plural, verb conjugations, etc.), syntactical variations (e.g., different word orders), and synonymy” (Bates, 1986b, p. 362). As well, it is known “that uncontrolled vocabulary fails to group related materials together” and as a result “much valuable material may be missed” (Bates, 2003, p. 14).

One may suspect the artificial and sometimes esoteric nature of CV might hamper its usage by novice searchers but research suggests otherwise. In a rigorous test on an online catalogue (Carlyle, 1989) user keywords matched a single LCSH 47% of the time and partial matches would have raised the figure to 74% (p. 44). Bates (2003) estimated that a third of first time subject searches exactly matched the assigned LCSH, and Drabenstott (1996a) reports “a little over half of the subject queries users entered into online catalogues exactly matched the catalogue's controlled vocabulary” (p. 722).

3.1.4.2 Browsing Subject Structures

One of the advantages of controlled vocabulary stems from its hierarchical structure of broad to specific concepts since “providing a subject hierarchy is a conventional way to help browse information in a digital library” (Zhu & Chen, 2005, p. 158). Edwards and Hardman (1988) showed users formed better mental models when a small hypertext collection was organized hierarchically as compared with one that allowed flexible networked access. Simpson (1989) reports that a hierarchical contents list was superior to an alphabetic index and concluded that searchers use elements of the visual structure to form mental maps of the document.

Supporting subject browsing by representing an ontology as a hierarchy is an "obvious approach" (Akrivi, *et al.*, 2006, p. 2) based on inheritance links (i.e., broad to narrower relationships between terms). This does not explicitly represent non-inheritance links such as related[1] or synonymous[2] terms because "these types of relationships between topics are not parent-child and therefore are not allowed in a strict" hierarchy (Dushay, 2004, p. 4). Explicitly representing non-inheritance links would provide a more complete subject structure but would quickly clutter the visual display and overwhelm the user (Akrivi, *et al.*, 2006; Julien & Cole, 2009). Visual clutter is a real issue for semantic hierarchy browsing interfaces because any useful ontology will likely contain a large number of terms; however, some subject terms are much more important than others.

Bates (2003) suggested that CV assignments by human indexers most probably follow Bradford's Law: few subjects are assigned to many items, most subjects are assigned to few. This essentially creates a few large groups or families of information items grouped under their respective subject terms. It is likely that most large groups "are not only large themselves but are also of interest to disproportionate numbers of users" and providing quick access to a "few large families may satisfy many users quickly" (p. 40). This suggests that browsing a large ontology could be greatly simplified by presenting only the few CV terms to which many items have been assigned.

Since the 1980s, LIS researchers have studied interactions between users and the subject ontologies offered by libraries (e.g., LCSH, MeSH) through their OPAC interfaces. As the following section shows, LIS research offers as an extensive source of empirical knowledge on subject browsing interface design. This LIS knowledge is likely relevant to Semantic Web ontology exploration tool designs.

---

[1] LCSH "see also" relations
[2] LCSH "see" or "used for" relations

3.1.5 Browsing Library Catalogue Subject Indexing

> *"The limited interface of an OPAC…fails to give users a qualitative sense of the resources of the library as a whole. A useful comparison is between two kinds of physical library; a closed stack and an open stack collection. In a closed stack collection…serendipity is eliminated." (Sanchez, Twidale, Nichols & Silva, 2005, p. 216)*

Sanchez (2005) refers to OPAC interfaces that offer keyword matching and alphabetical lists of CV subject terms. A *closed stack* collection refers to libraries whose holdings are hidden from patrons where access is provided strictly via library staff who require specific item references. These *closed stack* online library catalogues were mainly developed based on practices and technologies available in the early 1980s (Markey, 2007). Like the hidden holdings of closed stack collections these online search tools do not offer search suggestions and little or no browsing features. The systems wait for the user to adequately define the information need before a small subset of the collection can be retrieved.

As the only publicly available search technology, the once innovative OPACs were appreciated by a captive audience (Farber, 1984; Matthews, et al., 1983). This golden age of online catalogues was slowly hampered by numerous reports that users needed subject searching to be improved (Besant, 1982; Larson, 1991). After more than 40 years of research (Salton, 1989), "it is quite clear that traditional query-based search tools do not always satisfy the need of prospective users" (Lin, 1997, p. 41) and indeed problems associated with these system are well documented (Belkin & Croft, 1987; Borgman, 1996). It should be stated that current Web searching engines are also query-based search tools albeit with the added value of relevance ranking.

IR research (Korfhage, 1991) has shown that many IR user interfaces are suffering from problems concerning the user's perception and understanding of the dialog with the system. The searcher often does not understand what criteria the system uses to choose relevant results, and he/she cannot easily estimate if the information need can be satisfied by the content of the specific database (Hemmje, Kunkel & Willet, 1994, p. 249).

The current ubiquitous human-information interaction model based almost exclusively on keyword(s) entered in a search box followed by a linear textual result list was adopted by current generations of Web search engines from existing library OPACs (Smith, Czerwinski, Meyers, Robbins, Robertson & Tan, 2006). These private search tools competed with OPACs by providing storage and retrieval scalable to the vast amounts of Internet information, and more useful result ranking algorithms; however, the basic human-information interaction has not changed. Even the most popular Web search engines do not offer any initial suggestion as to the scope of the collection they cover and little or no browsing features beyond the long list of ranked results. These systems are almost exclusively textual and given the unprecedented availability of online digital information, "a text-based interface may become increasingly overwhelming to users, especially novices" (Smith, *et al.*, 2006, p. 797).

In semantically organized collections such as libraries and the SW, the search box model used by OPACs and Web search engines does not adequately exploit the organized whole of the collection through groupings of items assigned to CV subject terms maintained in ontologies. A search box favours "specific searches for approximate targets" (Smith, *et al.*, 2006, p. 797) where the user can enter precise keywords and is looking for adequate (i.e., good enough) as opposed to comprehensive answers. Semantically organized collections favour "approximate searches for specific targets" (Smith, *et al.*, 2006, p. 797) where the user knows broad generic terms but is looking for specific treatises on that subject.

This research assumes it is possible to better capitalize on investments in semantic organization of information and powerful computer hardware. As the following section shows, a new generation of OPACs has started this endeavour.

3.1.5.1 Next Generation Library Catalogues

The NCSU Endeca discovery interface is the pioneer of the next-generation OPAC movement (see review from Julien and Bouthillier (2008)). Most of these next-gen OPACs offer improved browsing through faceted searching (Anderson & Hofmann, 2006) which essentially helps the searcher to "broaden his/her search by setting limiters based on categories such as genre or topic" (Iglesias & Stringer Hye, 2008, p. 13). Facets are orthogonal sets of categories (Yee, Swearingen, Li & Hearst, 2003, p. 402) such as "subject", "format", "genre". Usage statistics[1] suggest facets are appreciated and increasingly used by searchers; however, they do not provide an initial overview of the collection and its predominant subjects.

Facets have not solved the issues related to CV maintenance and assignment. Facet values are assigned for each individual information item by LIS professionals. These interfaces "have simply replaced the problem of laboriously categorizing the data items with the problem of laboriously categorizing the items' metadata into specific facets" (Smith, *et al.*, 2006, p. 798). Just as users did not always recognize the nature of CV subject terms (see section 1.2.1 *Problems with Library Subject Browsing*), the distinction between content keywords and facets are not clear to users (Smith, *et al.*, 2006, p. 803).

3.2 Information Visualization

> *...orientation, visual search, and cognitive processing of complex subject matter may be enhanced if structures behind…information…as well as the relevance for the task at hand…are made explicit. (Keller & Tergan, 2005, p. 11)*

Keller (2005) states that information visualization (Card, Mackinlay & Shneiderman, 1999b) generally aims at facilitating interactive information retrieval. Specifically, a primary objective of IV is to visually reveal patterns in large data sets (Bederson & Shneiderman, 2003, p. ix; Ware, 2008, p. 172). This research direction is partly explained by "the importance of visual interfaces for human-computer interaction" (Newby, 2002, p. 49) to support information seeking (Veith, 1988).

From a cognitive perspective, the impetus towards visualization "is in line with the dominance of vision" (Raeithel & Velichkovshy, 1996, p. 203) which capitalizes on the ability of the human mind to rapidly perceive visual information (Keller & Keller, 1993)

---

[1] See http://www.slideshare.net/youthelectronix/prestamo and www.lib.ncsu.edu/endeca/presentations/200806-ala-pennell.ppt

in fast pre-attentive cognitive processing (Triesman, 1985) (see section 3.2.1 *Theoretical Foundations*).

The term "information visualization" as a technology for visualizing abstract data structures can be traced back to the Xerox Palo Alto Research Center (USA) at the beginning of the nineties (Keller & Tergan, 2005). Reflecting its multidisciplinary nature, the field is covered by publications from information science (e.g., JASIST), computer-science/engineering (e.g., ACM, IEEE InfoVis Symp), and human-computer interaction (e.g., International Journal of Human-Computer Studies, Interacting with Computers). Since 1995, the creation of the IEEE Symposium on Information Visualization has provided a more focused venue and suggests IV has become a distinct research avenue. IV also attracts researchers from other fields such as information management (Chung, Chen & Nunamaker Jr., 2005; Turetken & Sharda, 2005), digital government (Zhang & Marchionini, 2005) and electronic imaging (Cribbin & Chen, 2001).

Scientific visualization and information visualization are sometimes used synonymously but they are in fact two distinct research avenues. Scientific visualization is "always about physical objects" (Zhu & Chen, 2005, p. 144) using data "with an inherent spatial component (e.g., wind tunnel vector data or three-dimensional (3D) medical images)" (Tory & Moller, 2004, p. 151). Information visualizations "usually do not have inherent geometries by which to map information" (Zhu & Chen, 2005, p. 145) and "typically involve abstract, non-spatial data (e.g., financial data or document collections)" (Tory & Moller, 2004, p. 151) making the visual metaphor design an arbitrary choice (Tamara Munzner's statement in Rhyne *et al.* (2003)).

The role of the information metaphor is to adequately communicate the existing information and its organization structure. This does not mean the searcher will necessarily find the organization useful for his/her task. The choice of metaphor "is only as useful as the underlying structures and relationships" (Sebrechts, 2005, p. 139) it intends to communicate to the user.

Freitas *et al.* (2002) make the distinction between 1) techniques for displaying data characteristics and values vs. 2) techniques for displaying data structure and relationships. The latter is generally associated with the domain of information visualization and this may be where IV can offer significant benefits since

> *"over time, more and more of what we know has become abstract, related in language or symbolism rather than through the concrete aspects of reality. One way in which we can enhance our ability to "know" is by making it possible to visualize these otherwise abstract relationships." (Sebrechts, 2005, p. 136)*

The following section describes the domain of information visualization which is the largest section of the theoretical framework. Figure 3.2 below offers an overview of this section.

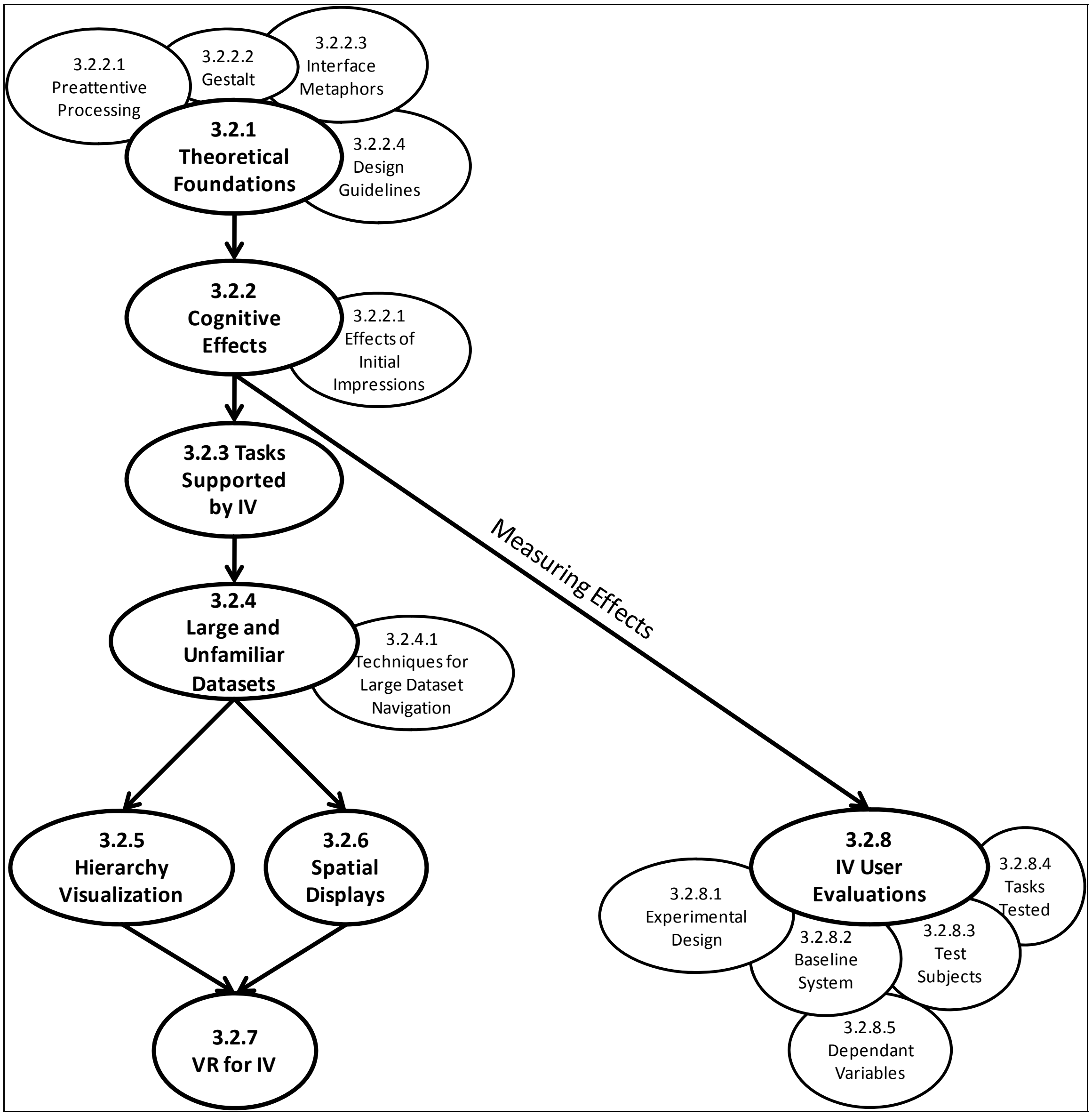


Figure 3.2. IV Summary

As shown in Figure 3.2, IV theoretical underpinnings are first provided (section 3.2.1) and they include preattentive processing, Gestalt, direct-manipulation interface metaphors, and IV design guidelines stemming from seminal works in the field. This is following by a summary of IV potential cognitive affects (section 3.2.2) and the associated IR tasks these visual tools aim to facilitate (section 3.2.3). IV benefits are mostly manifested when browsing large and unfamiliar datasets (section 3.2.4). This requires that the data be structured (see section 3.1.4.2 *Browsing Subject Structures*) before it can be visualized as a hierarchy (section 3.2.5), or as single-level hierarchies (also called *flat classifications*) often represented using spatial displays (section 3.2.6).

These IV techniques are followed by the potential advantages of the 3D virtual reality (VR) for IV although little research has been performed in this area. Finally, the effects of IV are measured during usability studies whose designs are described using an in-depth review (section 3.2.8).

3.2.1 Theoretical Foundations

IV relies on ongoing research in cognitive science and there is little theory to explain why any one design is preferred by users. Theoretical foundations of IV include preattentive processing which predicts some visual perception tasks are highly efficient. Gestalt laws of perception describe what the user might understand when specific visual designs are chosen. The field also offers guidelines and suggestions for IV design and testing.

3.2.1.1 Preattentive Processing

Preattentive processing refers to the human cognitive ability to automatically recognize basic features of objects such as colors, adjacent line ends, contrasts, tilt, curvature and size (Triesman, 1985). Typically, to be considered preattentive, visual tasks must be performed under 200 milliseconds (Healey, Booth & Enns, 1996), effortlessly and without specific attention. For example, detecting one red object within a group of blue objects is performed preattentively and can help rapidly draw the attention of the user to an object with a unique visual feature.

3.2.1.2 Gestalt

What users of information visualization perceive by specific design features can be partly described by the Gestalt laws of visual patterns perceptions (Wertheimer, 1925). The Gestalt approach is a descriptive set of theories and does not explain why humans perceptually process and organize visual stimuli. Gestalt emphasizes that we perceive objects as wholes rather than separate parts; indeed,

> *"when we open our eyes we do not see fractional particles in disorder. Instead, we notice larger areas with defined shapes and patterns. The 'whole' that we see is something that is more structured and cohesive than a group of separate particles" (Pedroza, 2005)*

There are several Gestalt laws (see review from Mullet *et al.* (1995)) which might very well act in combination. For example, relationships between items in information visualization displays are often drawn using "proximity, closure, continuity" (Koshman, 2006, p. 194), "connected lines, or color coding" (Shneiderman, 2003, p. 368).

Synnestvedt and Chen (2005) provide the following list of Gestalt laws relevant to computer interface design:

- **Spatial Proximity**: Things that are close together are perceptually grouped together.
- **Spatial Concentration**: Regions of similar element density are grouped together perceptually.
- **Similarity**: Elements of similar appearances tend to be grouped together.
- **Continuity**: Smooth and continuous connections between elements are easier to perceive than abrupt changes in direction. Assumes connectedness, which can be a more powerful grouping principle than proximity, color, size, or shape.
- **Symmetry**: Symmetrically arranged elements are perceived as forming a visual whole much more strongly than elements with lesser symmetry.
- **Closure**: A closed contour tends to be seen as an object.
- **Relative Size**: Smaller components of a pattern tend to be perceived as objects.
- **Figure and ground**: Object like figures are perceived as being in the foreground, the ground is whatever lies behind the figure.

For example, the VIBE prototype (Olsen, Korfhage, Sochats, Spring & Williams, 1993) shown in Figure 3.3 capitalizes on the law of proximity which states that elements tend to be grouped together depending on their closeness.

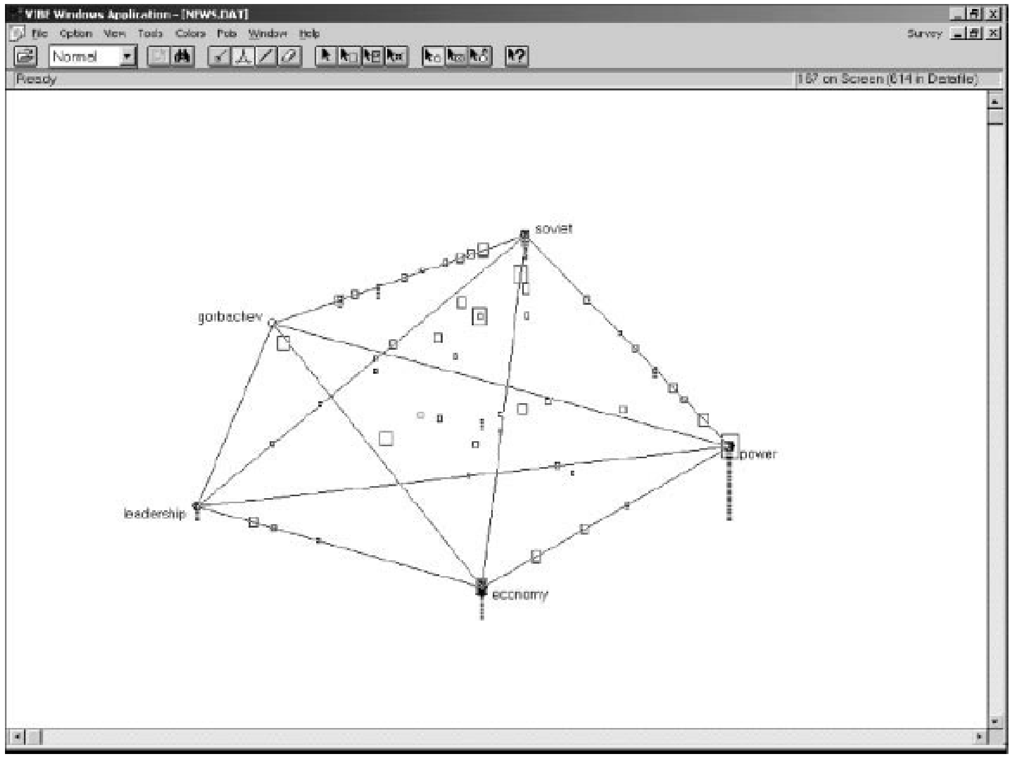


Figure 3.3: VIBE Prototype (Olsen, et al., 1993)

VIBE depicts user keywords as nodes (five nodes in Figure 3.3) and vertices (i.e. lines) between them on which each matching document is located. The visual distance between documents and keywords suggests the relative relevance of each concept to each document. This form of visual relationship communication is possible when the number of keywords is small so the number of relationship lines does not become too dense producing a kind of *spaghetti* effect (see example from Julien & Cole, 2009).

Beyond preattentive processing and Gestalt laws, IV is necessarily part of the shift from strictly textual command-based interfaces towards direct manipulation interface metaphors. Current keyword/result list search tools are descendents of the former era.

3.2.1.3 Direct Manipulation Interface Metaphors

> *"It is believed that the use of a metaphor in the construction of a visual space would (...) shorten learning process, (...) increase users' interest for the system, and make a full use of perception capacity of human being in navigation."(Nguyen & Zhang, 2006, p. 981)*

The commercial success of direct manipulation visual interface metaphors (e.g., Windows Desktop) over traditional command-based, text-only interfaces indicates “the power of using computers in a more visual or graphic manner” (Shneiderman, 2003, p. 364). Direct manipulation interfaces tend to create positive first reaction and "often evoke enthusiasm from users, and for this reason alone it is worth exploring their use” (Hearst, 1999, p. 282). This is especially the case for novice or occasional users.

Visual interface metaphor interaction techniques are especially apt “to provide orientation or context, to enable selection of regions, and to provide dynamic feedback” (Shneiderman, 2003, p. 365). Shneiderman and Plaisant (2005) list three principles of direct manipulation interfaces (p. 234):

1. **Continuous representations** of the objects and actions of interest with meaningful visual metaphors
2. **Physical actions** or presses of labeled buttons, instead of complex syntax.
3. **Rapid, incremental, reversible actions** whose effects on the objects of interest are visible immediately.

Early visual interface metaphor designs were supported by repeated observations that users tried to understand computers as analogical extensions of familiar contexts (Douglas & Moran, 1983; Mack, Lewis & Carroll, 1983). Visual metaphors have been shown to facilitate learning (Carroll & Thomas, 1982) by assisting in linking new information to existing knowledge (Indurkhya, 1992; Petrie & Oshlag, 1993) and structure the perception and the handling of the environment they refer to (Jih & Reeves, 1992; Kim & Hirtle, 1995; McKnight, Dillon & Richardson, 1990). Metaphors benefit the searcher by offering recognition over recall since "we can recognize that we have seen something before far more easily than we can reconstruct a memory" (Ware, 2008, p. 160).

Visual metaphors may reduce the learning curve by providing a set of basic interaction scripts (Schank & Abelson, 1976) to first time users. These *afforded* (Gibson, 1986) interaction mechanisms facilitate initial human-computer interaction and encourage learning by exploration (Ahlberg, Williamson & Shneiderman, 1992). This is analogous to the idea of context (Tversky, Zacks, Lee & Heiser, 2000), that is, a "set of structural properties that provide a framework for meaning" (Ziemkiewicz & Kosara, 2008, p. 1274). Metaphors are associated with the idea that meaningful mental models of how and why a tool works. Mental models are internal representations of tools that make it easier to remember tasks and spontaneously find more efficient ways to use these tools (Kieras & Bovair, 1984). IV researchers assume that a data metaphor aids data interaction in a similar way to direct manipulation metaphors aiding software interaction (Ziemkiewicz & Kosara, 2008, p. 1274).

Metaphors do have drawbacks. Halasz and Moran (1982) cautioned that teaching new users using metaphors may be an easy way to introduce a user to a new system but that they can eventually hinder the development of "an effective understanding of systems." (p. 383). Although it is often easier to describe the function of a design using a metaphor it "does not necessarily mean that whilst interacting with the product the user understands the design through one single, consistent metaphor" (Overbeeke, Djajadiningrat, Hummels, Wensveen & Frens, 2003, p. 12). Gentner and Nielsen (1996) summarize three "classic drawbacks" of metaphors:

- **Magic attributes**: the target domain has features not in the source domain.
- **Misleading attributes**: the source domain has features not in the target domain.
- **Violation of expectations**: some features exist in both domains but act differently.

Ultimately, a human-tool interaction metaphor is never *the* system itself and all designs choose to offer certain interactions rather than others, some information rather than other. As the Three Mile Island incident showed (United States, 1979), no matter how critical, designed and tested a system may be, there will inevitably come a point where these design choices will prove inadequate. Fortunately, few IR applications are as critical as a nuclear power plant.

Visible textual and symbolic information offered by visual metaphors serves as a reminder for what is and is not possible (Norman, 1988). This provides cognitive aids to memory in the form of "small images, symbols, and patterns (that) provide proxies for concepts" (Ware, 2008, p. 169). In this manner it is possible to place upwards of thirty (p. 169) concept proxies on a screen providing a very quickly accessible concept buffer. Visual icon proxies work once users have learned the association between the proxy and the object or action it hopes to represent (p. 169). This does not preclude traditional text-based interfaces that "are easier to use than other methods for many users in many contexts" (Hearst, 1999, p. 282). The choices between text, visual icons of various types and their combination bring up the question of how closely should the interface metaphor resemble its realworld counter-part.

3.2.1.3.1. Level of Realism

> *"It is only necessary for things to exhibit roughly the right physical behavior for objects to appear normal when we interact with them. (...) the models that are imbedded in our nervous systems are only crude approximate representations..." (Ware, 2008, p. 102)*

A fundamental interface metaphor design question concerns the level of realism or 'likeness' with the chosen analogy. The metaphor should be recognizable but it does not need to be exactly like 'the thing' to suggest it shares some of its behaviors (Laurel, 1993). In fact, metaphor designers are advised to restrict fidelity with reality (Stappers, Gaver & Overbeeke, 2003) by choosing "tools for 'expressive' rather than photorealistic rendering" (Sebrechts, 2005, p. 151).

Developing an interface metaphor with strict adherence to reality was found detrimental to the achievement of users' learning objectives (Loftin & Kenney, 1995). A series of studies have demonstrated "that virtual environments of moderate environmental fidelity can lead to effective learning and transfer of spatial layout" (Sebrechts, 2005, p. 151).

It is difficult to explain why too much realism in metaphor design may be detrimental. Sebrechts (2005) suggests that "excessive realism tends to impose more constraint than is desired at early visualization stages" (p. 151). Sherman and Craig (2003) suggest this may be because "attempting to render a world in a photo-realistic way can make mental immersion difficult, because any flaw in the realism will spoil the effect" (p. 383). Excluding specialized applications such as virtual architecture, most interface metaphors are not meant to behave exactly like their real world counterparts and this should be readily communicated to the user; otherwise, there is a risk of erroneously suggesting the interface 'is' the analogy and behaves exactly like its real-world counterpart, nothing less and nothing more.

The problem of realism in software environments has similarities with the problem of software visual icon design. Rogers (1989) lists four ways visual icons suggest its 'referent' (p. 110) or the effect they will produce if selected (see Table 3.2)

Icon Representational Form	Definition	Example	Level
Resemblance	Depicts the referent through analogy	 "Falling Rocks"	Highest
Exemplar	Depicts a typical example of a type of object or referent	 "Restaurant"	High
Symbolic	Depicts an image which must be abstracted to reveal the referent	 "Fragile"	Low
Arbitrary	No relationship, referent must be learned	 "Biohazard"	Null

Table 3.2: Visual Icon Referent Representation Ways (Rogers, 1989)

Rogers (1989) compared recall for four types of icons vs. command names alone. The experimental design was a between-subjects repeated measures of post-test memory task performance. The findings show that the most effective form of icon representation is "that which is the most direct (...) the set depicting concrete objects and abstract symbols" (p. 115). This means that resemblance icons would be the most effective while arbitrary representations would be the least effective.

Unfortunately, there are abstract concepts which have no obvious visual representation. For example, the concept of a "subject heading" might be referred to as a *container* image and textual label. Should the container be represented by the classic Windows 'File Folder' icon or is this a different kind of container? These types of abstract objects or functions are often found in HCI when it is "necessary to use less direct forms of representation" (p. 111).

Arbitrary forms are easier to learn if there are few of them. The advantage is that they produce little or no prior unwanted associations between the icon and the function or object it represents (p. 111). Abstract visualizations have a number of additional advantages: they are not limited by "real-life" constraints, they can be systematically designed, and well defined building blocks can be used to create larger visual representations.

Visual interfaces are not static images. They should allow dynamic iterative cycles of user request and clear system response (Norman, 1988). The clarity of the system response and its resulting status is accentuated by visually animating system state transitions.

3.2.1.3.2. Animated Transitions

Animated transitions are an important element of interactive visual metaphors because they can facilitate user comprehension of actions taken by the system and triggered by the user. The technique explicitly expresses changes in the state of the system which is said to be critical for a successful HCI (Norman, 1988; Sutcliffe, 2003). Bederson and Shneiderman (2003) report that "animation improved subject's ability to learn the spatial position…without a speed penalty" (p. 92). Nguyen *et al.* (2004) designed their Web result visualization system so that "each visual interaction is accommodated by an animation in order to preserve the cognitive-map of the user during the navigation" (p. 697).

Scientific understanding of human cognitive processes is incomplete and does not provide an explanation of why animation is beneficial. Robertson *et al.* (1991) suggest that interactive animation reduces cognitive load by utilizing perceptual system capacity. Bederson and Shneiderman (2003) state that the technique may help “maintain object constancy, and that without animation, users must spend time rebuilding an understanding of which object is which” (p. 93). Animation may be critical in 3D visual environments since “a strong depth cue is motion parallax” (van Ham & van Wijk, 2003, p. 35).

Preattentive processing, Gestalt and direct manipulation interface metaphors serve as the theoretical underpinnings for information visualization design. Beyond theory, designers require guidelines describing what techniques work in which context and how they can assess the quality of their design.

3.2.1.4 Design Guidelines

IV design guidelines are drawn from seminal works in visual perception. This includes the highly influential Bertin (1967) which identified basic elements of diagrams. Bertin (1967) lists the following main processes in visualization comprehension task analysis (from Trafton *et al.* (2002)):

1. **Encode visual elements of the display**: E.g., identifying lines and axes, influenced by pre-attentive processing and how easily each shape can be identified.
2. **Translate the elements into patterns**: E.g., relative sizes of pie chart slices or points strung together form a slope, affected by “distortions of perception and limitations of working memory” (Synnestvedt & Chen, 2005, p. 2).
3. **Identify the relation** between the patterns and the labels to interpret the relationships communicated by the graph. E.g., interpret the relative importance of one pie chart slice and the type of object it represents.

Edward R. Tufte (1983) provides two fundamental rules for visual display:

1. Within reason, maximize the data-ink ratio, i.e. every drop of ink, or pixel on your screen, ought to be information bearing. Anything that appears simply for decoration should be removed.

2. Within reason, maximize information density, i.e. prefer displays with more rather than less information.

Cleveland and McGill (1984) proposed a theory of graphical perception comprised of nine elementary information extraction tasks ordered from those most to least accurately performed by people. Their work "has served as the basis for guidelines on which low-level visual mappings are appropriate for which sort of data" (Ziemkiewicz & Kosara, 2008, p. 1274); however, their approach is strictly reductionist. They offer no way to assess the overall quality of a visual representation and there is no consideration of dynamic and interactive visualizations.

Arnheim (1972) makes a case for how to engage in *visual thinking* allowing perceptual stimuli to trigger thought processes such as selection and abstraction. Spence (2001) describes many techniques used to visualize multidimensional datasets without "inherent two or three-dimensional semantics" (Keim, 2002, p. 2). McKim (1980) states that graphical communication is "an explanatory process concerned with presenting fully formed ideas to others" (p. 122-23) which is far more appropriate to the limitations of language when symbolic or interpretive objects are described.

Preattentive processing, the descriptive Gestalt approach and design guidelines hope to support multiple uncharted cognitive mechanisms involved in IV usage. The following section offers some of these expected positive effects associated with IV tools.

3.2.2 Cognitive Affect

> *A key aim of information visualization research is to discover and develop ways of amplifying human cognition. (Pirolli, Card & Van Der Wege, 2000, p. 161)*

To amplify human cognition is a worthy objective; however, cognitive science research is ongoing (Ellis & Dix, 2006) and not yet able to predict how to 'amplify' the evolving set of interacting cognitive abilities comprised in the evolving definition of 'human cognition'. The exact cognitive mechanisms affected by IV are not known and, as the following section shows, the literature offers a wide array of conjectures and assumptions.

There are reports of positive performance effects for IV systems suggesting there are cognitive effects. For example, Veerasamy *et al.* (1997) conducted a controlled experiment where results suggest “that the visualization tool helps users in identifying more relevant documents (...) more quickly” (Veerasamy & Heikes, 1997, p. 244). According to Stephens *et al.* (2004), "numerous studies (show) the average performance of participants improved when using a visualization tool to discover knowledge hidden in data” (Stephens & Handzic, 2004, p. 1), and Monk *et al.* (1988) have shown that even a static, non-interactive graphical representation is useful.

Specifically, during a search task, “instead of wading through a long list of ‘hits’…visualization can serve to cluster similar documents and identify regions of potential interest” (Newby, 2002, p. 32). This “could enable set-at-a-time perusal of documents, rather than document-at-a-time perusal of text displays” (Veerasamy & Heikes, 1997, p. 237) while providing the user with new insights into their understanding of the information space (Lohse & Walker, 1993).

IV may allow “people to move from cognitive problem–solving to more natural sensorimotor strategies” (Chalmers, 1993, p. 377) which capitalize both on the visual and the spatial working memory system (Baddeley, 1998; Logie, 1995). This may reduce cognitive load (Sweller & Chandler, 1994) through *computational offloading* (Navarro-Prieto, Scaife & Rogers, 1999). For example, visual representations may provide a measurable gain in search task efficiency by shifting “the user’s mental load from slow reading to faster perceptual processes such as visual pattern recognition” (Zaphiris, Gill, Ma, Wilson & Petrie, 2004, p. 53). Some state that generally “people usually have an easier time in understanding information when it is visually presented” (Henderson & Card, 1986; Shneiderman, 1996); however the definition of *easier* is not absolute and may mean performance or preference which are not necessarily correlated.

Effects of IV are often found in post-test measures of user preference as opposed to accuracy and efficiency scores. Experimental results from Morse *et al.* (2000) show that purely textual displays “were extremely ill-preferred, regardless of performance” while “visual interfaces are associated with…user preference” (p. 659). Informal “user feedback (…) clearly indicates that users like the graphical nature of IV” (Chen, *et al.*, 1998, p. 600).

Most interesting are two controlled comparative studies (Becks, et al., 2002; Modjeska & Waterworth, 2000) that report user preference for the IV interface without significant effects on performance. Preference not correlated with performance measures was also reported in a single-tool study (Large, Beheshti, Clement, Tabatabaei & Tam, 2009; Sutcliffe, Ennis & Hu, 2000a) where users rated the system highly despite poor performance. A tool design needs to offer adequate performance beyond which other criteria seem to become significant to user preference.

3.2.2.1 Affects of Initial Impressions of a Design

IV may benefit from novelty effect but there are suggestions that agency provided by direct-manipulation and aesthetic appeal create positive initial impressions (Nielsen, 2003; Norman, 2004). There are strong suggestions that these positive first impressions have a positive effect on the learning curve, user preference and even measurable performance (Ashby, Isen & Turken, 1999; Tractinsky, 1997). The growing importance of the *affective* domain (i.e., emotions) is also recognized as a factor in information retrieval (Nahl, 2007).

Davis (1989) described their Technology Acceptance Model (TAM) and showed that in the context of a work setting, initial perception of usefulness and ease of use generated by a software interface are strong predictors of the acceptance of the tool over time (Davis, 1989; Davis, et al., 1989). Perceived usefulness had a significantly greater correlation with usage behavior than perceived ease of use. Further analysis suggested that ease of use may be a causal antecedent to perceived usefulness. This would be somewhat intuitive since users must be able to use a system before they can assess how useful it may be. The research suggest that, in an office workplace, users are more likely to keep using a software tool if they can quickly learn it and perceive it as useful over time.

In order to be useful, a software tool is designed to support anticipated tasks performed by an adequately known set of users. Like all HCI techniques, IV is meant to facilitate the completion of certain types of tasks.

3.2.3 Tasks Supported by IV

> *Information visualization is sometimes described as a way to answer questions you didn't know you had. (Plaisant, 2004, p. 111)*

The above quote from Plaisant (2004) refers to the serendipitous discovery of valuable information associated with browsing tasks (see section 3.1.1). Beyond unplanned discoveries, IV users are said to be interested in finding relationships between documents such as "discovering similar items, identifying patterns such as clusters, outliers, and gaps" (Bederson & Shneiderman, 2003, p. ix), "correlations" (Zhu & Chen, 2005, p. 145), and to recognize relevant documents (Lin, 1997; Lohse & Walker, 1993).

The interactive nature of direct manipulation IV may also support learning by exploring (Ahlberg, *et al.*, 1992; Chalmers, 1993) which enables "users to explore patterns, test hypothesis, discover exceptions, and explain what they find to others" (Bederson & Shneiderman, 2003, p. ix). Newby (2002) suggests this would support an iterative "search process, narrowing in on areas of potential value" (p. 32) and specifies that the IV tool should necessarily "retrieve documents [as opposed to being] simply a visual presentation of a set of data" (p. 37).

Browsing or explorative searching are the strategies cited most frequently as benefiting from visual support (Marchionini, 1995). Browsers are concerned with discovering patterns in the document space and getting an overview of available documents and their semantic relationships (Becks, et al., 2002). As described in section 3.1.1 *Browsing*, these types of tasks are broad and ill-defined which requires the ability to filter unpromising information and refine ones understanding of the collection through iterative cycles of inspection and query refinement. This is especially critical in today's quasi-infinite digital information collections.

3.2.4 Large and Unfamiliar Datasets

> *Visualization's primary goal is to make it easier for people to understand and use vast amounts of data. (Wiss, Carr & Jonsson, 1998, p. 137)*

The advantages of IV for IR tasks are often associated with large and unfamiliar data sets that can be more easily searched by visually representing the patterns in data. The benefits offered by information visualization displays are more prominent when interacting with large volumes of data (Card, Mackinlay & Shneiderman, 1999a; Chen, 1999; Spence, 2001; Tufte, 1983). This implies the information has been adequately structured (see section 3.2.5 *Hierarchy Visualization*). Subsequent visualizations of these structures may support understanding of the relations between information elements and visually searching relevant information (Keller & Tergan, 2005; Sebrechts, 2005).

Visualization techniques for information retrieval are said to be aimed at a non-technical general public (Borner, Chen & Boyack, 2003, p. 23) who are unfamiliar with a topic domain (McDonald & Stevenson, 1998). Everyone is a novice except in his/her own domain and guidance is often required "through newly accessible oceans of on-line information" (Morse, et al., 2000, p. 637).

IV is promising for large and unfamiliar datasets but one "long-lasting challenge" (Chen, 2005, p. 14) concerns real world practicality or scalability. Most IV techniques are demonstrated with small experimental systems of "hundreds to tens of thousands of data points (and may not apply to) text retrieval systems (that) often deal with hundreds of thousands or millions of items" (Newby, 2002, p. 38). As the following section shows, some techniques have been developed to facilitated navigation in densely populated visual interfaces for large data sets.

3.2.4.1 IV Techniques for Large Dataset Navigation

The magnifying lens approach provides a tool used to select a screen area to separately view in more detail or higher magnification. The technique may facilitate visual access to large data sets by improving performance on target finding tasks (Guiard & Beaudouin-Lafon, 2004; Leung & Apperley, 1994). Magnifying lens assumes there are predetermined levels of detail useful to all tasks and all users: overview and magnified. This applies well to certain contexts such as 2D schematics but some applications require multiple levels of magnification.

Zoomable User Interfaces or ZUIs (Bederson, Hollan, Perlin, Meyer, Bacon & Furnas, 1996) offer multiple magnification levels that users can dynamically scroll through. This type of *focus-plus-context* method is used to represent and manipulate large sets of data by managing the level of detail and separating the user point of interest area (focus) from the global view (context) (Guiard & Beaudouin-Lafon, 2004; Pietriga, Appert & Beaudouin-Lafon, 2007). Fabrikant (2001) demonstrated that users zooming into a spatial area understand that they are going deeper into a semantic hierarchy. ZUIs require an initial overview which does not overwhelm the user; consequently, very large dataset navigation requires additional HCI features that support detection of promising targets.

ZUIs can be combined with panning to facilitate the detection of promising areas of a large and complex dataset overview (Bourgeois, Guiard & Lafon, 2001). This technique is known as *Pan & Zoom* and provides a visual area smaller that the dataset overview through which only a partial view of the dataset is presented. The user browses the dataset through translations and zoom level modifications.

The FishEye technique presents the whole dataset at a low level of detail and utilizes a movable non homogeneous distortion (using the magnification lens metaphor) to a subset of the dataset shown at a higher level of detail (Gutwin, 2002). The major disadvantage of this distortion technique is that it is very hard to relate two sections of the same dataset "even when providing visual cues such as animation or coloring" (Abello, van Ham & Krishnan, 2006, p. 2).

Pan and/or zoom techniques are inadequate when the dataset overview may overwhelm the user, and the distortion caused by the FishEye hinders the construction of an overview with relates different sections of the dataset. An alternative to these techniques is to summarize groups of items using subject indexing (see section 3.1.4) or clustering techniques (see section 3.2.5 *Hierarchy Visualization*) to represent groups of documents using a few highly salient terms. This creates a "browseable hierarchy" (Smith, *et al.*, 2006, p. 798) of general to specific terms as described in the following section.

3.2.5 Hierarchy Visualization

All ontologies including those used by the Semantic Web contain semantic hierarchies. These broad to narrow structures are ubiquitous and seem to arise from the human need to organize collections of objects (Olson, 2004; Simon, 1996). Hierarchies are used for IR when “one is simply interested in gaining an overview, or has a general question, one peruses the table of contents, which lays out the logical structure of the text” (Cutting, Karger, Pedersen & Tukey, 1992, p. 319). For example, hierarchies are found in the broader/narrower subject terms of library information organization systems (e.g., LCSH, MeSH), and using a mathematical analysis, Resnikoff (1989) demonstrated that the hierarchical structure of the common library catalog card system minimizes manual search time (p. 112-117).

Explicitly representing ontology concepts and their relations in a visual manner may facilitate retention of the ontology as a whole. Simon (1996) cites the "well-known" (p. 71) experiments (de Groot, 1966) and others on chess perception. These suggested that grand masters could recall the exact positions of chess pieces based on the relations between the pieces; as opposed to some special gift of visual memory. In the same manner, users of visually represented subject hierarchies may learn and recall subjects based on their relations; as opposed to using a classic list of alphabetically sorted terms.

There are some drawbacks to organizing information in broad to narrow subject hierarchies. As suggested by Smith *et al.* (2006), it can be "laborious" (p. 798) to organize a large information collection and there is never a single agreed upon best categorization. As the collection grows it is difficult to maintain a balance between the branches of the hierarchy. A well balanced hierarchy is said to be "essential" (p. 798) for effective narrowing of the collection when the searcher travels down towards narrower subjects.

Hierarchies are in fact generally "large, arbitrarily shaped, and often used by people to make decisions" (Bederson & Shneiderman, 2003, p. 229). Navigating a large hierarchy is cognitively demanding since it "requires retaining in memory all potentially interesting paths…that the user did not pursue but might consider returning to if the information sought had not been found" (Roussinov & Chen, 2001, p. 797). This makes these "large structures (…) much harder to grasp" (van Ham & van Wijk, 2003, p. 31) and creates an "increased cognitive load for users who are forced to make selections among the hierarchical branches, especially when the whole hierarchy is not displayed on the screen" (Lin, 1997, p. 43). Current systems tend to avoid this problem by limiting the number of visible items to about 10,000 (Akrivi, *et al.*, 2007).

Hierarchies are ubiquitous and often large which makes these structures good candidates for information visualization; indeed, hierarchy visualization (also called *tree* visualization) has become one of the "most mature" (Chen, 2004, p. 90) and "important topic in the visualization community" (van Ham & van Wijk, 2003, p. 31). Hierarchical organizations are universally recognized and there are multiple hierarchy navigation interfaces.

A partial explanation behind "the popularity of strictly hierarchical interfaces is that tree layout is a much more tractable computational problem than general graph layout" (Risden, Czerwinski, Munzner & Cook, 2000, p. 697). Hierarchies are computationally easier to deal with because they are a constrained form of a general directed graph (Battista, Eades, Tamassia & Tollis, 1994) (i.e., nodes and arrows between them) where each node has strictly one parent (i.e., one arrow pointing to the node) but may have multiple children (i.e., multiple arrows stemming from one node to others).

Very few studies have applied hierarchy visualization to textual information retrieval (notable exceptions: Hearst & Karadi (1997), Cribbin & Chen (2001)). This type of research is ambitious since applying hierarchy visualization techniques to a collection of textual documents implies two unresolved difficulties (Lin, 1997, p. 43):

1. Generating the hierarchy: which means organizing the collection using either

- Clustering algorithms from computer science (Iwayama & Tokunaga, 1995) based on some measure of similarity (e.g., common vocabulary, citations)
- Manual labor from LIS professionals assigning controlled subject terms (e.g., LCSH, MeSH)
- Automatic classification requiring both LIS expertise to build and maintain a model hierarchy and computer science machine learning algorithms (Krowne & Halbert, 2005)

2. Displaying the large hierarchy: choosing hierarchy visualization and interaction techniques, dealing with visual clutter

Clustering and automatic classification address the fact that manual organization of information into pre-determined subject hierarchies (e.g., LCSH, MeSH) is not scalable to quasi-infinite online collections. Clustering demands high computational capacity and computer science research has produced multiple algorithms offering various combinations of cluster quality and computational efficiency (see review by Iwayama & Tokunaga (1995)).

Computer generated subject hierarchies have their set of HCI problems. Clustering methods often suffer from labelling issues: 1) the label has little meaning or is simply not suggestive of contents (Abello, *et al.*, 2006) or 2) the groups are not “at the same level of abstraction” (Chen, *et al.*, 1998, p. 597). This is due to the “unsupervised nature of clustering” (Hearst, 1999, p. 223) which generates groups and extracts their labels without considering their level of description. For example, a group might be labelled “Study of”, or a hierarchy level can show groups labelled “American History” and “Bill Clinton” which should not be placed on the same level of abstraction.

Beyond computer generated subject hierarchies "there are in fact many existing, important collections whose contents already have hierarchical metadata assigned" (Yee, et al., 2003, p. 402). Metadata refers to information that *stand for* or represent other information such as the title of a book, its assigned subject headings (see section 3.1.4 Subject Indexing), or its semantic contents (see section 3.1.3 *Semantic Web*). These highly salient information pieces are critical to the development of visual interfaces for digital libraries (Beagle, 1998; Mitchell, 1999). This research aims to visually represent an existing hierarchically organized collection in the hope of facilitating access to these browseable structures.

Assuming a hierarchy is available, graph visualization and interaction techniques should be investigated. Early works emphasized aesthetic and easy-to-read layouts (Battista, et al., 1994). Today, commercial applications offer “two large categories of solutions…space-filling techniques and node-link techniques” (Plaisant, Grosjean & Bederson, 2003, p. 287). Space-filling refers to techniques inspired by Treemap (Shneiderman, 1992) (see section 4.2.2.4.2) and exemplified by Newsmap[1] which shows the top news stories by region and subject. Node-link techniques are the ubiquitous indented outline view offered by disk management tools in computer operating systems such as MS Windows, Linux Ubuntu or Mac.

Ontology visualization research suggests that browsing large hierarchies should be integrated with keyword search tools (Akrivi, *et al.*, 2007). For example, the Flamenco system (Yee, et al., 2003) showed that a significantly more efficient and enjoyable user experience (as compared to keyword search alone or pure categorization browsing) can be achieved by integrating browsing and filtering features.

A visual hierarchy traditionally represents parent-child relationships treating members of a single level as equals; however this is not always the case. Spatial displays are a family of IV techniques which address this issue by visually representing more complex relationships between items on a single level of the hierarchy.

3.2.6 Spatial Displays

> *Spatial visualization is used to communicate information search and browsing activities in a natural way by applying metaphors of a spatial navigation and attraction to abstract information spaces. (Hemmje, et al., 1994, p. 250)*

[1] see http://newsmap.jp/

Spatial displays use spatial metaphors (also referred to as *spatialization* or *information landscapes* or *map visualizations*) to represent semantic relationships between information items. The technique generally produces a topographical map data metaphor where similar documents are presented as single groups of varying sizes placed on a surface such that “closely semantically related items are spatially proximate and semantically unrelated items are spatially distant” (Westerman & Cribbin, 2000, p. 766). The assumption is that information retrieval will be facilitated if semantic relationships are adequately conveyed to the user (Wise, 1998).

The value of using a geographic metaphor was first suggested by the SPIRE landscape visualization (Wise, Thomas, Pennock, Lantrip, Pottier, Schur & Crow, 1995). Spatial displays are an example of “creating visual cues in a retrieval interface to allow visual and perceptual inferences during the search process” (Lin, 1997, p. 52). The technique is scalable since it represents multiple items as a single visual group and it has become one of the most popular in information visualization (Chen, 2004, p. 107). Examples of spatial metaphors for IR include:

- Bead (Chalmers & Chitson, 1992)
- Starfield (Ahlberg & Shneiderman, 1994b)
- VxInsight (Davidson, Hendrickson, Johnson, Meyers & Wylie, 1998)
- LyberWorld (Hemmje, et al., 1994)
- StarWalker (Chen, 1999)
- SPIRE (Hetzler, Harris, Havre & Whitney, 1998)
- Financial Viewpoints (Strausfeld, 1995)
- Vineta (Krohn, 1995)
- TripleSpace (Mariani & Lougher, 1992)
- those stemming from the geographical domain (Fabrikant & Skupin, 2003; Skupin, 2000)

Comparative experiments between spatial displays and text-only displays (Becks, et al., 2002) or Web search engines (Chung, et al., 2005) have been performed. These show little or no significant results on measures of effectiveness and efficiency but once again (see section 3.2.2 *Cognitive Affect*), test users prefer the visual display over text only equivalents. Westerman and Cribbin (2000) compared 2D vs. 3D spatializations and their results “support the view that participants were relying on spatially mapped semantic information as a means of locating items within the database” (p. 781).

Spatial displays may capitalize on human spatial processing capacities. Jackendoff (1983) argued that all semantic information can be spatially mapped in this way, and Lakoff *et al.* (1980), argued "persuasively" (Ware, 2008, p. 62) that spatial metaphors are fundamental to the way language works. A spatial display of information can represent a collection as a terrain where "it is possible to literally see the structure that otherwise might take very substantial cognitive resources to extract” (Sebrechts, 2005, p. 139).

Spatial displays assume that the visual structure offered by the interface matches the structure the user would have otherwise extracted; something impossible to verify for any single user. Nonetheless, the technique could stave off *information overload* (Blair & Maron, 1985) through

> *“continuity of movement over a landscape, coupled with perspective viewing allowing one to incrementally refine one’s attentional focus down to more local areas, while smoothly adding more information to the context or periphery of view”(Chalmers, 1993, p. 384).*

Spatial visualizations support browsing (Borner, et al., 2003, p. 24) or “tasks which rely on consideration of the overall relationships of the themes and words active in a corpus, as well as the individual elements” (Chalmers, 1993, p. 379). As seen in section 3.1.1 *Browsing*, these types of tasks require broad exploration of the corpus in order to formulate and refine a query. This process often requires considerable navigation and scanning which can be supported by interface features that represent the structure of the information, possibly as a visual map (Card, *et al.*, 1999a; Sutcliffe, 1996).

Tufte (1990) also observed that maps are a superior means of applying his visualization design guidelines (see section 3.2.1.4 *Design Guidelines*) as they tend to present more information per display area than other techniques. Finally, spatial displays

> *"could be used both as an overview tool and an access or exploration tool, and interactive tools implemented on the map displays might be particularly useful in assisting the user to see and interact with the rich information revealed by the map displays"(Lin, 1997, p. 52).*

Although mostly a 2D technique, spatial displays suggest the information space can be visualized in multiple dimensions. As the following section shows, virtual reality (VR) applied to information visualization is still in its infancy but it promises multiple advantages over text-only and 2D displays while having its own set of design limitations.

3.2.7 Virtual Reality for IV

> *With its interactive focus, VR emphasizes procedural over declarative knowledge, activity more than description. This interactive model of learning is applicable to the exploration of information and knowledge spaces (Sebrechts, 2005, p. 157)*

Sebrechts (2005) makes a case for VR information visualization in part because, even in non-immersive desktop VR, "it helps place the focus on task rather than tool" (p. 156), supporting an invisible computer interface (Norman, 1998). Sebrechts (2005) does not suggest all IV should be in VR but states it may be suited to "extend visualization into experiential learning" (p. 138). Just as direct manipulation visual metaphors should not strictly adhere to their physical reality counterparts (see section 3.2.1.3.1 *Level of Realism*), VR "can provide non-existent, transformative worlds in which many of the properties of the physical world are modified or deliberately violated" (Sebrechts, 2005, p. 149). In fact, there is a "vast literature on advantages and disadvantages of 3D vs. 2D with somewhat conflicting results" (Teyseyre & Campo, 2009, p. 90)

One of the major appeals of VR is to capitalize "on the 'natural' navigation skills of users" (Plaisant, *et al.*, 2003, p. 362) acquired through "series of interactions that are highly over learned" (Sebrechts, 2005, p. 140). These include ubiquitous actions such as approaching, grasping objects, repositioning them in space, or rotating them, as well as repositioning ourselves. Transfer of these real-world skills to a VR interface is likely for many users since “most children, and even many adults, already embrace 3-D and know these interfaces quite well primarily through games” (Chang & Said, 2003, p. 4).

Experimental results suggest 3D is preferred by users (Levy, Zacks, Tversky & Sciano, 1996; Perez & de Antonio, 2004; Smallman, St John, Oonk & Cowen, 2001) and Card and Robertson (1996) have shown that 3D interfaces can be more powerful than 2D equivalents to manage overlapping windows. These findings, though are not supported by the demonstrated performance advantages (Keller & Tergan, 2005, p. 11). In fact, utilizing a third dimension in a desktop environment may be detrimental by imposing additional human navigational costs (Leuski & Allen, 2000; Sebrechts, Vasilakis, Miller, Cugini & Laskowski, 1999) which may in part be explained by the added cognitive load required to mentally infer simulated depth on a 2D computer screen (Westerman, Collins & Cribbin, 2005, p. 716).

There are reports of negative results of 3D visualizations (see review from Katifori *et al.* (2007)). Hearst (2009) reviewed the usage of 3D for IR (see Hearst (2009) section 10.10) producing a two paragraph section containing four studies. Teyseyre (2009) noticed the same lack of evaluation studies in the field of 3D software visualization; thus, 3D for IR and other applications is still in its early days. Hearst (2009) suggests 3D interfaces may improve with the evolution of hardware and software technology; however, current research suggests 3D does not produce superior performance as compared to 2D or purely textual displays.

Kobsa (2004) states that in some cases negative 3D IV performance results are due to the lack of other features such as an effective search tool, highlighting of search results, filtering, or navigation. These types of purely 3D displays are often

> *“complexly structured with many occlusions and obstructions of view. Without references such as a horizon and a consistent ground plane, information gained by overview and exploration is more difficult to come by”(Chalmers, 1993, p. 384).*

The landmark description of the BEAD information landscape (Chalmers, 1993) partly explains the problems users experience with 3D. Chalmers observes that “our experience is of a world with greater extent in the horizontal than the vertical: one might even call our everyday world ‘2.1–dimensional’” (p. 378). Although airplane pilots and underwater divers may develop a strong 3D perception, most of us are adequately served by a 2D eye-level horizontal plane. This helps explain observations of user difficulties when judging values in 3D scatter plots or point clouds (Kosara & Hauser, 2003), and findings that "changes in physical area on the surface of a graphic do not reliably produce appropriately proportional changes in perceived areas" (Tufte, 1983, p. 71).

These reports suggest that 3D should be used for intuitive navigation and appeal as opposed to precise value judgments. These types of VR applications should support navigation by offering conspicuous referential elements such as a static ground and/or walls. These were design choices made for the Virgilio music browsing application (Costabile, Malerba, Hemmje & Paradiso, 1998) where “the floor represents a type of music…the corridor provides access to different rooms, each one associated to a band” (p. 52). Chang (2003) applied a similar concept to a movie collection where automatically generated rooms represent genre and posters on the walls provide access to a specific movie (see Figure 3.4).

Figure 3.4: 3D movie collection browsing by genre (Chang & Said, 2003)

Representing a complex ontology in a 3D environment shows promise since 2D can be restrictive while 3D offers the possibility of a richer visual representation (Bosca, Bomino & Pellegrino, 2005; Bosca, Bonino, Comerio, Grega & Corno, 2007). 3D Ontology visualizations have "not yet been applied extensively" (Akrivi, *et al.*, 2007, p. 34); thus, knowledge as to its effectiveness is still inconclusive.

Studies that evaluate effects of visualization tools on performance and preference are grounded in the domain of HCI and usability. The following section provides a thorough overview of IV testing which informs the design of the controlled experiment in order to verify the thesis (see section *Chapter 2:Thesis*).

3.2.8 IV User Evaluations

IV research is still ongoing but "the role of visualization for IR systems…has not been proven" (Newby, 2002, p. 49). There are "hundreds" (Morse, Lewis & Olsen, 2002, p. 39) of information visualization systems in development but very little usability testing (Andrews, 1995; Chen & Czerwinski, 2000b). Simply put, "there is not yet a body of knowledge on information visualization evaluation" (Santos, Zamfir, Ferreira, Mealha & Nunes, 2004, p. 812), and it is not known which set of techniques are better suited to which types of tasks or how integrated these techniques should be.

IV testing is a highly complex endeavour since multiple factors and their interactions can create effects on various measures. An interactive visualization is a dynamic HCI that can be influenced in surprisingly significant ways by context (Jarvenpaa, 1989; Tversky, et al., 2000), the specific IR task demands (Gattis & Holyoak, 1996), verbal instructions (Oakhill & Johnson-Laird, 1984; Spivey, Tyler, Eberhard & Tanenhaus, 2001), and the user's internal understanding or mental model (Kieras & Bovair, 1984). The initial appearance of the visualization itself can affect the interpretation of its informational content (Zacks & Tversky, 1999). These complexities entail that the experimental design must "acknowledge and minimize these factors when evaluating the usability of a visualization method" (Ziemkiewicz & Kosara, 2008, p. 1275).

In the HCI tradition (Nielsen, 1993), IV applications are generally evaluated using one of four types of methods (Plaisant, 2004):

1. Controlled experiments comparing design options
   - The studies in this category compare different widgets (Ahlberg & Shneiderman, 1994a) or different visualization schemes (Modjeska & Waterworth, 2000first experiment).
2. Usability evaluation of a single tool
   - Studies aim to gather feedback and lead to design improvements (Sutcliffe*, et al.*, 2000a), what Byrd (1999) called *formative* evaluations.
3. Controlled experiments comparing two or more tools
   - The most common type of study. Studies aim to compare novel IV system to a state-of-art baseline (Cribbin & Chen, 2001; Swan & Allen, 1998; Veerasamy & Heikes, 1997).
4. Case studies of tools in realistic settings
   - The least common type of study. Studies produce "naturalistic" in-context results but tend to be time consuming, difficult to replicate and generalize (Trafton, Kirschenbaum, Tsui, Miyamoto, Ballas & Raymond, 2000)

All methods can produce rigorous and useful results but controlled comparative experiments offer the promise of replication and generalization. This type of data supports the production of IV design guidelines based on experimental results. This method can also provide statistical measures of IV effects on performance measures which can then be compared and combined using meta-analysis techniques (Rosenthal, 1991). These are desirable because statistical meta-analysis can reveal the suggestions from a body of literature as opposed to unitary result sets.

An in-depth literature review (Julien, Leide & Bouthillier, 2008) of controlled IV usability experiments showed design considerations include choice of experimental design, baseline system, test participants, the tasks being tested and their measurements (i.e., dependant variables). These are described in the following sections.

3.2.8.1 Experimental design: between vs. within subjects

In a classical experimental between-subject design two groups are formed and each is assigned to perform tasks on either the test or baseline system. This assumes control and treatment groups initially share similar significant attribute values (e.g., cognitive abilities, experience, motivation, etc.). Unfortunately it is not known which measurable traits to balance between groups (Swan & Allen, 1998, p. 180); as a result, it is always possible that performance differences between baseline and treatment group were present before the test. This is often addressed by randomly assigning each participant to a group and assuming characteristics will be evenly distributed across both groups.

A within-subject design, the most popular choice, solves the issue of balancing control and test groups, since it allows each subject to act as his/her own control. In its purest form, all participants are asked to perform all tasks on all systems, which demands more time from each subject. Since every subject acts as his/her own baseline this type of experiment does not require a separate control group; consequently, within-subject designs can produce solid data with half the users required by a between-subject design. Typically system/task order is counter balanced but this it is not always practical when testing a logically sequential range of IR tasks within a single collection (e.g., Cribbin *et al.* (2001)).

3.2.8.2 Baseline system

A comparable baseline can produce measures of IV tool effect on search performance as compared with the chosen baseline system, and allows the combination of results from multiple studies. Researchers use various types of baselines for different purposes. The review conducted by Julien *et al.* (2008) showed that the majority (22/31) of IV for IR evaluations compared their prototype with a text-only baseline system (e.g., Becks *et al.* (2002); Chung *et al.* (2005)).

Choosing existing commercial solutions as a baseline (e.g., Chen *et al.* (1998); Turetken *et al.* (2005)) provides information about IV acceptance among users. This is desirable since

> *"no matter how usable a software system proves to be during formative evaluations, its usefulness in a real world environment depends on the alternative software systems available and of course on their quality." (Turetken & Sharda, 2005, p. 178)*

3.2.8.3 Test Participants

Julien *et al.* (2008) revealed that using undergraduate students as test participants is by far the most popular option among the studies analyzed (26/29). Hedman (2004) attempted to verify if this user population is representative of other sampling frames, concluding that undergraduate students "can be an indicator of differences in the general population" (p. 358).

The level of pre-existing user experience with the tested tools is reported as a confounding factor on performance measurements. This concerns the fact that baseline systems are often known and used *a priori* by the test subject as opposed to the novel IV system which "could put the visual interface at a disadvantage, or create a need for extensive training" (Newby, 2002, p. 39). Novelty can also create an advantage for the tested system. The time spent by participants to familiarize themselves with an unknown tool may significantly affect performance measures. Effects may be in favour of the baseline system simply because participants were proficient in their use prior to the test. This partly explains reports of user preference for the IV tool without measurable effects on performance (see section 3.2.2 *Cognitive Affect*).

Attempts should be made to control the proficiency levels of the participants on both baseline vs. test system lest the results simply "suggest that (the) visualization cannot be adequately evaluated using only short term studies of novice users" (Sebrechts, et al., 1999, p. 9). Longitudinal studies can address this but they are expensive and actual training time is difficult to control beyond the walls of the experiment location. Some experimental contexts may allow the recruitment of reasonably novice users through pre-selection.

3.2.8.4 Tasks

Studies have shown that “different tools (are) better for different types of tasks” (Kosara & Hauser, 2003, p. 132); in other words, interaction between the type of task and the type of interface (Sebrechts, et al., 1999). Calls for the development of an IR task taxonomy have been made (Chen & Czerwinski, 2000a). A review from Li & Belkin (2008) shows there are numerous ways to describe and classify information retrieval tasks. IR task classification models provide dozens of task attributes and their values which can combine into hundreds of task types. This complexity is in part due to the interactive and iterative nature of the IR search process (see section 3.1 *Information Retrieval*).

3.2.8.4.1. Information Retrieval Tasks

The most recognized effort is Shneiderman (1996) who provides a taxonomy of seven elementary tasks performed with IV systems:

1. **Overview**: general knowledge of the contents of the collection and its major subject areas
2. **Zoom**: ability to focus search from general to more specific subject areas
3. **Filter**: equivalent to the logical operator not “NOT”, exclude specified classes of items from the search scope
4. **Details-on-demand**: ability to efficiently inspect groups or individual item contents
5. **Relate**: view relationships among items (e.g., semantic, co-authorship, co-citation)
6. **History**: inspect the search path explored
7. **Extract**: collect and save groups or individual items for later use or query refinement.

This taxonomy was used as a framework by Cribbin and Chen (2001) to design their test tasks in a collection of 200 articles. They describe four sequential tasks starting from task A) a broad theme yielding 20-25 documents, task B) a more specific aspect of the previous theme reducing the set to between six and ten documents, task C) within the previous set, find a pair of articles that discus two directly related subjects, and task D) find a pair of articles that discuss two directly related subject from a completely unrelated theme. This last task was deemed more difficult than task C “because the users must reorient themselves within the space without the benefit of progressive query refinement of the kind offered in tasks A and B" (p. 202).

Designing a range of IR task types to test is becoming an increasingly common empirical approach (Sebrechts, et al., 1999). This is also the current TREC (Voorhees & Harman, 2005) approach "in order to gain a more generalisable impression of the capabilities of an interface" (Cribbin & Chen, 2001, p. 201). Mann and Reiterer (2000) defined their tasks as "specific fact finding" when a clear success criterion is provided, and "extended fact finding" when the success criterion is diffuse requiring a longer broader search (p. 588). Pirolli *et al.* (2000) chose tasks from the Great CH1 '97 Browse-off (Mullet, Fry & Schiano, 1997) and organized them into the following task types (Pirolli*, et al.*, 2000, pp. 163-164):

- **Simple retrieval**: required finding a leaf node in the tree; e.g., "Find Lake Victoria."
- **Complex retrieval:** also involved finding leaf nodes, but involved some ambiguity and lack of familiarity; e.g., "Which army is lead by a Generalissimo?"
- **Local relational**: involved examination of several nodes that were reasonably close together in the tree structures; e.g., "Which religion has the most holidays in this list?"
- **Complex relational**: required examination of several nodes in disparate parts of the tree; e.g., "Which Greek deity has the same name as a space mission?"

3.2.8.4.2. Hierarchy Navigation Tasks

Barlow and Neville (2001) performed comparative experimental tests of various visualization glyphs (organization chart acting as a baseline, icicle plot, Treemap, and tree ring) for a file hierarchy. The range of tasks tested were restricted to comprehension of the file structure and file attributes as follows:

1. **File size**: Users had to select the three largest leaf nodes
2. **Tree topology**:
   - Users had to indicate level of a predetermined node
   - Users had to indicate the total number of levels in the tree

- Users indicated the deepest common ancestor of two predetermined nodes

3. **Node memory**: Users were asked to memorize the positions of two predetermined nodes, the view was closed and shown at a different size, users then had to indicate the positions of both nodes.

3.2.8.4.3. Task Complexity

Experiment designers may also consider the level of task complexity since "keeping tasks simple makes it easier to attribute differences in task performance directly to the different types of visualization, and helps eliminate confounding factors" (Kobsa, 2001, p. 129); however, "tasks should not be so simple that their ecological relevance is unclear" (Santos, et al., 2004, p. 814). Positive test results using basic perceptual tasks (e.g., which icon is closer to another) or simple known-item tasks (e.g., what is the capital of Norway?) may not generalize to common information tasks (e.g., what type of refrigerator brand should I purchase?). Some wonder "how frequently these low-level tasks actually occur in real-world settings, and how significant are they in the overall task solution process" (Kobsa, 2001, p. 129)? Choosing "higher-level tasks is perhaps more appropriate for a qualitative analysis" (Swan II, Gabbard, Hix, Schulman & Kim, 2003, p. 265) at the expense of strong statistical results, but allows the study to "simulate realistic seeking processes" (Grun, et al., 2005, p. 180) which could apply to real-world application domains.

Results may also be affected by test subject familiarity with the task topic because "individual performances can overwhelm browser design for the overall task" (Pirolli, *et al.*, 2000, p. 165). For example, a medical student may know *a priori* the location of a subject within the MeSH hierarchy and overwhelm the effect of an IV interface. Plaisant *et al.* (2002) also recognized this issue and "to avoid measuring users knowledge about the (topic) they were asked to find (e.g. kangaroos), (the researchers) provided hints to users (e.g. kangaroos are mammals and marsupials) without giving them the entire path to follow" (p. 62). Pirolli *et al.* (2000) went a step further and controlled the familiarity of each test subject with each search concept by asking the participants to "1) rate their familiarity with the term on a 7-point scale, and (2) to identify their top choices of categories for locating the answer" (p. 164). Unsurprisingly, their results confirm that users perform better when searching for terms that are familiar to them.

3.2.8.5 Dependant Variables

The choice of measurements expresses the effects of the IV interface as compared to the chosen baseline system. The vast majority of comparative IV experiments use various flavours of effectiveness and efficiency (Chen & Czerwinski, 2000a; Julien*, et al.*, 2008). Effectiveness expresses the level of task success (i.e., whether the question was accurately answered, to what extent) or using traditional information science measures of recall (i.e., number of relevant items found over total number of relevant in the collection) and precision (i.e., number of relevant items found over total items found). Efficiency indicates the cost to achieve success expressed in terms of resources (e.g., time).

The level of user satisfaction is often measured but using various methods and sources. For example:

- Becks *et al.* (2002) devised their own set of five Likert questions
- Chen *et al.* (1998) extracted satisfaction events using verbal protocol analysis
- Chung *et al.* (2005) used one Likert scale and written comments
- Grun *et al.* (2005) used the SUS quick and dirty usability scale (Brooke, 1996)
- Rivadeneira & Bederson (2003) adapted the *QUIS: Questionnaire for User Interaction and Satisfaction* (Harper & Norman, 1993)

In fact, in the review performed by Julien *et al.* (2008), all 12/31 studies measuring satisfaction did so using different measurement tools and/or methods. The disparity among operational definitions of this dependant variable makes it difficult to compare the results from multiple studies.

3.2.9 Information Visualization Summary

Section 3.1 described information retrieval knowledge and the elusive concept of browsing which is ill-supported by keyword search boxes and long result lists. This section concerned potential solutions offered by the field of information visualization.

Theoretical foundations of IV show that visual representations of data can be cognitively efficient when design utilize preattentive processing (section 3.2.1.1) and Gestalt laws of visual perception (section 3.2.1.2). Since the early 1980s, direct-manipulation interface metaphors (section 3.2.1.3) have used these concepts and their mass-market acceptance has shown visual interfaces are preferred by users as opposed to purely textual displays.

User preference for interactive visual interface metaphors has also been observed in the context of IV data metaphors (section 3.2.2). This preference may be partly due to positive first impressions created by these pleasing and dynamic interfaces. Positive first impressions positively affect perceived ease of use and usefulness, both of which are correlated with tool acceptance over time (section 3.2.2.1). IV tools offer much of their anticipated benefits when browsing is necessary to retrieve documents (section 3.2.3) in large and unfamiliar datasets (section 3.2.4).

Large datasets require structure to visualize and there are many existing collections whose contents have hierarchical subject organization (section 3.2.5); for example, semantic Web ontologies (section 3.1.3) and library collections (section 3.1.4). These are the types of collections this research proposes to visualize integrated with keyword searching.

Section 3.2.7 showed that virtual reality offers advantages for IV in terms of its immersive qualities and intuitive navigation for generations of 3D game users. Little research has been performed in VR for IV and the few negative reports may be due to the lack integrated search or filtering, and complex navigation without consistent horizon or ground plane.

This overview of IV ended with an in-depth review of usability studies (section 3.2.8) which showed that there is not yet an accepted testing method for these types of applications. Comparative IV testing is difficult because users require time to learn the novel visual IR tool as opposed to the text-only baseline. The tasks tested have much effect on performance measures and the best approach may be to test and range of IR task types including hierarchy navigation, simple and complex information retrieval.

Section 3.2.5 described how navigating large hierarchies can be demanding because searchers must remember all potentially interesting branches – just in case the information sought is not found along the initially chosen exploration path. This is analogical to gatherers who must continuously remember the location of promising food patches while they extract berries from one that is being depleted. The following and final section of this theoretical framework describes a human-information behaviour process as searchers *foraging* for information. This information foraging model frames the design of a hierarchically structured information retrieval system and offers specific measures to assess its quality.

3.3 Information Foraging

> *Internet users are seen as "informavores" maximizing gains of valuable information per unit cost associated to their searching efforts (Emond & West, 2003, p. 527)*

The information foraging (IF) model (Pirolli & Card, 1999) has generated new understandings concerning interactions between searchers and large information collections. Based on the optimal foraging theory (Charnov, 1976), inspired by Marcia Bate's berry-picking model (Bates, 1989) and related to ASK (Belkin, 1980), IF treats "adaptations to the flux of information in the cultural environment in much the same manner as biologists study adaptations to the flux of energy in the physical environment" (Pirolli & Card, 1999, p. 643).

Shneiderman and Plaisant (2005) and Emond and West (2003) describe Internet user behavior as foraging for information. Sandstrom (2001) shows

> *"that universal principles such as prey-choice models from optimal foraging theory (…) can be successfully applied in the bibliographic microhabitat to explain information seeking and use behavior"(Borner, et al., 2003, p. 50).*

IF represents information as being locally grouped in *patches* searchers sift through to select valuable items. A patch can be a document, a group of Web pages, a few words on the desktop, "piles of documents, file drawers, office bookshelves, libraries, or various on-line collections" (Pirolli & Card, 1999, p. 645). Any patch can also be seen as a member of one or many other patches until the level of semantic abstraction provided by the patch label is suitable for the task at hand. This dynamic definition of an information patch suggests different tools are necessary for different types of target patches; indeed a searcher often

> *"has to navigate from one…on-line collection or WWW site to another or from one search engine result to another. Often the person is faced with decisions much like (an) imaginary bird: How should time be allocated to between-patches foraging tasks versus within-patch foraging tasks" (Pirolli & Card, 1999, p. 646).*

IF suggests that the design of IR tools should provide the highest possible value for each unit of time spent using the tool. Information seekers cycle through moments of foraging a patch until little value remains at which point they move on to another promising patch. IF states that IR tools should:

1. aim to reduce navigation between patches and
2. maximize the extraction of value within each patch.

Value is defined as content harvested per unit of time. As a patch is mined the remaining potential value is surpassed by higher potential value of other patches. This creates a continuous and dynamic flow between evaluating the relevance of an item and the assessment of other patches (e.g., query reformulation) since "value is defined in relation to the embedding task, and this often changes dynamically over time" (Pirolli & Card, 1999, p. 645).

Designing an IR system which supports information foraging should allow the searcher to find more valuable information in less time. In practice this means an information retrieval interface that 1) offers efficient value assessment tools such (e.g., keyword-in-context, relevancy measures), 2) strives to reduce the time and effort to switch from groups of documents (e.g., LCSH subject labels) and their specific information items, 3) switch from one topic to another, and 4) extract/save valuable information.

The objective of IF is to design a useful and efficient IR system based on the assumption that most information seekers behave in a predictable manner most of the time. For example, if a result set yields little value, most searchers, most of the time, will move on to another set of search results, unless they've had their fill. Consideration of ease of learning and intuitiveness of the IR tool is implicitly promoted by maximizing value extraction. This becomes clear when considering that any time spent learning the tool is necessarily value lost; in other words, a pleasing and appreciated tool should be more easily and quickly learned which in turn will save time. This creates more value per unit of time.

One criticism of IF is based on its initial intent as a cognitive engineering model to automatically evaluate browsers (Pirolli, 1998) and Web-site design (Chi, Rosien, Suppattanasiri, Williams, Royer & Chow, 2003). As such, it tests "simulated users" (Emond & West, 2003, p. 530) assumed as "automatons" (Toms, 2002, p. 856). IF is an optimization model thus necessarily deals with trends in the behavior of populations of searchers which means it

> *"should not be taken as a hypothesis that human behaviour is classically rational, with perfect information and infinite computational resources. A more successful hypothesis about humans is that they exhibit bounded rationality or make choices on the basis of satisficing (Simon, 1955). (Pirolli & Card, 1999, p. 645)*

Katz and Byrne (2003) examined how principles from information foraging could be used to enhance the navigability of online stores. Their findings showed that Web site structure elements of *breadth* (i.e., number of menu options) and *information scent* (i.e., how suggestive labels are as to their potential value) both influenced browsing behaviour. Information scent is an important element of the IF model of searcher behaviour and deserves further description in the following section.

3.3.1 Information Scent

The concept of *information scent* has recently received much attention in relation to web site design (Card, Pirolli, Van Der Wege, Morrison, Reeder, Schradley & Boshart, 2001). Information scent describes how much a label suggests its content; in other words, “the ability of proximal cues to create in the mind of the user associations related to the content looked for” (Hornbaek & Frokjaer, 2003, p. 498). IF describes human-information behavior as people foraging for information in much the same way animals must balance the energy expended in finding and consuming food vs. the energy obtained. Information searchers are constantly making decisions about "what *information scent* (...) to follow, and they try to minimize how much work they must do" (Ware, 2008, p. 176).

Chi *et al.* (2007) measured the effects of document highlights of concepts related to user keywords on information scent. Their study showed that users were more efficient and accurate in finding, comparing, and comprehending material with their *ScentIndex* system as compared to regular reading of a paper book. Larson and Czerwinski (1998) evaluated the information scent of various breadth vs. depth Web site navigation structures. They showed that users were faster at finding targets within broader, shallower categories with distinctive category labels. Using Smith's measure of *lostness* (Smith, 1986), they also report users were lost in the hierarchies with the most levels. Pirolli *et al.* (2003) evaluated how hyperbolic views (see section 4.2.4.1 *Hierarchy Parts (Row 3, Column C)*) compared with Windows File Explorer (see Figure 1.1) when conducting visual searches. Using a within subject repeated measures design they compared the efficiency of users performing various types of IR tasks. They found that performance was affected by a combination of scent level (i.e., how familiar and suggestive the labels are of their content) and the system used (i.e. Hyperbolic vs. Windows File Explorer).

3.3.2 Ranked Results for Information Foraging

The classical ranked list of results is ill-adapted to information seeking behaviour in terms of patch foraging. Presented in a ranked list,

> *each element of the result-set is a singleton; relationships between elements cannot be determined, other than the common relevance of each one to some query. Therefore, identifying relevant patches requires an exhaustive search of the result-set"(Hoare & Sorensen, 2005, p. 235)*

A more suitable approach to IF has long been provided by pay-per-use database tools such as Dialog or OVID – query nesting allowing searchers to refine their queries by limiting their search to a subset of the previous search results.

The inadequacy of the ranked list is sometimes confounded by low precision result sets caused by inadequate initial queries (Jasen, Spink, Bateman & Saracevic, 1998). The searcher must spend time evaluating a significant proportion of irrelevant items until, depending on the size of the result set, cognitive capacity is exceeded and he/she quits the search in frustration often excluding valuable information. Since Web searchers rarely inspect beyond 40 individual items (Chen, *et al.*, 1998), searching through larger result sets is not well supported by ranked lists.

3.3.3 Integration of Search/Browse

IF environments should offer integrated search strategies since foraging "users interleave directed structured behavior (i.e., known-item searching) with opportunistic and unstructured behavior (i.e., browsing)" (Olston & Chi, 2003, p. 181). They report that several participants stated that the test interface "permitted them to narrow down some aspects of the task with keywords, while honing in on other aspects by browsing" (p. 193).

Hoare and Sorensen (2005) also define IF as an iterative integration of keyword searching and browsing. This conclusion was reached after considering search engine query log analysis studies (Jasen, et al., 1998; Silverstein, et al., 1999) which suggest users are likely to use adhoc search strategies which do not strictly conform to either search-by-query or search-by-navigation strategies (Jul & Furnas, 1997).

3.4 Theoretical Framework Summary

To summarize the presented theoretical framework, the following predominant and convergent themes should be considered:

- Information retrieval is an iterative process which integrates periods of specific searching for known-items and unstructured browsing.
- Information visualization is promising for browsing in large and unfamiliar datasets because it visually reveals patterns in the data.
- Browsing large datasets requires structure and hierarchical subject structure is currently part of many existing collections such as libraries and semantic web ontologies.
- Virtual reality is promising for hierarchy visualization and navigation because the 3rd dimension can convey subject hierarchy depth while allowing the use of 2D visualization techniques for each individual hierarchy level.
- Users interacting with information can be described as foraging for information which involves iterative cycles of specific item value extraction followed by assessment and navigation to other more promising patches of information.

These elements frame the design of a solution to the issue of large hierarchy navigation for information retrieval presented in the following section.

Chapter 4: DESIGN SPACE

The purpose of this research is to develop and test an interactive information visualization tool to support information foraging in a semantic hierarchy. Specifically, a prototype application based on a library bibliographic collection and its LCSH organization are chosen as a test case (see section 3.1.4.2 *Browsing Subject Structures*). This context may offer valuable insights into issues that will eventually be faced by the Semantic Web initiative as it inherits significant ontologies from the library world (see section 3.1.3 *Semantic Web*)

The following sections enumerate specific design requirements stemming from the previous theoretical framework. This is followed by a review of existing commercial and experimental IV solutions for hierarchy navigation. Finally, the proposed Subject Explorer 3D (SE-3D) is described and compared to these existing systems.

4.1 Solution Design Requirements

The purpose of the design is to facilitate information retrieval by providing visual explorative access to hierarchically organized information collections. Prototype development requires specific design requirements which are believed to support the stated purpose. Design requirements are provided below within the theoretical framework elements of information retrieval (section 3.1), information visualization (section 3.2) and information foraging (section 3.3).

4.1.1 Information Retrieval

> *…shift smoothly from browsing to searching and back again, all the while maintaining continuity and context by reference to the static landscape framework. (Chalmers, 1993, p. 387)*

Chalmers (1993) describes his ideal IR system which perfectly integrates keyword searching and visual browsing without disorienting the searcher. Knowledge from IR research strongly suggests specific keyword and browse searching should be tightly integrated so users can effortlessly switch between the two (see section 3.1.2 *Integrate*

Search and Browsing). This design requirement is also consistent with information foraging theory (see section 3.3) since it supports relevance or value assessments while facilitating transitioning between information patches (i.e., from broad exploration to specific keyword matching and back).

Specifically, IR research suggests that (Driessen, et al., 2006, p. 217):

1. ideal information retrieval systems support both search strategies and allow the user to switch between them without effort, and
2. information visualization may be more useful if querying is tightly integrated with browsing.

When first shown an interactive IV interface users may have "no idea where they should start to look for interesting features" (Abello, *et al.*, 2006, p. 7) but this can be alleviated by providing the ubiquitous keyword search box. There is evidence from text-only interface research that when semantic hierarchy information is combined with keyword search, results are slightly improved over using either on its own (Henzler, 1978; Lancaster, 1986; Markey, Atherton & Newton, 1982).

Building a <u>purely</u> visual information browsing environment may very well produce low performance scores due to the lack of keyword search features. For example, Katifori *et al.* (2006) showed that locating a specific subject label may be accomplished by browsing a visualized semantic hierarchy but it is "much quicker and more effortless to do so using a search tool" (Akrivi, *et al.*, 2007, p. 34). Kobsa (2004) compared the performance of four kinds of hierarchy visualizations and concluded that extremely low scores of the Tree Viewer (Kleiberg, van de Wetering & van Wijk, 2001) and BeamTrees (van Ham & van Wijk, 2003) "have mostly been caused by a lack of functionality beyond the pure visualization" (p. 16).

An IV interface should consistently offer "equal facility in search and browse along any criteria the user has in mind" (Smith, *et al.*, 2006, p. 798). Design should strive to seamlessly integrate keyword searching and visual browsing features throughout the application (Yee, et al., 2003, p. 402); in other words, enable the user to "understand the structure (...) as well as attributes of data" (Kleiberg, et al., 2001, p. 88).

4.1.2 Information Visualization

IV is useful when it explicitly reveals data patterns that would otherwise require costly analysis. Revealing patterns is an iterative and explorative process best supported by an interactive search environment. Interacting and navigating in a hierarchy visualization application requires interface features that facilitate manipulation of a large hierarchy and its inspection. Design requirements derived from this IV knowledge are described in the following sections.

4.1.2.1 Reveal Patterns in the Data

IV is concerned with revealing patterns in large datasets (see section 3.2.4 *Large and Unfamiliar Datasets*). As seen in section 3.2.5 *Hierarchy Visualization*, all ontologies are in part semantic hierarchies and a visual representation of an information collection could explicitly represent this pattern of broad to narrow relations between groups of information items. Hierarchical structures are useful because they lessen the learning curve of the application by facilitating the formation of the users mental model of the application (see section 3.1.4.2 *Browsing Subject Structures*).

The Bradford's Law pattern of subject assignments (see section 3.1.4.2 *Browsing Subject Structures*) can also be explicitly revealed by an IV representation of a collection. The visual area allotted to subject classes can be relative to their respective predominance in the collection; in other words, the number of items assigned to each subject should be directly proportional to the size of the visual subject symbol (Tufte, 1983, p. 56). These visual subject areas should also be clearly identified with their textual labels (Tufte, 1983, p. 56).

Representing Bradford's Law pattern of subject assignment provides a mechanism to simplify the otherwise overly complex hierarchical structure by visually representing only the few most significant subject classes at each level. This means that at any level of subject abstraction, although a particular level of the subject hierarchy might have dozens of narrower subjects, only the few assigned to many items should be visually represented. Those that contain few items can be visually omitted with minimal impact on the accuracy of the collection representation.

Sacrificing some information in order to gain simplicity is sometimes justified (Moya-Anegón, Vargas-Quesada, Herrero-Solana, Chinchilla-Rodríguez, Corera-Álvarez & Munoz-Fernández, 2004, p. 136; Small, 2000) since "what is sought in designs for the display of information is the clear portrayal of complexity" (Tufte, 1983, p. 191). This could be described as an instance of the "empty world hypothesis" (Simon 1986, 209) where most areas of the subject hierarchy contain a tiny fraction of the collection, and only a tiny fraction of the hierarchy needs to be taken into account for a "tolerable description" (p. 209) of the collection reality. A similar philosophy is applied in the PNASLINK application (White, Lin, Buzydlowski & Chen, 2004) to emphasize "only the most prominent links" (p. 5299).

4.1.2.2 Interactive Search Environment

The advantages of direct manipulation interactive interfaces (see section 3.2.1.3) are beneficial in the context of an IR environment. Through intuitive exploration of an appealing display the user constructs a mental model of the information in the visualization (Eades, Lai, Misue & Sugiyama, 1991). This permits examining the environment from multiple paths which has been shown to decrease “the specificity and increase the flexibility in the environment” (Sebrechts, 2005, p. 153). Using an interactive environment, users can explore, discover, and analyze information (Nguyen & Zhang, 2006, p. 981) to “immerse themselves in a body of information” (Toms, 2002, p. 855).

4.1.2.3 Hierarchy Visualization Navigation

Katifori *et al.* (2007) provide an extensive review of ontology visualization tools and their predominant design elements. Most of the applications support retraction and expansion of subjects from broad to narrower terms, change of viewpoint, rotation and zooming. The authors also note that overview tools and back and forward options are also useful navigation aids (p. 35).

4.1.3 Information Foraging

An optimized foraging environment should save time since reducing navigation of any kind should result in a net time savings. Assuming total relevant information is an imposed parameter of the information collection, the interface can only maximize within patch value by reducing time to find, evaluate and retrieve relevant items.

The ideal is a system which provides exactly the information needed at the exact moment the need arises. In practice this means providing the searcher with the most relevant information at a given moment. This would also mean minimizing the cost of getting more information "related to something already discovered (...) sometimes called *drilling down*" (Ware, 2008, p. 176).

It is useful to constrain the dynamic definition of a patch (see section 3.3). In the context of a semantic hierarchy, the most general patches are the top level classes (e.g., science, arts, literature, etc.). Within-patch foraging refers to inspection of either

1. labelled routes to more specific subject patches (except the most specific subjects or leaf nodes), and/or
2. individual bibliographic records (the most specific patch) explicitly assigned to the current subject.

Navigation between patches refers to switching between subjects and/or sets of bibliographic records whether on the same level or from broad to narrow subjects. Maximizing the value of a patch refers to the time required to assess the relevance of elements of a single patch (i.e., bibliographic records or more specific subjects).

4.1.3.1 Information in Space

> *Just as explorers need maps and compasses to aid their travel, so too do (searchers) need effective aids for exploring (information). (Fox, et al., 2006, p. 53)*

A mapped i*nformation space* may be a useful metaphor for online information foraging environments. IF describes searchers mining patches in succession with a constant evaluation of navigational costs to reach one identifiable area of information to another (Pirolli & Card, 1999). According to Lakoff *et al.* (1980), these kinds of spatial analogies are fundamental to human reasoning, which is "ultimately grounded in human experience gained by interacting with the environment" (Ware, 2008, p. 63). Hierarchy navigation has been described as gardening operations (Robertson, *et al.*, 1991) where the user can manually prune and grow the view of the tree. This would "concretize the abstract" (Sebrechts, 2005, p. 139) and may lower information extraction costs since searchers can manipulate more information with less effort.

An adequate spatial representation of an information collection may reproduce positive and pleasurable effects long associated with traditional physical libraries. A spatial analogy applied to online collections implies that "imperfect information at intermediate locations is used by the forager to decide on paths through a library or an on-line text database to target information" (Pirolli & Card, 1999, p. 646). LIS professionals and library users know that when moving freely among the books and other resources

> *we walk among organized, labeled bookshelves, we get a sense of the information space—we take in clues, perhaps unconsciously, as to the scope of the collection, the currency of resources, the frequency of their use, etc. We also enjoy unexpected discoveries. (Dushay, 2004, p. 1)*

Katifori *et al.* (2007) warn against movement and rotation of the visualization by the user as opposed to the user moving around a static visual structure. They state the former may disorient the user and "does not help the creation of a cognitive model of the ontology as nodes continuously change position" (p. 35)

4.1.3.2 Intuitive Navigation in 3D Space

A space designed for IF could make good use of the 3rd dimension offered by virtual reality interfaces (Section 3.2.7: *Virtual Reality for IV*). This may reduce between-patch navigation by providing point-to-move commands such as those provided by first-person avatar games (e.g., Doom, Call of Duty), and direct manipulation grouping and filtering capabilities offered by point-to-select. VR has the potential to make software easier to explore and learn because it

> *"simulates the way we learn about the world around us in real life. The transfer of knowledge can allow us to absorb spatial information with little or no training in a virtual prototype"(Shiaw, Jacob & Crane, 2004, p. 133)*

Using 3D permits the communication of hierarchical relationships using depth (see Chapter 2: *Thesis*) to optimize screen usage and offer a hierarchy overview. This also allows each hierarchy level to be represented by a spatial display (see section 3.2.6 *Spatial Displays*) where subjects are placed "semantically close if they share many objects" (Fluit, et al., 2003, p. 39). This graphical representation of the semantic relationships between topics provides an overview of each hierarchy level and "improves a user's ability to find relevant information when foraging" (Hoare & Sorensen, 2005, p. 236).

Users of tools are more likely to accept new designs if they recognize features from prior exposure with similar tools (Bornstein, 1989; Lidwell, Holden & Butler, 2003, p. 70). Beyond supporting easy switching between specific and browse searching (see section 4.1.1 *Information Retrieval*), the ubiquitous keyword searching paradigm should also be provided since "it is arguable that keyword searching has been the access mode to which most users have become accustomed" (Tang, 2007, p. 1999).

4.1.3.3 Subjects as Patches in Static Space

When a keyword search is performed the results will often contain a few subject headings found in many information items; in other words, most items will cluster around a few popular subject headings (see section 3.1.4.2 *Browsing Subject Structures*). Using search keywords, the information seeker prunes the visual semantic hierarchy and result clusters create visually conspicuous information patches acting as landmarks.

In a static visual structure fixed in space (see section 4.1.3.1 *Information in Space*), keyword searching matches would consistently appear at the same locations and drawn in a predictable manner. These known patches become stable and conspicuous landmarks that support navigation in physical space (Lynch, 1960) and may do the same in information spaces (Dillon, McKnight & Richardson, 1990). Travelling back to a previously visited object in space has been shown to be more effective when the application maintains "a constant positioning of the nodes and allows quick browsing at the same time" (Akrivi, *et al.*, 2007, p. 34)

An interactive yet static visual metaphor means that the representation does not significantly change through time. The visualization serves as an overview of the major collection subject terms which should not change rapidly. Analogically, although trees fall and grow all the time, barring catastrophic events, it takes years for a forest to look significantly different. This is important since

> *“an inherent advantage of spatial visualizations (is) that users can incrementally build a mental map of places and locations and navigate by visual memory, (which would) be lost (if) the information space continually changed its form” (Andrews, 1995, p. 97).*

This allows users “to develop new habits (because) the interface (is) predictable…information appears in the same place every time the program is used” (Hornbaek & Frokjaer, 2003, p. 490).

4.1.3.4 Textual Labels as Named Patches

> *Words and pictures belong together. Viewers need the help that words can provide. (Tufte, 1983, p. 180)*

As Tufte (1983) suggests above, IV application design does not proscribe the use of text. Fekete and Plaisant (2003) state that text labels are not preattentive (see section 3.2.1 *Theoretical Foundations*) but "important to understand the context in which visualized data appear" (Akrivi, *et al.*, 2007, p. 37). Subject indexing (see section 3.1.4) creates a semantic hierarchy of textual subject term labels each acting as conspicuous attractors shown even at a distance (Darken & Sibert, 1996), and “conceptually, the document is stored within the category label” (Hearst & Karadi, 1997, p. 250).

These named patches may support value extraction by providing a quick preview of neighbouring contents. This supports continuous and inexpensive assessment of remaining potential value of the current patch versus navigating to other promising patches.

4.1.3.5 Visited Patch History

An IF interface should visually represent past search paths; indeed, "information exploration is inherently a process with many steps, so keeping the history of actions and allowing users to retrace their steps is important" (Risden*, et al.*, 2000; Shneiderman & Plaisant, 2005, p. 597; Sutcliffe*, et al.*, 2000a). Accessible search history features support patch value assessment by facilitating travelling back to known high value patch after assessment or mining of other patches.

4.1.3.6 Overview of the Patch Structure

Providing an overview of the information space is one of the most popular recommendations for designers of hypertext systems (Borner, et al., 2003). Gestalt theory (see section 3.2.1 *Theoretical Foundations*) emphasizes that the whole is larger than the sum of its parts, and parts (details) are more meaningful when viewed within the whole (the context). Providing an overview of the information collection may allow "the user to quickly grasp the overall gestalt of the data and begin to focus on regions of interest" (Brady, Pixton, Baxter, Moran, Potter, Carragher & Belmont, 1995, p. 489). Yee *et al.* (2003) present a broad overview of their image collection to suggest multiple stating paths for exploration and quickly "familiarize the user with the high-level information structure of the collection" (p. 403).

Bederson (2003) reports that the absence of an overview of a hierarchy made users extremely uncomfortable and "the provision of adequate overview strategies is a useful criterion to judge (IR) interfaces" (Shneiderman & Plaisant, 2005, p. 592). When people explore hierarchies, repeatedly switching between different levels of the hierarchy can become disorienting and confusing (Bederson, Hollan, Stewart, Rogers, Vick, Ring, Grose & Forsythe, 1998). This "could be eliminated if the interface provided an idea of how the immediate details fit in with the overall context of information” (Turetken & Sharda, 2005, p. 274).

Sometimes it is useful to assess an information item in its global context (Dumais, Cutrell & Chen, 2001). For example, a search for “River Jordan” would yield, at the very least, biblical accounts of the River Jordan, its geological and hydrologic description and geopolitical issues involving Israel, Jordan and the Palestinian Authority. An IR interface which explicitly represents these multiple contexts might offer an advantage over a ranked list of results with these multiple and mixed contexts.

For a library collection, an overview is provided by subject hierarchy organizations (e.g., LCSH). They provide searchers with groupings of documents at multiple levels of abstraction (e.g., Science, Physics, Quantum Physics, Experimentation, etc.). This facilitates between-patch navigation in the same manner that a geographical map helps travel from one country to another, an unknown city and neighbourhood. This type of collection overview may be enhanced if the visual screen area for each subject is proportional to the number of information items assigned which may draw greater attention to subjects "with wider applicability" (Smith, *et al.*, 2006, p. 799).

4.1.3.7 Relevance Sorted List

Relevance sorted lists of short results is the primary means of within-patch value extraction in current Web and library search tools. For the purposes of this research the search box and relevance sorted list of results will be identical and available on both test and baseline system. The only difference between the systems is the presence of the 3D visualization of the semantic hierarchy and its navigational features.

4.1.4 Design Requirements Summary

The design requirements were presented in the context of the theoretical framework and they are now summarized and listed as follows:

- The visualization of an existing semantic hierarchy must be integrated with keyword searching.

- The solution should simplify the complex semantic hierarchy by trimming the structure based on the Bradford's Law pattern of subject assignment.
- The solution must be interactive and offer basic hierarchy navigation features such as expansion/retraction of subjects, change of viewpoint, rotation and zooming.
- The visualization of the semantic hierarchy should be static in 3D space i.e., the user explores the structure by moving/flying around it.
- Subjects are represented as labelled shapes sized proportional to their relative importance in the collection
- The searcher can inspect his/her exploration path history
- Bibliographic records belonging to individual subjects or matching from a keyword search are shows in a ranked result list

These requirements may be partly or fully met by existing solutions. The following section provides an extensive review of hierarchy visualization systems. None of them meets all the above requirements and very few were ever tested in a controlled setting; however, they provide relevant information concerning working designs potentially recognized by users.

4.2 Review of Existing Systems

This section provides a review of interactive experimental and commercial hierarchical visualization applications for information retrieval. All systems described aim to represent single or multi-level hierarchies of information items. Information is defined broadly and includes full-text, bibliographic records, short labels, multi-media or computer file names. The review is restricted to applications cited in peer-reviewed publications.

4.2.1 A taxonomy of hierarchy visualization applications

Existing IV taxonomies (e.g., Lohse, Rueter, Biolsi & Walker, 1990; Shneiderman, 1996; Tory & Moller, 2004) aim to organize all types of visualizations and are too broad for a research constrained to hierarchy visualization. Bosca (2007) lists four types of ontology visualization techniques (i.e., network, tree, neighbourhood, and hyperbolic) based on the kind of visual interface technique used to represent the relations between concepts. Based on the design requirements stated in section 4.1.4, this research offers a taxonomy of interactive hierarchy visualization techniques which is described in the following section.

Categorization criteria selection is based on the review by Katifori *et al.* (2007) and research by Smith *et al.* (2006). The latter remarked that visualization research has progressed

> *"primarily in two ways: static, pre-computed visualizations of the large dataset itself, or more dynamic visualizations of a smaller, post-query subset, usually of a few hundred items resulting from an initial search on the original larger dataset" (p. 798).*

This suggests IV applications vary in terms of the percentage of the total information collection they visually represent and how users can modify the visualization through keyword search or filtering.

The resulting taxonomy of hierarchy visualization applications is characterized by combinations of amount of hierarchy overview (see section 4.1.3.6 *Overview of the Patch Structure*) and level of search/browsing integration (see section 4.1.1 *Information Retrieval*). These two criteria and their finite set of ordinal scale values are defined as follows:

- Number of displayed hierarchy levels: describes to what extent the application provides a complete hierarchy overview using one of the following values:
    1. The application offers strictly a single level hierarchy (also called *flat classification*)
    2. The application allows users to interactively navigate between levels but displays only one level at a time
    3. The application displays parts of the hierarchy
    4. The application provides a complete overview of the hierarchy
- Keyword Searching Integration: describes to what extent the application integrates keyword search in addition to the visual representation using one of the following values:

1. No keyword searching features
2. User keyword(s) restrict the corpus to be visualized
3. User keywords are dynamically searched within the labels of the structure and affect the visualization but the actual information items are not part of the structure
4. User keywords dynamically affect the visual representation which fully integrates structure and content

Informed by the original publication and the live application when available, each hierarchy visualization application is categorized in the resulting 4x4 taxonomy shown in Figure 4.1.

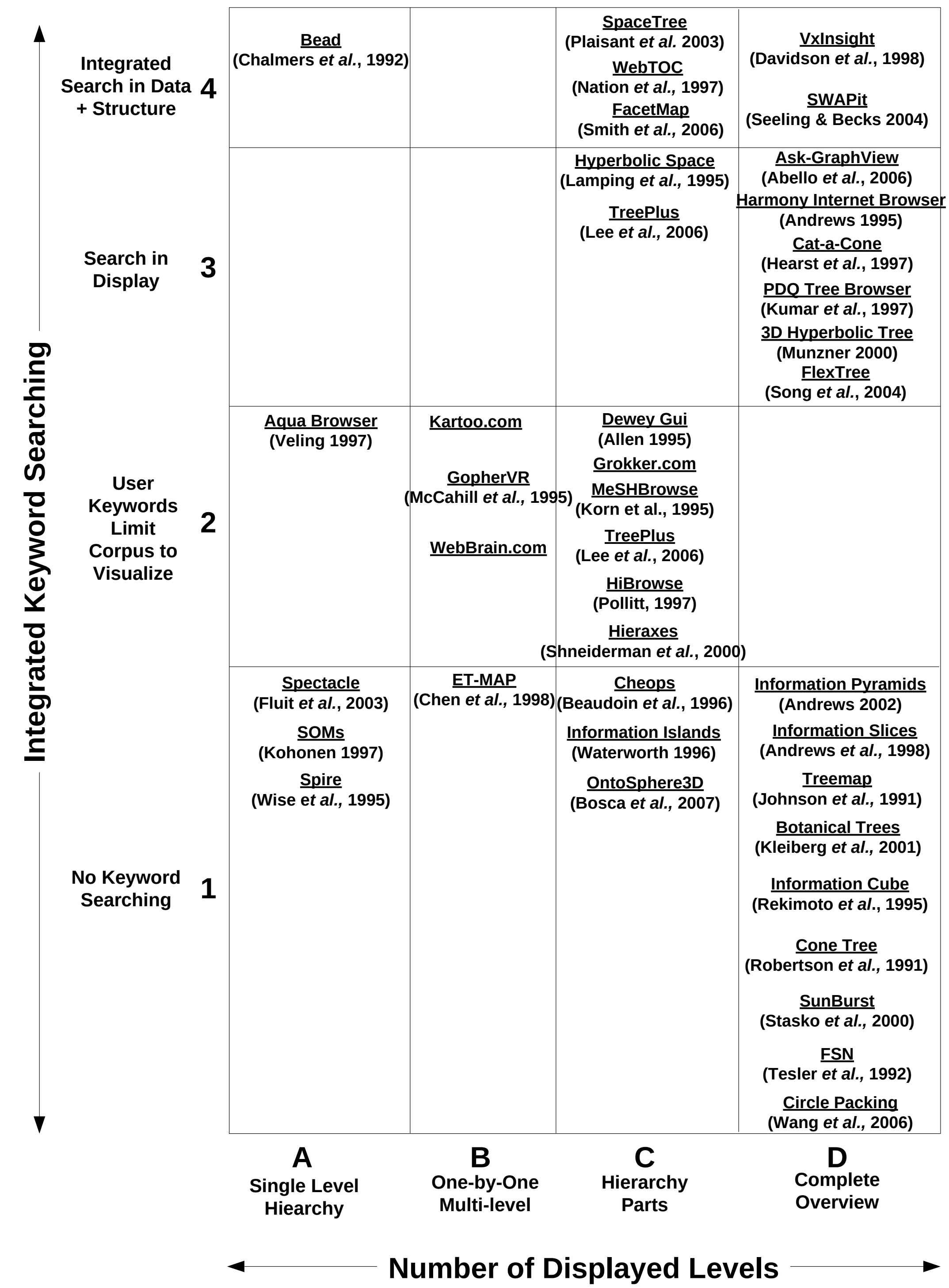


Figure 4.1: Categorization of interactive hierarchy visualization applications

Figure 4.1 reveals that over 43% of the applications (14/32, see row 1, columns A to D) were developed strictly as structure browsing tools with no keyword searching features. As well, 78% (25/32, see rows 1 to 4, columns C and D) offer visualisations of more than one hierarchy level in a single display. Only two applications (see row 4, column D) offer a full overview of the hierarchy and fully integrated searching features within the display and contents. This suggests there is room for novel applications of this type.

By row and then column, the following sections briefly describe the applications pertinent to this research contained within each cell of Figure 4.1. Further differentiation between the SE-3D solution design and comparable applications is also provided. The original version of Figure 4.1 was produced in 2006 and updated in early 2011. The patterns shown by the original version were not changed by the update. These later additions are not included in the detailed descriptions provided by the sections below. They include:

- ASK-GraphView (Abello, *et al.*, 2006)
- OntoSphere3D (Bosca, *et al.*, 2007)
- Botanical Trees (Kleiberg, *et al.*, 2001)
- TreePlus (Lee, Parr, Plaisant, Bederson, Veskler, Gray & Kotfila, 2006)
- FacetMap (Smith, *et al.*, 2006)
- FlexTree (Song, Curran & Sterritt, 2004)
- Circle Packing (Wang, Wang, Dai & Wang, 2006)

4.2.2 No Keyword Searching (Row 1)

Applications in this section were developed strictly as structure browsing tools with no keyword searching features. Some provide varying levels of visual hierarchy overview.

4.2.2.1 Single Level Hierarchy (Row 1, Column A)

The three applications described in this section offer a spatial representation of a single level hierarchy (also called flat classification). The visual size of each classification object suggests its relative prevalence and its placement reflects its pair-wise similarity based on lexical analysis.

4.2.2.1.1. *Spectacle*

Spectacle (Fluit, et al., 2003) provides a unique way to represent items with multiple subject terms i.e., that belong to more than one class or hierarchy branch. Users select initial subject terms of interest which the application represents as labeled dots surrounded by offshoots leading to groups of documents assigned exclusively to the term or in conjunction with other terms (see Figure 4.2).

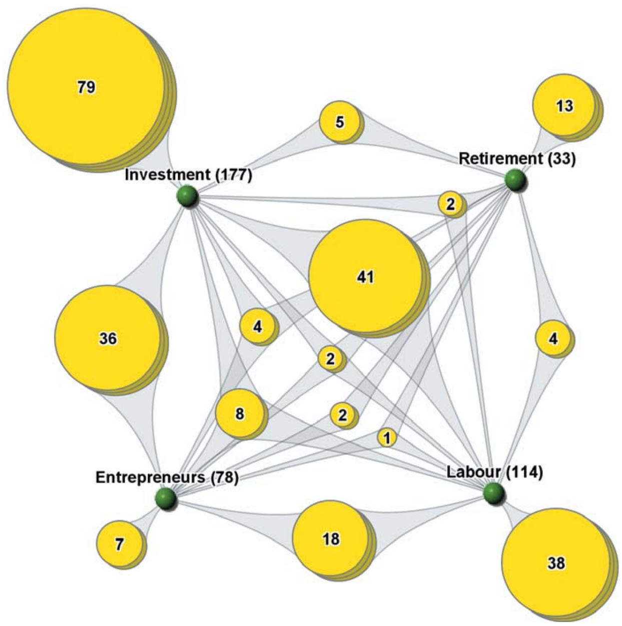


Figure 4.2: Spectacle Ontologie Visualizer (Fluit, et al., 2003)

For example, Figure 4.2 states the term “Retirement” (green dot, top right) has been assigned to 33 documents of which: 13 were exclusively assigned the term (yellow circle, top right), four were assigned both “Retirement” and “Labour” (yellow circle, middle right), and one document was assigned both previous terms as well as the term “Entrepreneurs” (small yellow circle close to “Labour” green dot), etc. The solution is not scalable since the display becomes cluttered with multiple overlapping offshoots when too many subject terms are selected.

Spectable is unique because it addresses the issue of representing multiple assigned subject terms without creating a copy of the item for each subject term assigned. The application also provides a practical layout method by placing subject terms closer if they share many documents (Fluit, et al., 2003, p. 39), as opposed to computationally demanding calculations of semantic distance based on word occurrences.

4.2.2.1.2. Self-Organizing Maps (SOM)

Self-Organizing Maps (Kohonen, 1997) are computer generated visual representations of a full-text documents organized in a single level of clustered groups. The visualization algorithm generates a 2D tile map such as the one shown in Figure 4.3.

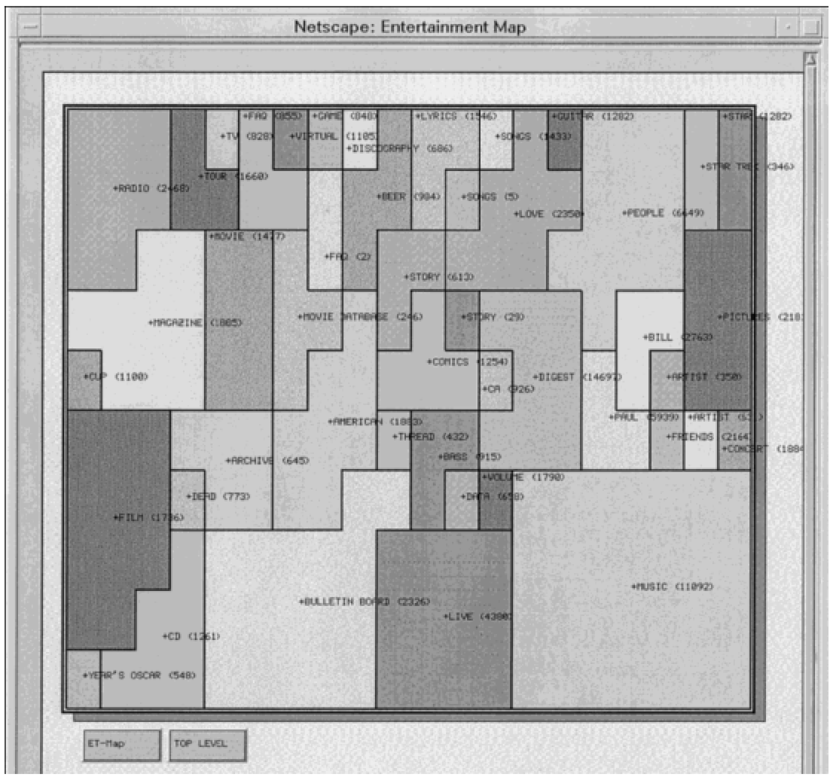


Figure 4.3: Self-Organizing Map (Chen, et al., 1998)

For example, Figure 4.3 is an organization of entertainment related Web sites. The visualization shows a large group (lower right) labeled "MUSIC" which contains 11 092 documents. Groups are placed according to their pair-wise semantic similarity, labels are statistically chosen as being most representative of the group, and a "title can be found in several places, depending on how many concepts or terms it associates with" (Lin, 1997, p. 50).

This kind of visualization may help "the viewer clarify some terms without even reading any document" (Lin, 1997, p. 46) and suggest valid search terms for the collection. It also provides a sense of the subject distribution within the collection since the area of each group is proportional to the number of documents it contains.

SOMs suffer from labelling issues associated with unsupervised clustering (see section 3.2.5 *Hierarchy Visualization*) and offer only a single level hierarchy. This application is pertinent to this research since it provides a systematic and informative way of determining group membership, placement and size.

4.2.2.1.3. SPIRE

Spire (Wise, et al., 1995) was the first clustering and spatial visualisation application developed at Pacific Northwest National Laboratory and now commercialized by ThemeMedia (see Figure 4.4).

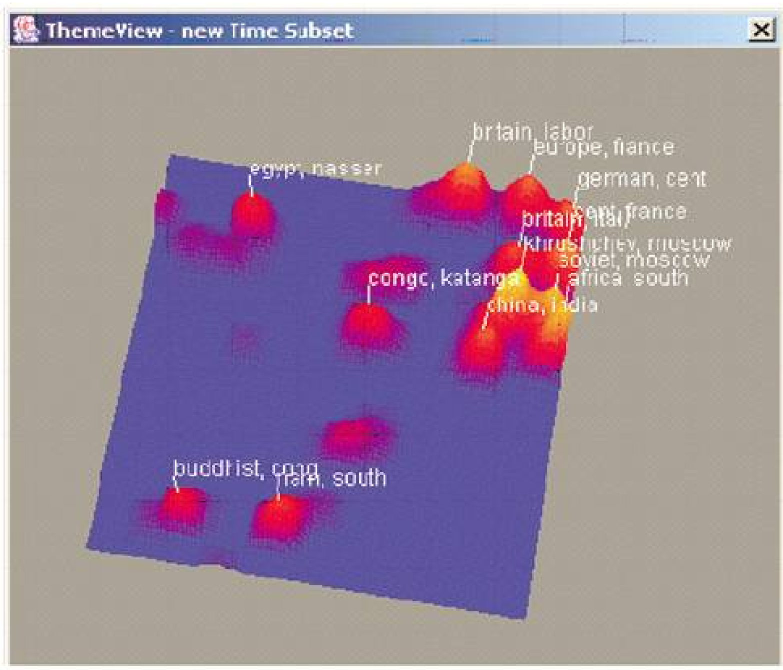


Figure 4.4: SPIRE in ThemeScape (Hetzler & Turner, 2004)

SPIRE suffers from the same problems as SOMs (see section 4.2.2.1.2) namely: labeling issues and single level hierarchy. This application is historically pertinent since it inspired a generation of spatial visualizations.

4.2.2.2 One-by-One Multi-Level (Row 1, Column B)

ET-MAP (Chen, *et al.*, 1998) adds a form of multi-resolution zooming (see section 3.2.4.1 *IV Techniques for Large Dataset Navigation*) to SOM visualizations (see section 4.2.2). As Figure 4.5 suggests, the user is presented with an initial SOM, he/she then clicks on a region of interest. This removes the current SOM and replaces it with sub-map showing "greater informational resolution through a finer degree of categorization" (Borner, et al., 2003, p. 20). Lin (1997) states that

> *"the user of such an interface can decide at what level of details the map display should be shown, knowing that the levels can always be changed as needed." (p. 50)*

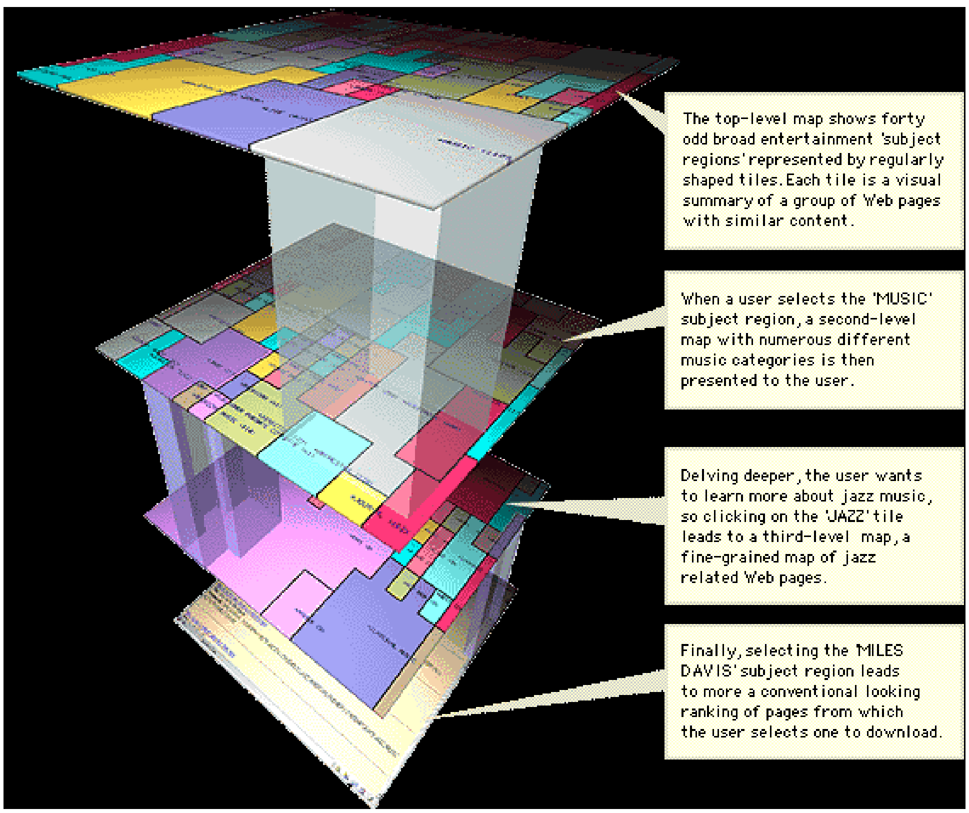


Figure 4.5: ET-MAP (Chen, et al., 1998)

Like SOMs, ET-MAP suffers from labelling issues and usability studies indicate that users tend to get lost when browsing multi-level SOM maps and continued to prefer to use a conventional text-based alphabetic hierarchy (Chen, *et al.*, 1998). This may be due to the fact that ET-MAP does not provide an overview of the hierarchy nor information about the users' current location within the structure. This experimental data are pertinent to this research as it suggests the importance of an explicit representation of the user's current location within the hierarchy in order to stave off cognitive disorientation and support branching decisions.

4.2.2.3 Hierarchy Parts (Row 1, Column C)

The two visualizations described in this section display parts of multi-level hierarchies by collapsing certain branches into a single icon or by selective occlusion.

4.2.2.3.1. Cheops

The Cheops application (see Figure 4.6) is intended for "browsing and exploration tasks in complex hierarchical information structures. It is not intended for use in analyzing these structures" (Beaudoin, Parent & Vroomen, 1996, p. 88).

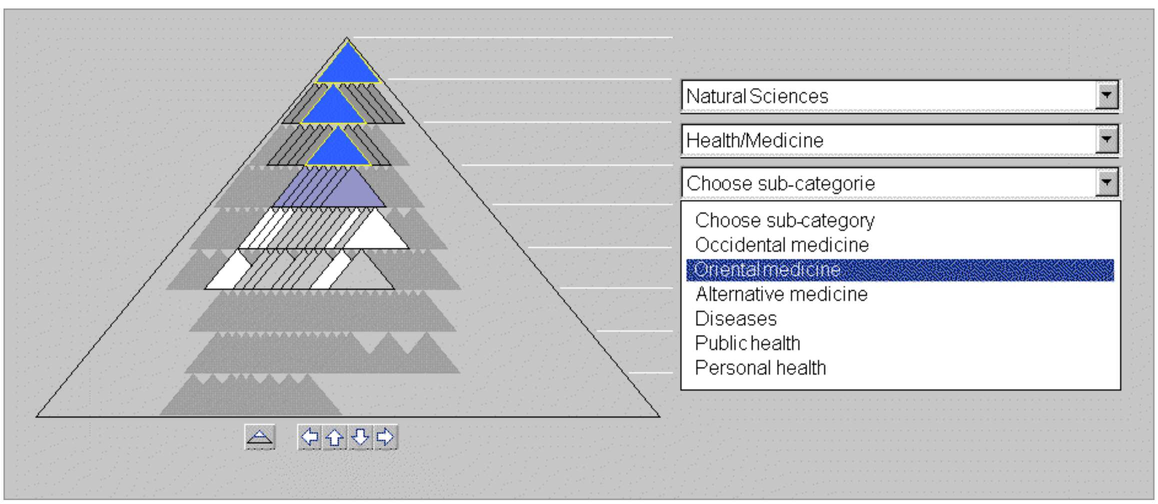


Figure 4.6: CHEOPS (Beaudoin, et al., 1996)

As shown in Figure 4.6, Cheops provides a pyramid-like visual representation divided into multiple layers of subject specificity. Cheops is unique because it achieves a compact display by partially occluding non-selected branches. Informal usability tests showed users required half an hour to understand the rich interface but did see potential in the prototype application (Beaudoin, et al., 1996, p. 91).

The Cheops prototype was developed at the Computer Research Center of Montreal (CRIM) and is pertinent to this research because it offers a unique approach to 2D screen space optimization using overlapping shapes.

4.2.2.3.2. Information Islands

Information Islands (Waterworth, 1996) uses a pre-existing subject hierarchy and attempts to provide an aesthetically pleasing spatial visualization of a few hierarchy levels (see Figure 4.7)

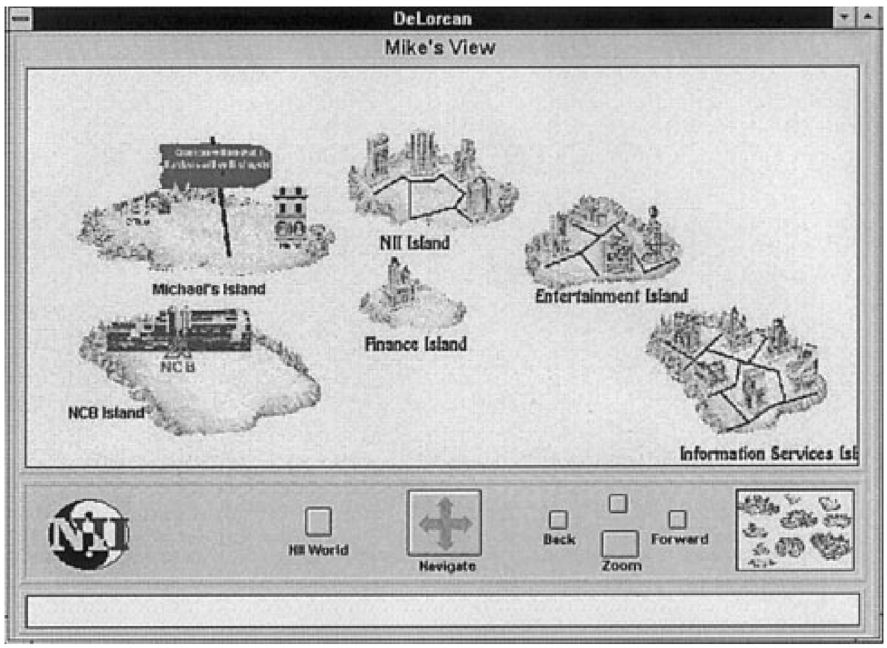


Figure 4.7: Information Islands (Waterworth, 1996)

Figure 4.7 shows the original manually generated interface comprised of a central view window, a small overview window at the bottom right and navigation buttons.

Based on this concept, Modjeska and Waterworth (2000) developed a VRML[1] version of the Yahoo! entertainment directory shown in Figure 4.8.

[1] Virtual Reality Markup Language

Figure 4.8: Yahoo! Entertainment Directory Islands (Modjeska & Waterworth, 2000)

Comparative user tests between versions of the interface showed a simplistic world comprised of basic geometrical shapes (as shown in Figure 4.8) resulted in higher efficiency and no effect on effectiveness as compared with a more ‘naturalistic’ world (Modjeska & Waterworth, 2000, p. 219). As described in section 3.2.1.3.1 *Level of Realism*, this suggests that a VR world does not need to closely resemble physical reality to produce the same performance.

4.2.2.4 Complete Overview (Row 1, Column D)

The four applications described in this section provide no keyword searching features but display the whole hierarchy structure by providing users with a form of dynamic zooming.

4.2.2.4.1. Information Pyramids

Information Pyramids (Andrews, 2002) is a VR view of a large hierarchy built using pyramid-like structures which grow upwards as the hierarchy deepens (see Figure 4.9).

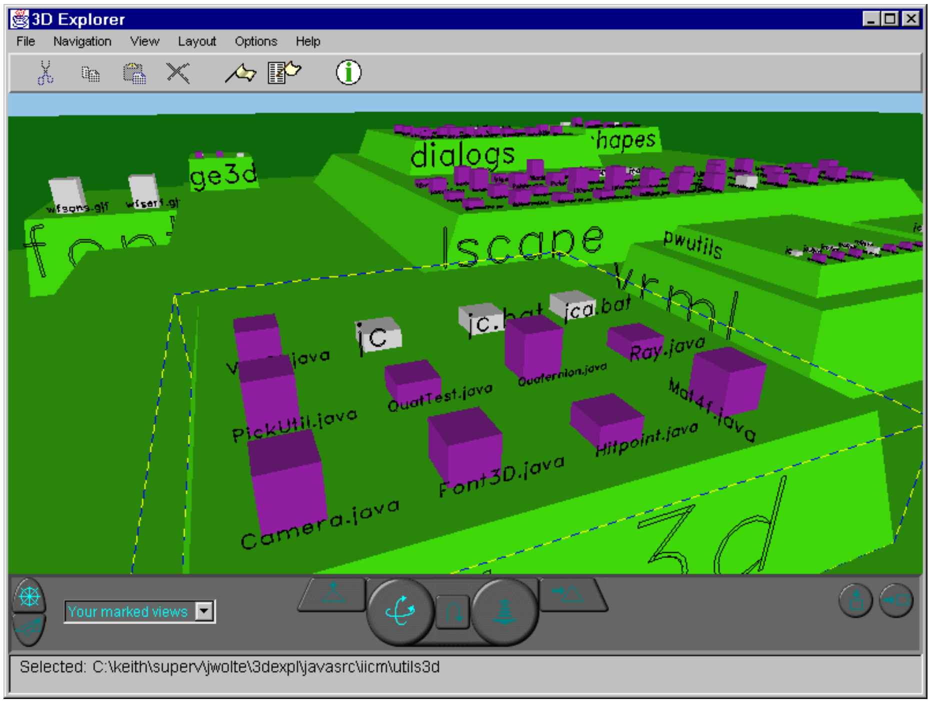


Figure 4.9: Information Pyramids (Andrews, 2002)

A plateau seen in Figure 4.9 represents the root of the hierarchy; smaller plateaus arranged on top of it represent its branches. Separate 3D shapes are used to represent documents. Fly-through navigational facilities are provided for interactive exploration. This application is unique and significant to this research because it utilizes 3D containment to suggest parent-child relationships within a large hierarchy.

4.2.2.4.2. Treemap

Treemap (Johnson & Shneiderman, 1991) is a popular compact *space-filling* algorithm for the display of pre-existing large hierarchies. Each root node is drawn as a square sized proportional to the number of items the branch contains. Sub-nodes are then recursively drawn as smaller squares within its parent node square (see Figure 4.10). The term 'space-filling' describes the way Treemap progressively fills all screen space allotted.

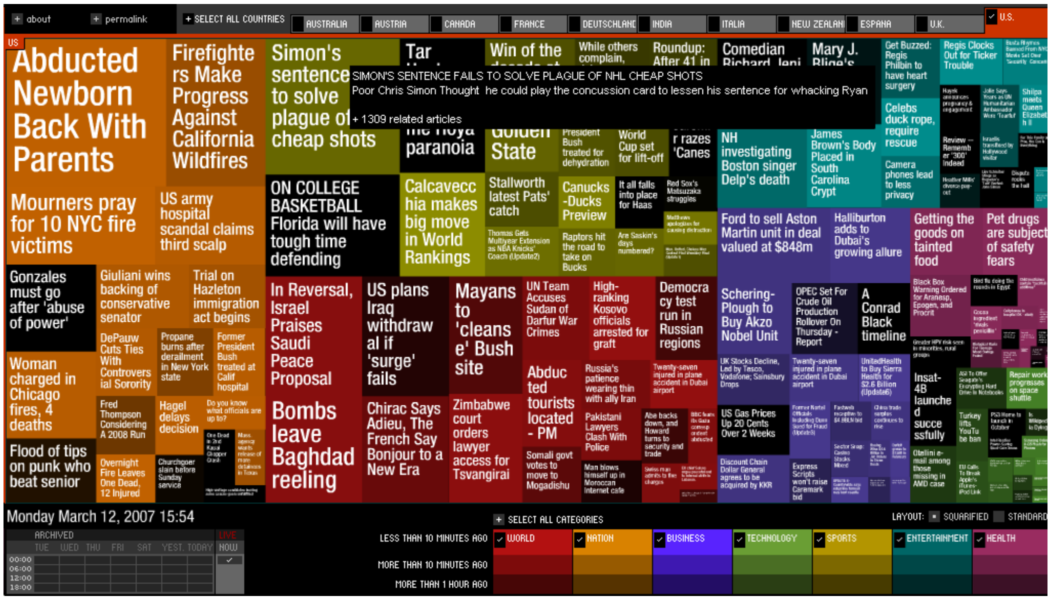


Figure 4.10:Treemap from Maramushi Google NewsMap[1]

Figure 4.10 shows a Treemap of the Google News information service[1]. Class of news story (e.g., World, Nation, Business, etc.) and geographical source can be used as filters using the color legend (bottom right) and the source tabs (top right). The size of each square is proportional to the number of articles on the subject which provides an efficient indication of major news stories.

The visual complexity of the Treemap display has caused issues with users. Experimental studies "have demonstrated benefits but acknowledged that Treemap required training before users could use it effectively" (Plaisant, 2004, p. 113). Barlow & Neville (2001) conducted a controlled comparative experiment of four hierarchy visualization strategies and report that "Treemap was uniformly disliked" (p. 130)

Space filling techniques are said to be particularly appropriate "when users care mostly about leaf nodes and their attributes (e.g. outlier stocks)" (Plaisant, *et al.*, 2003, p. 287). One important drawback is that the hierarchical structure is hard to discern because all space is used for the display of items and the structure is suggested implicitly via containment. Beamtrees (van Ham & van Wijk, 2003) is an attempt to address this known drawback using a simulated 3rd dimension to convey parent-child relations. Kobsa

[1] Available online at http://newsmap.jp/

(2004) restates the importance of specific and browse search integration by suggesting Treemaps could benefit from keyword search features and result highlighting.

Treemap is pertinent for this research since this popular hierarchy visualization algorithm may be part of pre-existing user knowledge. Empirical knowledge of Treemap also suggest that such a densely populated screen may create a negative first impression of the application – less might indeed be more.

4.2.2.4.3. Cone Trees

In the most widely cited publication in the domain of IV (Chen, 2004, p. 4), Robertson *et al.* (1991) describe the 3D Cone tree visualization of a pre-existing hierarchy (see Figure 4.11).

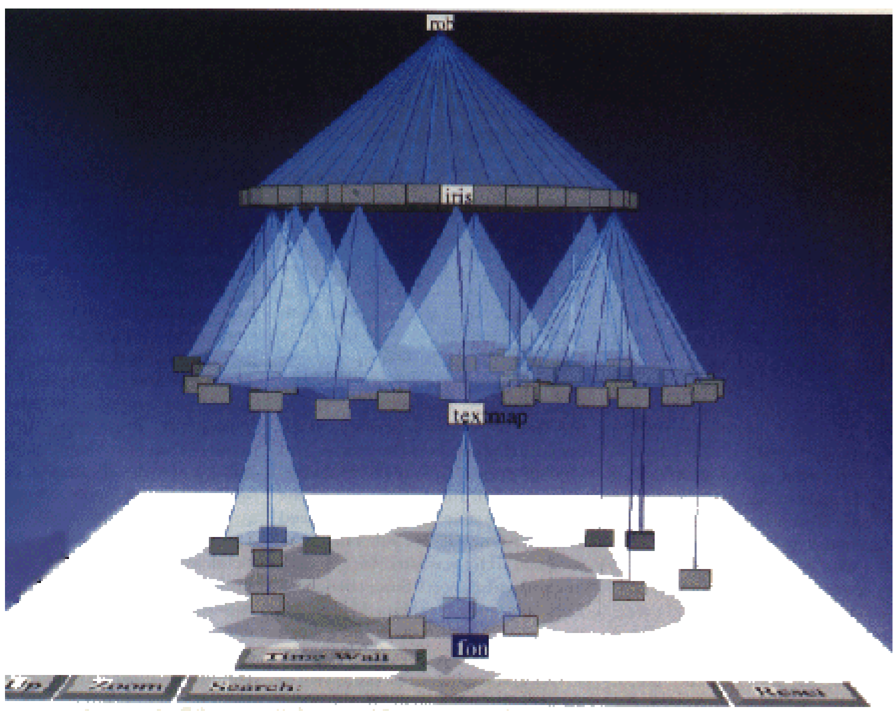


Figure 4.11: Cone Tree (Robertson, et al., 1991)

The visualization is built from the top of the hierarchy placed near the top of the screen, "and is the apex of a cone with its children placed evenly spaced along its base. The next layer of nodes is drawn below the first, with their children in cones" (Robertson, *et al.*, 1991, p. 190). In this manner a complete overview of the hierarchy is provided which users can explore by rotation “to bring interesting parts to the front and pruning to remove non-relevant information” (Andrews, 1995, p. 98).

Cone trees is highly pertinent to this research for its historical significance and its use of interactive 3D to explicitly represent the structure of a large hierarchy. One of the main problems with Cone trees is that users must navigate around the structure in order to see the nodes which are hidden or occluded (Teyseyre & Campo, 2009, p. 93). The original Cone trees uses a "cascading rotation of the 3D cones to bring the desired child to the front" (p. 95). This makes the technique well suited for overall structure comprehension but requires efficient navigation controls for specific node finding.

4.2.2.4.4. File System Navigator

The File System Navigator (FSN, or “Fusion”), written by Joel Tesler and Steve Strasnick at Silicon Graphics, visualises a Unix file system as an information landscape (see figure Figure 4.12).

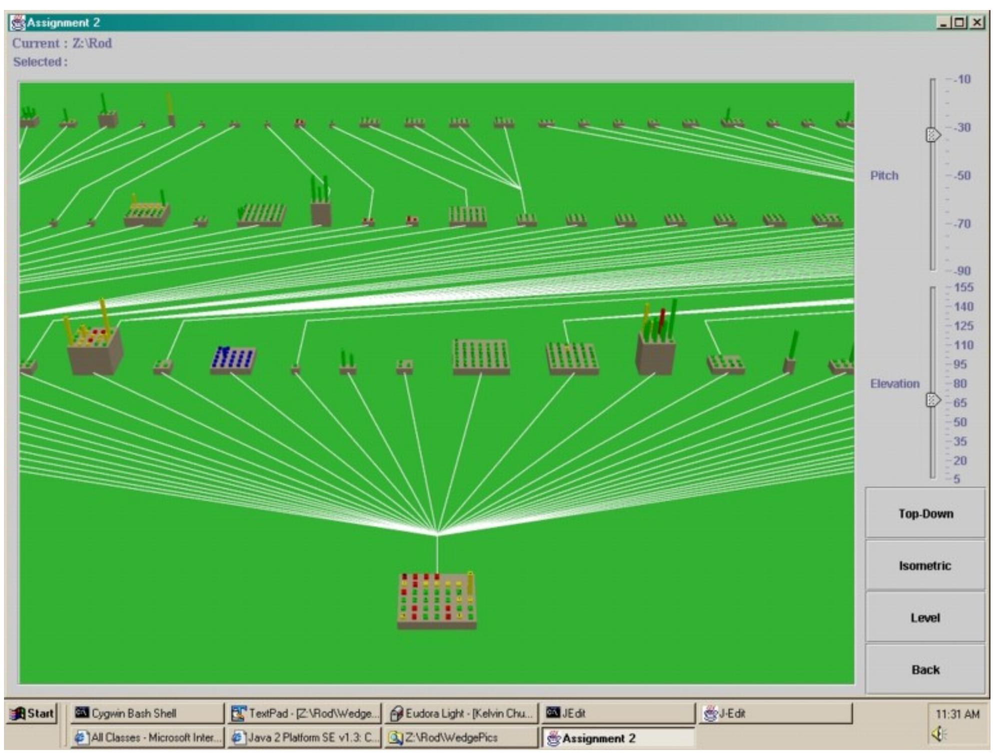


Figure 4.12: File System Navigator (Tesler & Strasnick, 1992)

Figure 4.12 shows directories represented by blocks laid out on a plane, their height representing the cumulative size of the contained files. Smaller blocks atop the directories represent files in the directory (their memory size mapped to their height). FSN provides navigation controls where

> *“users can “fly” over the landscape, taking it in as a whole, or swoop down to a specific directory. Clicking on the arc to a subdirectory results in an invisible hand grabbing you and leading you through space to that subdirectory. Clicking on a file block brings a virtual spotlight to bear on that block, double-clicking opens the file for editing, etc” (Andrews, 1995, p. 98)*

This application is pertinent for its historical significance as the first to offer a 3D fly-through visualization of a large hierarchy. The application was shown in the 1993 film “Jurassic Parc”[1] and has inspired other 3D prototypes such as the Harmony Internet Browser (Andrews, 1995).

4.2.3 User Keywords Limit Corpus (Row 2)

Applications in this section require user keyword(s) to constrain the corpus to visualize. This is also called *term seeding* (Buzydlowski*, et al.*, 2002) and these tools can be described as query result visualization applications. This produces a "dual-mode interface" (Smith*, et al.*, 2006, p. 798) and implies the user must recall relevant keyword(s) before the IR interaction can begin.

4.2.3.1 Single Level Hierarchy (Row 2, Column A)

AquaBrowser (Veling, 1997), an OPAC discovery interface, is part of the current wave of next generation OPACs (see section 3.1.5.1) commercialized by Medialab Solutions[2] from the Netherlands (see Figure 4.13). Using a bibliographic database, the application generates an interactive, single level hierarchy visualized as a network plot or classical undirected graph (Battista, et al., 1994).

[1] “This is Unix, I know this!”
[2] Company Web site and customer list available at http://www.medialab.nl/index.asp

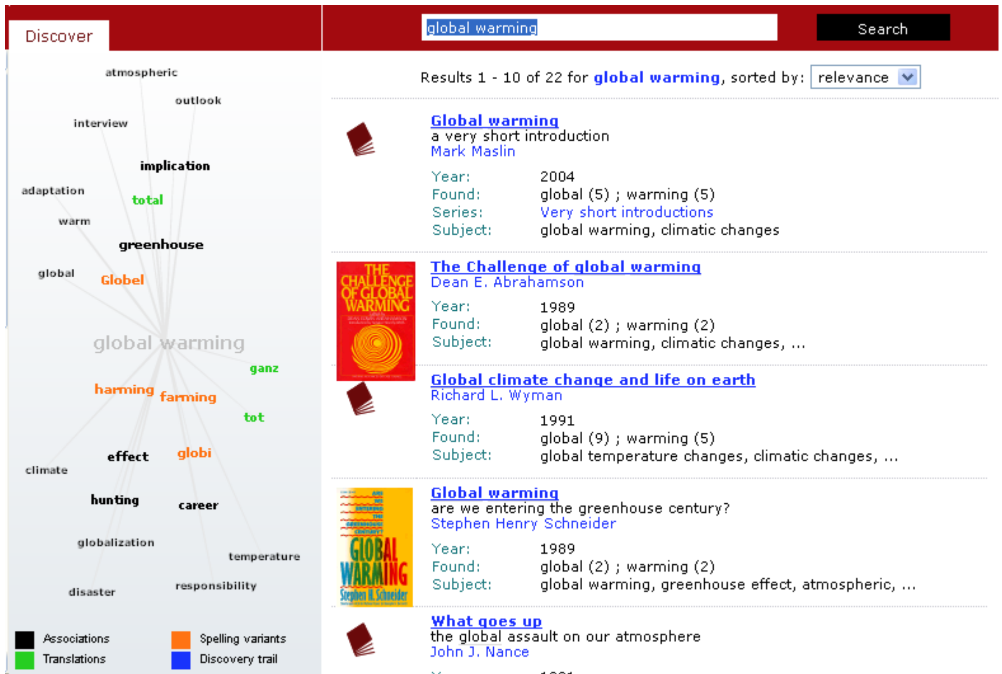


Figure 4.13: AquaBrowser (Veling, 1997)

The visualization shown on the left pane of Figure 4.13 is called a *word constellation* and comprises linked terms centered on the user's keyword(s) which can be interactively browsed for related concepts. Clicking on a word label within the constellation automatically updates the result list (right side pane of Figure 4.13) and redraws the word constellation centered on the selected term. The visualization is highly interactive and suggests valid related concepts for further browsing and searching.

The interface does suffer from limitations associated with undirected network visualizations. A complete overview is not available and the visualization is not bounded such that users can get 'lost' endlessly clicking from one term to another. As well, the relative placement of each term is arbitrary.

Since this application is widely available on existing OPACs it is important to consider it as a browsing solution part of searcher pre-existing knowledge and expectations.

4.2.3.2 One-by-One Multi-Level (Row 2, Column B)

Applications in this section allow users to navigate multi-level subject hierarchies strictly one level at a time. To some extent they also suffer from the limitations described for ET-MAP (see section 4.2.2.2). Users have to remember the explored branches and the future branches to investigate. Both applications in this section are publicly available and are considered existing browsing tools which may be part of searcher pre-existing knowledge and expectations.

4.2.3.2.1. Kartoo

Kartoo[1] is a highly interactive clustering and Web search result visualization (see Figure 4.14).

Figure 4.14: Kartoo.com

The central pane of Figure 4.14 offers a spatial visualization (see section 3.2.6) of the result space populated by irregularly shaped labelled groups (e.g., “collections”, evaluation” and “presence” in Figure 4.14) and their highly representative Web sites (i.e., the page icons in Figure 4.14). Clicking a group label removes the current map and generates the map of items contained within the selected group providing a drill-down type interaction.

The application is aesthetically pleasing, interactive and responsive. It suffers from labelling issues associated with automatic clustering mentioned in section 3.2.5

[1] Available online at http://www.kartoo.com

Hierarchy Visualization (i.e., various levels of abstraction, non-descript labels) and placement of the icons on the map is arbitrary.

4.2.3.2.2. WebBrain

WebBrain[1] is a single level node-link type visualization of the Open Directory[2] hierarchy (see Figure 4.15).

Figure 4.15: WebBrain.com

The visualization initially presents the top level of the subject hierarchy in a quasi outline view and searchers can enter keyword(s) to visualize related terms within the structure. For example, Figure 4.15 shows a search for the term "visualization" (see bottom entry box to the left of "GO" button); as a result, the visible structure is redrawn to show subjects related to the term along with a list of associated Web sites (not shown). Clicking on a textual label centers the visualization on that term surrounded by its related subjects and updates the list of Web sites.

WebBrain capitalizes on the human generated Open Directory hierarchy which provides better abstraction control and higher quality labelling. The visualization suggests valid terms for further browsing and searching; however, the structure of the hierarchy is not apparent and users may assume they are navigating a flat network plot or undirected graph. Additionally, WebBrain does not provide information about the searchers' current location within the hierarchy.

---

[1] Available online at http://www.webbrain.com
[2] Available online at http://www.dmoz.org/

4.2.3.3 Hierarchy Parts (Row 2, Column C)

Applications in the following section provide visualizations of multiple levels of a subject hierarchy within a single display. In general, from a user specified entry point, the visualization presents no more than 2-3 hierarchy levels.

4.2.3.3.1. Dewey Gui

Dewey GUI (Allen, 1995) presents an outline view of the Dewey library classification system. The application positions the user-specified search term(s) within the local region of the Dewey hierarchy. Like WebToc (see section 4.2.5.2.2) this application inherits the limitations of 2D hierarchy displays (Shneiderman, et al., 2000, p. 58): inefficient usage of screen space. Dewey Gui is historically significant since it is the first to explore ways to visually represent a library collection subject hierarchy.

4.2.3.3.2. Grokker

Grokker[1] was a publicly available clustering and hierarchy visualization application produced by Groxis, Inc. (see Figure 4.16).

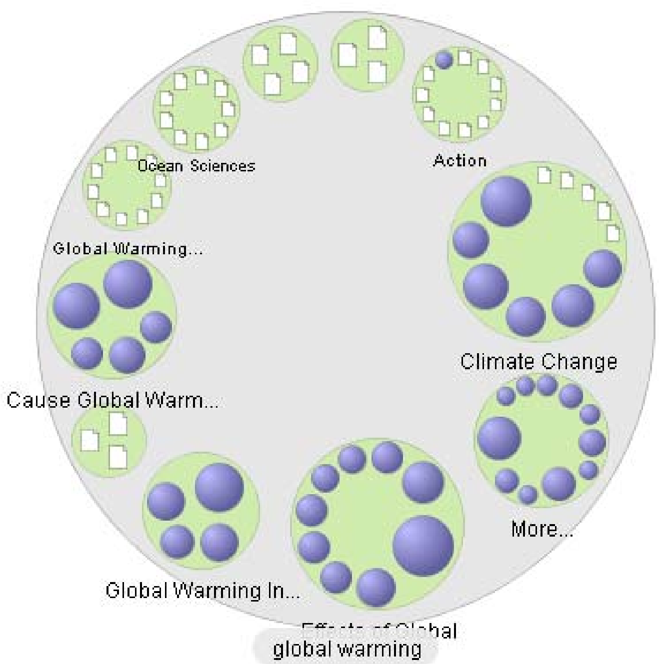


Figure 4.16: Grokker.com

[1] Available online at http://www.grokker.com

After receiving user keyword(s) the application generates clusters and sub-clusters represented by recursively embedded circles and individual Web sites as paper page icons. For example, Figure 4.16 shows the 'Grok' for the term "global warming" (bottom label) contains 11 more specific subject groups such as "Climate Change" and "Action" (middle right side circle and above). The blue circles contained within the "Climate Change" group signify deeper hierarchy levels while the page icons represent individual Web sites the user can choose to inspect. The application initially presents 2-3 hierarchy levels but clicking in one sub-category smoothly animates the view towards a more focused and refined selection of the hierarchy.

Grokker is said to be "impressive…scalable and fun to use" (Chen, 2004, p. 139) but some design problems were reported by Rivadeneira *et al.* (2003). In a controlled experiment comparing Grokker, Grokker Text and Vivisimo's outline view clustering[1] the authors report no significant differences in terms of accuracy and efficiency but participants preferred Vivisimo. Issues with Grokker included lack of information concerning previously visited nodes and the meaning of circle colors.

Grokker is pertinent to this research as it is a clear utilization of Gestalt principles (see section 3.2.1 *Theoretical Foundations*) of closure, relative size and spatial concentration. Unfortunately it suggests misleading similarity by spatial proximity since the relative position of the groups and Web sites is arbitrary. Grokker is also an effective utilization of animation to indicate movement through the hierarchy.

4.2.3.3.3. MeSHBrowse

MeSHBrowse (Korn & Shneiderman, 1995) is a 2D hierarchy representation which allows users to interactively browse a subset of the MeSH hierarchy. The application requires a user specified starting point and does not provide content searching capabilities.

[1] Available online at http://search.vivisimo.com/

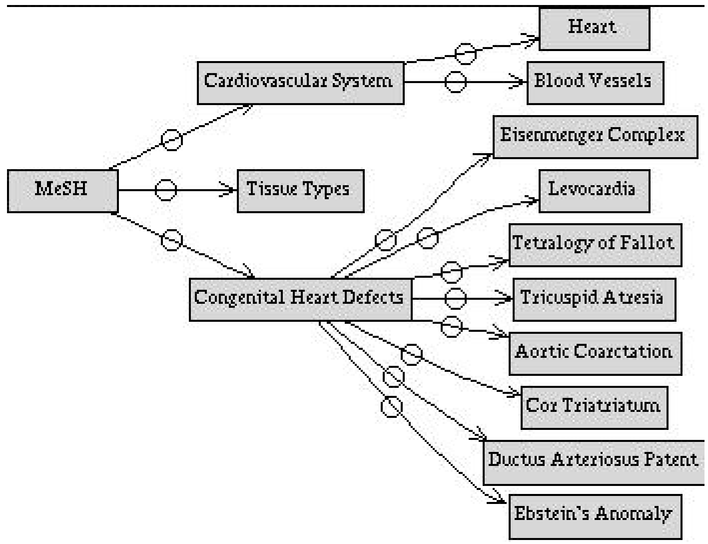


Figure 4.17: MeSHBrowse (Korn & Shneiderman, 1995)

MeSHBrowse is described here because it is one of few visualizations based on a library subject hierarchy. The prototype suffers from inefficient screen usage associated with 2D outline views.

4.2.3.3.4. Hieraxes

Hieraxes (Shneiderman, et al., 2000) was designed specifically for digital libraries and their hierarchically organized collections. The interface organizes the results of a user query in a customizable grid. At each grid point of the display a cluster of color coded dots or a bar chart represents the matching documents (see Figure 4.18).

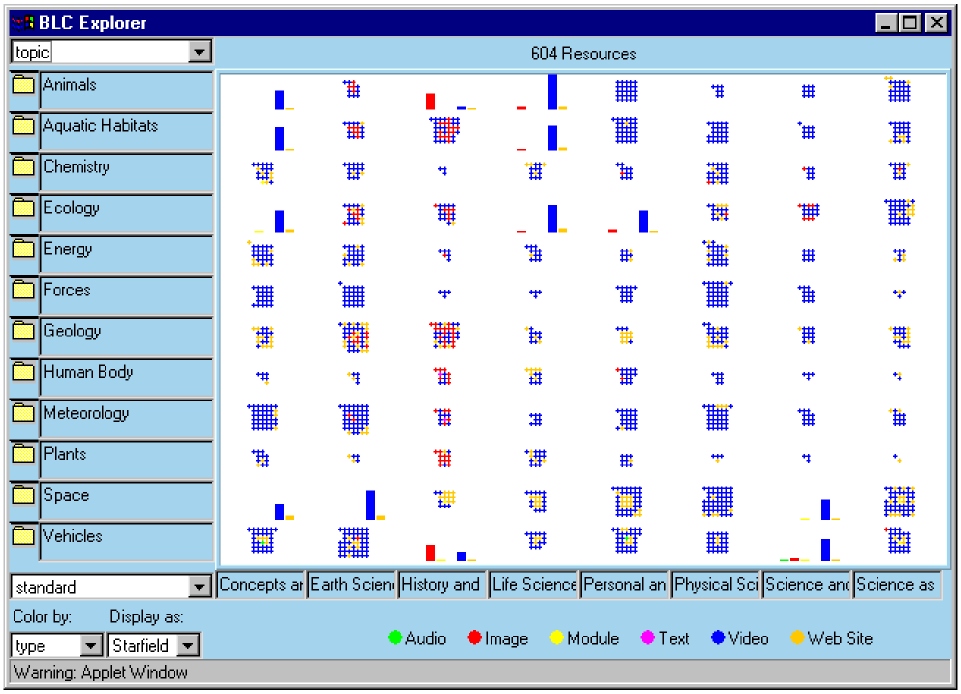


Figure 4.18: Hieraxes (Shneiderman, et al., 2000)

Usability tests (Shneiderman, et al., 2000) show users require considerable training to understand the information provided and the high level of user control over the display. Hieraxes is included in this list because it is a unique approach to the manipulation of a library information collection organized by subject. The approach also suggests there is an upper limit to the visual complexity imposed on novice users.

4.2.4 Search in Display (Row 3)

Visualizations in this section provide searching capabilities within the structure but not the individual information items. A keyword 'match' in this case strictly signifies the query is used as a term or part of a term within the subject hierarchy. This type of searching is problematic since the very nature of controlled vocabulary used in semantic hierarchies implies a highly restrictive and often specialized vocabulary (see section 3.1.4 *Subject Indexing*); in other words, searchers may have difficulty entering keywords that are part of the subject labels.

4.2.4.1 Hierarchy Parts (Row 3, Column C)

Hyperbolic space (Lamping, Rao & Pirolli, 1995) is a projection of a hierarchy onto a non-Euclidean hyperbole space which distorts the size of nodes further from the chosen central and non-distorted node. The advantage it provides stems from the fact that "hyperbolic space is infinite but can be projected into a finite ball of Euclidean space" (Risden, et al., 2000, p. 697). This provides information about depth and shape of branches which would otherwise grow beyond the screen area (see Figure 4.19), and "because more of the tree structure is accessible on the display, the Hyperbolic browser is expected to accelerate users' browsing performance over conventional tree browsers" (Pirolli, *et al.*, 2000, p. 161).

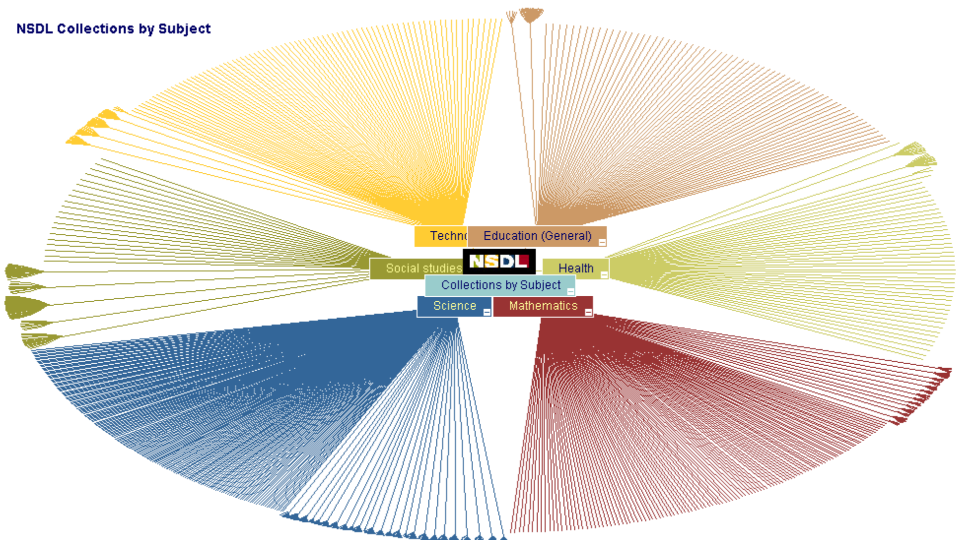


Figure 4.19: Hyperbolic Space (Lamping, et al., 1995)

A laboratory experiment (Lamping, et al., 1995) compared using the hyperbolic browser with a conventional outline view layout. Although users preferred the hyperbolic visualization, there was no performance advantage on the task of finding specific node locations. Kobsa (2004) reports that Hyperbolic view performance was average on every task.

This visualization is significant to this research because it is commercially available from Inxight Software[1] and has been implemented by LexisNexis[2] to represent its collections and their relations. Additionally, it is a well known example of context-plus-focus presentation using distortion (see section 3.2.4.1 *IV Techniques for Large Dataset Navigation*).

[1] http://www.inxight.com/
[2]http:// http://www.lexisnexis.com/startree/startree.asp

4.2.4.2 Complete Overview (Row 3, Column D)

The following applications offer a complete overview of the hierarchy either using distortion or powerful zooming and view point displacement techniques.

4.2.4.2.1. Harmony Internet Browser

Harmony Internet Browser (Andrews, 1995) uses FSN (Tesler & Strasnick, 1992) and VR to provide a visualization of a hierarchically organized Web collection. The author states that searching can be performed but it is not clear whether this is matched in the structure, the content or both. This application is mentioned because it is a utilization of VR for visualization and searching of hierarchies.

4.2.4.2.2. Cat-a-cone

Based on a mix of Cone trees (see section 4.2.2.4.3) and the WebBook (Card & Robertson, 1996), Cat-A-Cone's designers (Hearst & Karadi, 1997) expected that "the context preserving display of ancestor and sibling information provided by (Cone trees) can help the user see the general meaning of a term" (p. 251). Users of their application (see Figure 4.20)

> *"can view different levels of description simultaneously, so more familiar concepts can be viewed in more detail, and less familiar at a more general level. Category labels are disambiguated by their ancestor/descendant/sibling representation." (p. 253).*

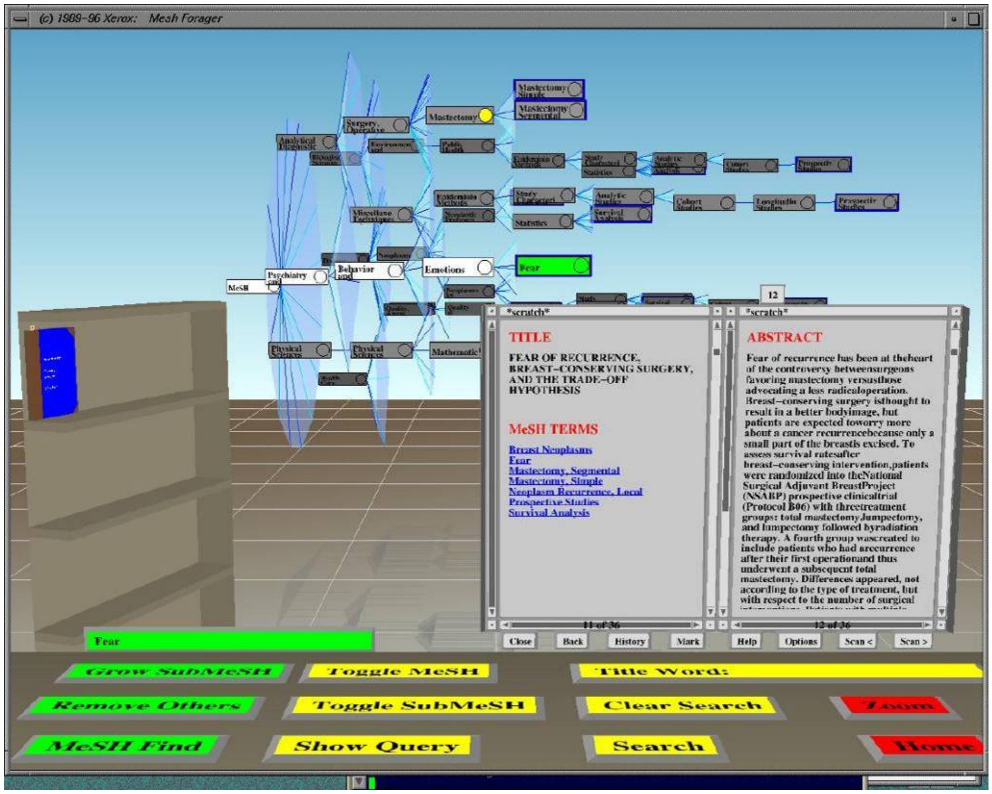


Figure 4.20: Cat-a-cone (Hearst & Karadi, 1997)

Users of Cat-a-Cone first specify a search parameter (e.g., title, keyword) and each matching document is represented as a book placed on a shelf (left side of Figure 4.20). Selecting a book displays its bibliographic information (right side of Figure 4.20) and its position(s) within the subject hierarchy is (are) illuminated in the Cone Tree shown in the background.

Cat-a-Cone is highly pertinent to this research because it is the only visualization of a library controlled subject hierarchy which integrates searching and subject browsing. One of its drawbacks stems from its use of unmodified Cone Tree which does not provide information about the relative importance of branches for the current collection. As well, the user must specify search terms and cannot manipulate the tree directly.

4.2.4.2.3. 3D Hyperbolic Space

3D Hyperbolic Space (Munzner, 1997) is a 3D extension of the 2D hyperbolic distortion technique (see section 4.2.4.1) which provides a distorted view of the complete hierarchy (see Figure 4.21).

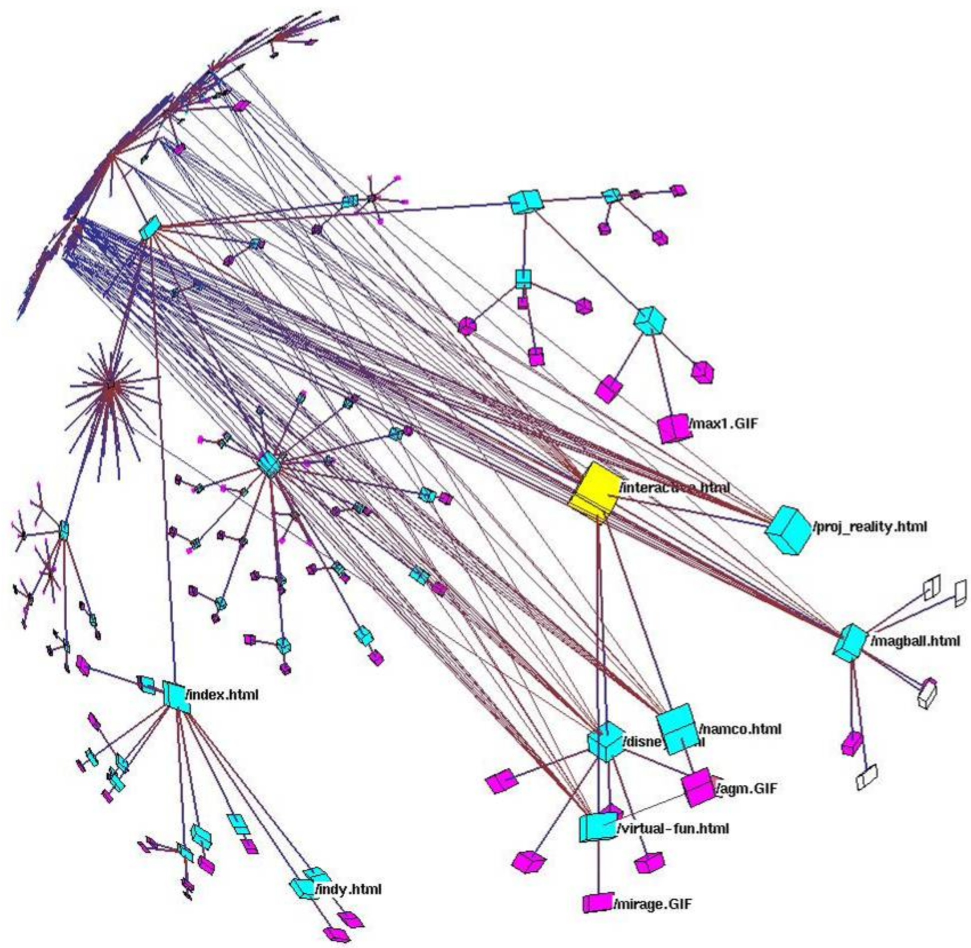


Figure 4.21: 3D Hyperbolic Space (Munzner, 1997)

This impressive visualization technique is pertinent to this research because it makes use of the third dimension and illustrates related issues of visual occlusion. Katifori *et al.* (2007) suggest weaknesses of this 3D technique concern the dense visual display which "provides only part of the sphere, labels are not visible away from the center and sometimes the animation may be disorienting" (p. 30).

4.2.5 Integrated Search in Data + Structure (Row 4)

Finally, applications described in this section offer fully integrated searching through content and visual structure. This is where the SE-3D hopes to be situated and this makes the following application highly pertinent.

4.2.5.1 Single Level Hierarchy (Row 4, Column A)

Bead (Chalmers, 1993) is the first description of an interactive landscape corpus visualizations where user can ‘fly’ through a text document corpus and inspect individual bibliographic records within a 3D world. Individual document placement is based on semantic similarity. Bead also allows the user to specify a keyword search and ‘highlights’ matching documents in the display (see Figure 4.22).

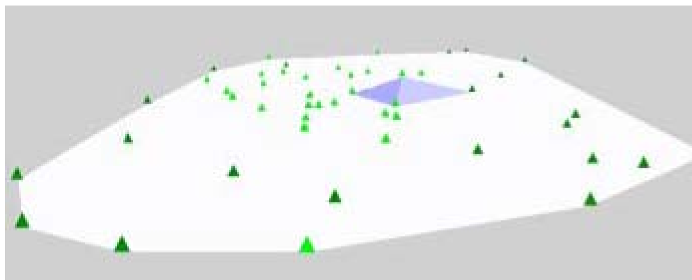


Figure 4.22: BEAD (Chalmers, 1993)

Figure 4.22 represents individual documents as small green pyramids and a large central cluster of documents as a large blue pyramid. Documents matching a query are marked in light green and cluster tightly around the center cluster. Bead provides only a single level hierarchy, navigation in large collections can become difficult due to visual occlusion, and the lack of textual labels forces the searchers to inspect each item individually.

Beyond its historical significance, Bead is an early example of 3D fly-through navigation which integrates keyword searching and matching document highlighting.

4.2.5.2 Hierarchy Parts (Row 4, Column C)

Applications in this section display parts of hierarchies using innovative icons to represent whole branches or using an enriched outline view. Both of these applications are pertinent to this research because they address shortcomings of classical 2D hierarchy visualizations and provide fully integrated search capabilities.

4.2.5.2.1. SpaceTree

SpaceTree (Plaisant*, et al.*, 2002) aims to enhance the top-down node-link hierarchy view classically used to represent organizational charts by describing the topology of hidden branches through a stable and consistent layout (see Figure 4.23).

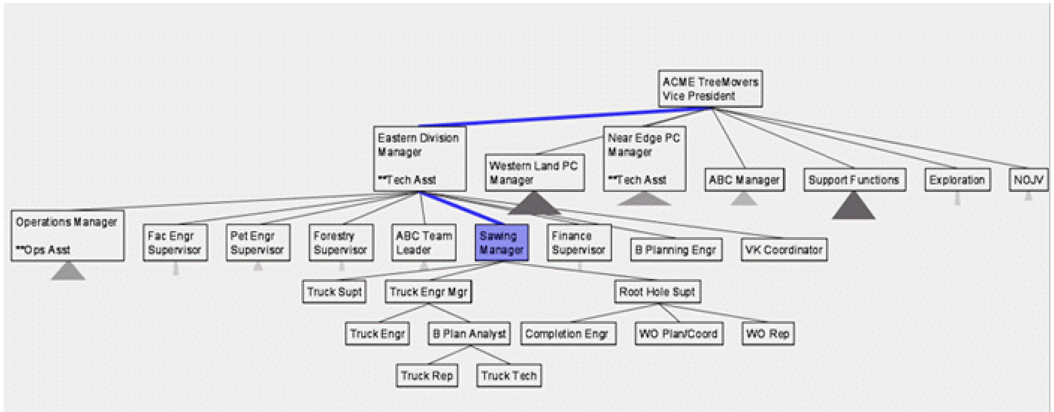


Figure 4.23: SpaceTree (Plaisant, et al., 2002)

This is visually communicated using triangles of variable width and height mapped respectively to the relative breadth and depth of the branch. Darker shades correspond to branches with more items. Plaisant *et al.* (2002) performed a controlled experiment which showed benefits for navigation to previously visited nodes and estimation of tree typology. The 3×7 (3 interfaces × 7 tasks) repeated measures within subject design produced a wide variability on measures of speed and preference leading to only a limited number of statistically significant effects. The authors state the main study value stemmed from observations during the experiments that suggested the three attribute coding of the triangles (i.e., depth, breadth, density) was often misinterpreted and had to be explicitly clarified.

4.2.5.2.2. WebTOC

WebTOC (Nation, Plaisant, Marchionini & Komladi, 1997) provides an outline view representation of a Web site (see Figure 4.24[1]). The application provides keyword searching and a rich set of information about the content of the web site (e.g., predominant file types, number or size of items).

[1] Available online at http://www.cs.umd.edu/projects/hcil/webtoc/fhcil.html

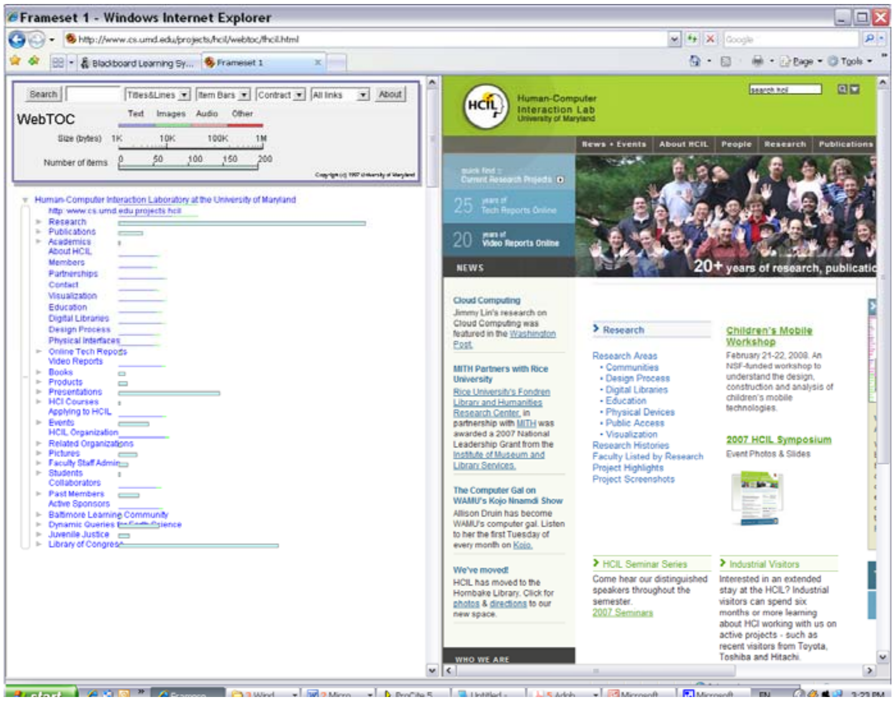


Figure 4.24: WebTOC (Nation, et al., 1997)

WebTOC inherits issues related to 2D node-link hierarchy representations (see section Chapter 2: *Thesis*); for example, the fully expanded tree cannot fit on a single screen and the collapsed display provides little information about the overall structure of the hierarchy.

4.2.5.3 Complete Overview (Row 4, Column D)

The following two applications are most closely comparable to SE-3D since they offer a complete visual overview of a hierarchical organization by subject and fully integrated searching features.

4.2.5.3.1. VxInsight

VxInsight (Davidson, et al., 1998) is a commercial spatial visualization application inspired by SPIRE (see section 4.2.2.1.3) which supports browsing and searching through an automatically clustered database (see Figure 4.25).

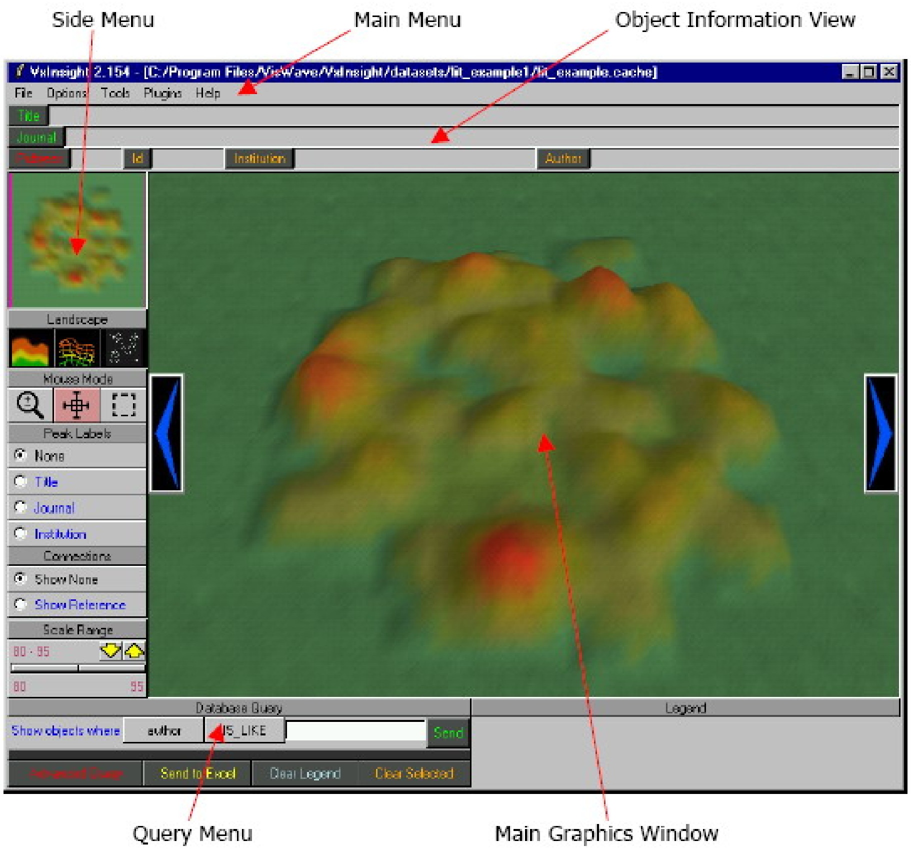


Figure 4.25: VxInsight (Davidson, et al., 1998)

The application is unique because, unlike SPIRE (see section 4.2.2.1.3), it provides the "ability to zoom into the mountains to see more structure" (Davidson, et al., 1998, p. 263). This can be described as progressively traveling from the general view shown in Figure 4.25 towards a more focused search where the individual documents are visually represented. The application also provides full searching capabilities with matching documents highlighted on the map.

VxInsight is an impressive visual IR application which capitalizes on existing user knowledge of interaction with geographical information systems such as Google Earth[1]. Its drawbacks are those associated with clustering algorithms (see section 3.2.5) and lack of an explicit representation of a semantic hierarchy. The hierarchy is implicitly communicated by the height of the mountains which suggests denser branches, and there

[1] Available from http://earth.google.com/

is no way to inspect broad to narrow subjects without zooming into every area of interest and inspecting some of its contents.

4.2.5.3.2. SWAPit

SWAPit (Seeling & Becks, 2004) is a tight integration of a query box with relevance ranking, a document list, an outline view of a subject hierarchy and a 2D spatial representation (see Figure 4.26). This synchronized multi-tool approach is a response to the general finding that

> *"there is not one specific method that seems to be the most appropriate for all applications and, consequently, a viable solution would be to provide the user with several visualizations, so as to be able to choose the one that is the most appropriate for his/her current needs" (Akrivi, et al., 2007, p. 39)*

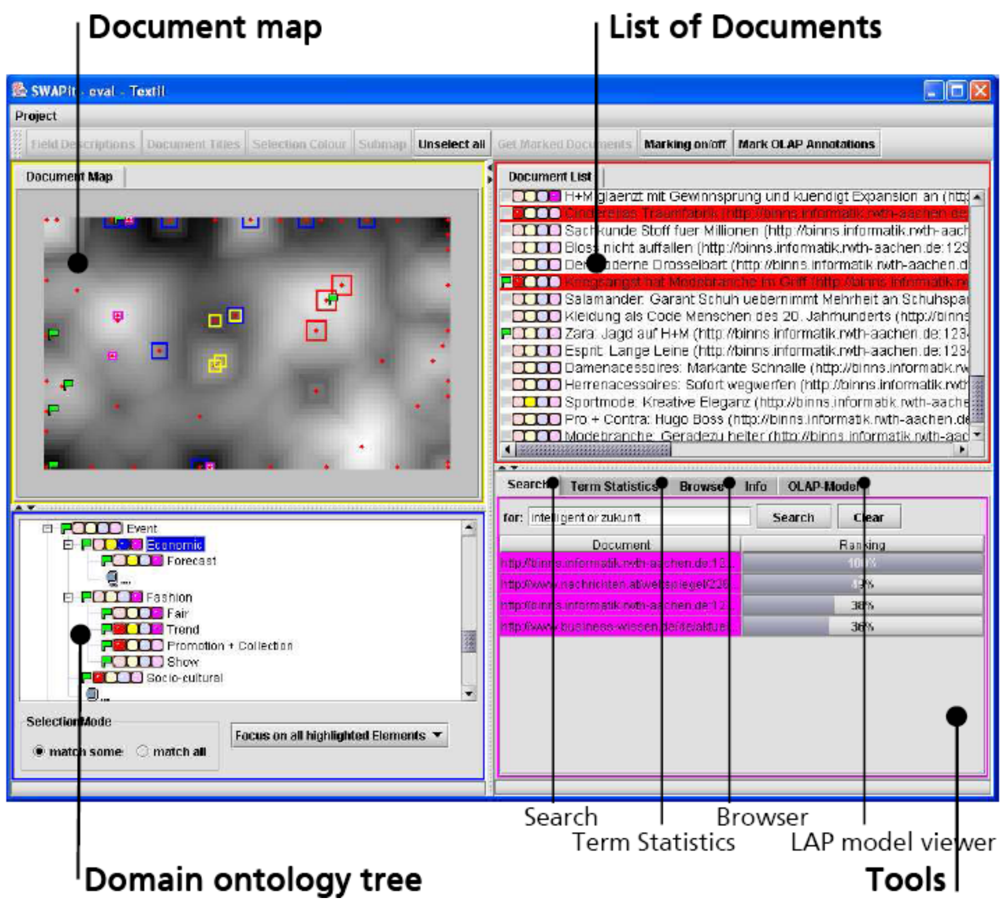


Figure 4.26: Multiple Views System (Seeling & Becks, 2004)

SWAPit shows a keyword-search tab (bottom right) whose results are displayed as a ranked list (top right) and highlighted in the map display (top left). The domain semantic hierarchy (bottom left) is also highlighted to reveal where matching documents are situated in the structure. Selecting an item in any of the views automatically updates the others.

Seeling and Becks (2004) performed an expert-based evaluation of their multiple views SWAPit system using two business intelligence professionals. Informal findings suggest the interface could be successfully used but participants "would have liked a possibility to scale down the number of views and the functionality for specific tasks" (Seeling & Becks, 2004, p. 69). This finding is a likely result of the flexibility vs. usability tradeoff (Norman, 1998) where satisfying a larger set of requirements (e.g., many kinds of visualization tools) necessarily means more compromises and complexity.

4.3 Subject Explorer in 3D (SE-3D)

This research aims to develop and test SE-3D which can be conceptually described as a mix of Conetrees (see 4.2.2.4.3) where each node (small grey/white rectangles in Figure 4.11) is replaced by a simplistic spatial visualization (see section 3.2.6 *Spatial Displays*). Users navigate around a static structure in space while inspecting floating textual labels of LCSH subjects offered by a large organized collection. Keyword searching and ranked lists of results are always available (see section 4.1.1). Tight integration of searching allows filtering of the visible parts of the structure and automated movement to a subject location when the user clicks a hyperlinked subject in the result list. The objective is to produce a semantic hierarchy explorer allowing equivalent or better performance but preferred by users as compared to a purely textual hyperlinked ranked result list and semantic hierarchy navigation.

4.3.1 Simplifying the LCSH Hierarchy

SE3-D hopes to capitalize on the immersive potential of an appealing VR application (see section 3.2.7 *Virtual Reality for IV*) and its suitability to represent hierarchies (see section Chapter 2: *Thesis*). 3D Cone Trees is one of few visualization techniques that have been systematically integrated into designs of information systems (Chen, 2004, p. 1). For example, they are used in LyberWorld (Hemmje, et al., 1994), Cat-a-Cone (Hearst & Karadi, 1997), and VizNet (Fairchild, Serra, Hem, Beng Hai & Tin Leong, 1993). Cockburn and McKenzie (2000) tested Cone Trees compared to traditional outline views. They report slower performance on specific data location tasks using Cone trees and a rapid decrease in performance as the hierarchy complexity increases. Qualitative results indicated that

> *the subjects were enthusiastic about the cone tree visualization and that they felt it provided a better 'feel' for the structure of the information space. (p. 425)*

A visualized semantic hierarchy inspired by Cone Trees offers structure at multiple scales which "supports location memory and makes it easier to revisit places that have been looked at only seconds ago" (Ware, 2008, p. 41). For example, the distribution of items within the broadest subjects in the LCSH hierarchy acts as the highest scale overview (see section 4.1.3.6) of the whole collection. As the user travels further down towards more specific subjects the scale and the number of items represented decrease. This effectively allows users to browse for known subjects and hone in on promising subsets of the collection. The user might not recognize all subjects but this can be "alleviated by the ability to type a random keyword into a search box and have the system mark" (Abello, *et al.*, 2006, p. 7) all areas of the structure that contain the keywords.

Studies have begun to suggest the upper complexity limit of the hierarchy which can be adequately visualized with Cone trees. Carriere & Kazman (1995) suggest the technique tends to lose its effectiveness for hierarchies beyond approximately 1000 items. Robertson *et al.* (1991) state the shape of the hierarchy is also important and suggest the technique is constrained to 30 items on a single hierarchy level beyond which Cone Trees becomes cluttered and ineffective.

The LCSH semantic hierarchy is necessarily too large to visually represent; furthermore, this complexity grows rapidly when considering multiple inheritance (i.e., subjects that have multiple broader subjects) which is often supported by "replicating child nodes under all their parents" (Akrivi, *et al.*, 2007, p. 5). This duplication of branches necessarily increases the size and complexity of the semantic hierarchy which prevents a complete visual representation of the structure. There are simply too many visual objects to display for limited computer hardware and human cognitive capacity.

Fortunately the design does not have to accurately represent the complete semantic hierarchy; on the contrary, as stated in section 3.2 *Information Visualization*, the primary aim of IV is to visually reveal *patterns*. Beyond the hierarchical pattern, subject indexing likely follows a Bradford's Law: few subjects contain many items, most subjects are assigned to very few items (see section 3.1.4.2 *Browsing Subject Structures*).

Bradford's Law distribution of subject indexing is a statistically verifiable fact that may very well affect all organized collections and provides a clearly defined procedure to greatly simplify the semantic hierarchy to visually represent. At any level of semantic scale (i.e., whether broad or a very specific subject) the IV application will likely satisfy many users quickly by only showing the few large subject families of information items. Based on Pareto's Principle or the 80/20 rule (Juran & Godfrey, 1999), as few as 20% of subjects may well contain as much as 80% of items at every level of the semantic hierarchy. In essence, an IV application may very well represent over 80% of the collection by visually showing less than 20% of the semantic hierarchy; in other words, simplify to reveal the essential.

4.3.2 Semantic Maps for Each Hierarchy Level

Using the 3rd dimension to communicate broad to narrow subject term relationships allows the use of 2D maps or spatial displays (see section 3.2.6) to represent subjects of common semantic abstraction level or at each depth of the hierarchy. The visual 2D design can suggest the relative importance of subjects through shapes sized proportionally with the number of items assigned to each subject. Each of these 2D maps placed in 3D space provides a set of specific terms belonging to the broader parent subject term, their relative importance, and the few highly predominant subjects that deserve to be further subdivided into their own map of more specific subjects.

2D map type spatial displays are traditionally used to communicate relationships between groups of documents using an arbitrarily chosen measure of similarity (Lin, 1997). Bates (2003) suggests revealing the term co-occurrence pattern which assumes a relationship between subject terms if both are used in the same bibliographic record. The approach is based on the well-established bibliometrics practice of author-co-citation analysis (McCain, 1990; White, 1990; White & McCain, 1998). The user may not explicitly understand the meaning of the co-occurrence relationships yet they receive suggestions as to which subjects <u>will</u> likely be read together since they are covered by the same documents. As these relationships are generated from the actual data the visualization may also have "considerable promise as an aid to browsing in a digital library space" (Buzydlowski, *et al.*, 2002, p. 140).

4.3.3 SE-3D vs. Existing Solutions

In terms of information representation, usage of a third dimension by SE-3D is limited strictly to depth representing broad to narrow subject term relationships; therefore, SE-3D is not prone to negative results associated with precise evaluations of quantities and distances in simulated 3D environments (see section 3.2.7 *Virtual Reality for IV*). Studies of existing 3D applications (see review by Hearst 2009, p. section 10.10) suggest special care should taken to ensure the design provides legible textual labels, and intuitive navigation solutions in a fully 3D environment.

Using VR to represent hierarchies is potentially superior to existing 2D tree representations (e.g., Grokker.com, Windows File Explorer) because it makes more efficient use of screen space by utilizing the third dimension to convey hierarchy depth. As opposed to other 3D solutions (Robertson, Card & Mackinlay, 1993), SE-3D incorporates structure and data (Fluit, et al., 2003; Hearst & Karadi, 1997). This provides an integrated IR environment where the “subject hierarchy not only facilitates browsing but also suggests appropriate query terms for searching” (Zhu & Chen, 2005, p. 160). These term suggestions may mitigate the vocabulary difference problem when users cannot come up with valid keywords (Borner, et al., 2003), and supports those who are not native English speakers (Chen, *et al.*, 1998, p. 598). Cat-a-Cone (Hearst & Karadi,

1997) also aimed to present a library subject hierarchy but does not directly integrate the data into the structure.

SE-3D usage of map type displays is unique as it exploits and facilitates access to existing subject headings assigned by librarians as opposed to computationally demanding automatic clustering and labelling technique (e.g., Davidson, et al., 1998). SE-3D also combines the advantages of semantic maps with keyword searching (Hearst, 1999; Yee, et al., 2003) within a 3D hierarchy representation. Finally, as the review of existing systems showed (see section 4.2), there have not been any prior attempts to develop a virtual reality IR application with fully integrated search and browsing features for a real-world large semantic hierarchy.

4.3.4 SE-3D Support in the Searching Process

Search tasks aim to satisfy an information need by finding a small set of items in a large collection (Marchionini, 1995). The tasks supported by SE-3D are wider in scope than most existing IV applications which usually aim to facilitate navigation within the information organization structure, not the information itself. The anticipated benefits of using SE-3D are now described within the six usage scenarios listed in section 3.1 *Information Retrieval*:

1. <u>Formulate query</u>
   - Valid subject search terms, chosen and maintained by library professionals, are explicitly shown which facilitates vocabulary choice.
2. <u>Identify the composition of search results</u>
   - Items with terms matching a keyword search are highlighted within the visual structure explicitly showing multiple contexts associated with the concept of interest.
   - Branches with concentrations of highlighted matches suggest highly relevant subject areas or knowledge domains.
3. <u>Understand the interrelations between retrieved documents</u>
   - Items in a ranked textual list of results can be selected for travel to the location of the subjects to which they belong.
   - Documents assigned to multiple subjects are duplicated and each copy shows all subjects it belongs to.

4. Refine the search
    - Users can filter the visual structure by entering keyword search terms.
    - Users can progressively restrict searching to more specific subject areas which contain fewer items.
    - Matching documents positioned deeper in the hierarchy suggest more specific subjects and keyword search terms.
    - Visual size of subject areas suggest relative coverage of subject domains by the current collection. This can save the time of the searcher by communicating what the collection can offer and potential information needs it may fulfill.
5. Gain an overview of the coverage
    - Users are initially placed at the top of the semantic hierarchy which provides a quick overview of the main sections or branches in the collection.
    - Floating textual subject labels provide a quick way of assessing collection contents.
    - Visual structure of the subject hierarchy suggests depth and breadth of coverage for each subject. This suggests if the collection is appropriate to the task.
6. Browse
    - At any time, the user can inspect a chronological list of items assigned to a single subject
    - The searcher can fly around, through and up/down the visual structure to inspect the various subjects and their broad to narrow relationships

Developing and testing stable data-driven VR prototype such as SE-3D requires analysis of the database collection and its organization, software development and usability testing. These stages are described in the following chapters.

Chapter 5: CORPUS ANALYSIS

Over 120,000 bibliographic records and their subject organization structure (over 280,000 authority records) from the domain of science and engineering supplied by a large academic library were analysed and transformed. The distribution of subject assignments was statistically analysed and the LCSH hierarchy extracted. Results show data scarcity: a relatively small group of subjects contains most of the collection. These results are compared with those of previous analyses of information organization structures.

5.1 Introduction

There is a consensus that explicit CV subject hierarchies can provide a benefit to the searcher and should be incorporated into online catalogues (see review by Fischer, 2005, p. 74). The alphabetical ordering of CV terms found in many traditional OPACs does not provide the "broader context of a topic and its relationship to similar topics" (Larson, 1992, p. 130). Placing subject terms in a broad to specific structure places each work "in a systematic hierarchy and array of related subjects" (Chan, 1989, p. 530). The sum of the relations between subjects is the *syndetic structure* of the controlled subject vocabulary. The LCSH syndetic structure has known issues described as inconsistencies, non-hierarchical cross-references and the large number of *orphan* terms without broader or narrower subjects (see review by Shubert, 1992, p. 54).

Online browsing using broader and narrower subjects is difficult with the unmodified LCSH structure in part because the hierarchical structures "do not cascade systematically" (see review by Shubert, 1992, pp. 59-60). These LCSH structure issues require that user resources be allocated to extract and clarify the LCSH hierarchy; as a result, the explicit "hierarchical nature [of the subject vocabulary] is largely ignored" (Frank & Paynter, 2004, p. 214) by the searcher.

The subject structure may be useful for information retrieval if the "embedded" (Piggott, 1988, pp. 240-241) structure of LCSH could be made explicit and "more complete " (Cochrane, 2000, p. 85). These relationships may benefit online searchers by providing "query expansion and browsing" (Larson, 1992, p. 130) using related subject terms at different levels of conceptual abstraction found in the hierarchy of CV terms (Cochrane, 1986; Markey, 1984).

5.1.1 Sparse Subject Hierarchies

Liu *et al.* (2005) analyzed the Yahoo! web directory and reported typical problems found in real world collections:

- **Depth**: the subject hierarchy is very deep which can make it seem like an endless network as opposed to a finite set of branches. This can be disorienting for users who cannot perceive the top or bottom of the subject hierarchy.
- **Skewed distribution**: information items are not evenly assigned throughout the subject hierarchy. This is understandable since some subjects tend to receive more attention than others. In practice this means that a few areas of the subject hierarchy contain most of the collection while the rest contains very little.
- **Sparse categories**: many subjects contain very few documents. This is explained by the fact that library cataloging rules specify a document should be classified as precisely as possible even if this makes it the only one of its kind.

A sparse subject hierarchy means a few subject branches contain most of the collection. The probabilities are low that searchers are often interested in subject structure areas which contain few or no items. Arguably, "trees with...too few nodes (implying overly specific concepts) are less useful for providing semantic information" (Yi & Chan, 2010, p. 686). Information retrieval may be facilitated by lessening the importance of these empty or near empty subjects—a form of semantic hierarchy noise-reduction.

Resolving these subject hierarchy item distribution issues may be critical. Based on theoretical analysis and empirical evidence, Yang, Zhang & Kisiel (2003) showed that the shape or *topology* of the subject hierarchy dictates the scalability of an automatic text classifier (TC) method. An adapted version of a subject hierarchy was done by Wang & Lee (2007) for their reconstruction of the DDC hierarchy. They assumed that "an automatic (TC) process should follow a path different from a human classifier to traverse a category hierarchy" (p. 137). They believed that "the topology and structure of taxonomies needs to be transformed. This is as yet an unexplored research area." (p. 140).

This study is a subject hierarchy reconstruction (Wang *et al.* 2007) made necessary by the fact that subject ontologies are used by different types of users for different types of tasks. Searchers are interested in collection coverage of subjects. Human content classifiers require relevant and specific subject term(s) for a given work. These are not the same types of tasks and they may require different kinds of subject hierarchies. Unmodified LCSH is a broad and sometimes deep hierarchy designed for classifiers. Information seekers may benefit from using a simplified subject hierarchy showing the few highly populated subjects which contain most of the collection.

5.1.2 Integrating Structure and Information Collection

In the library world the separation of the *pure* information organization structure from its multiple implementations in information collections is important (see Chan, 2000, p. 167). It involves two independent processes: 1) maintenance of the official LCSH list performed by LC, and 2) building individual subject term strings performed by each library based on local needs. This separation may not always be advantageous for searchers.

A subject heading list independent from its collection creates the problem of "searchers being led down a blind alley, when in fact no entries are posted to the related terms" (Shubert, 1992, p. 61). The assumption is that a searcher might gain from the knowledge that a subject exists but the current collection does not provide information for it. This might be useful for some but this information should be clearly communicated "with the number of postings in the particular library given for each heading" (Shubert, 1992, p. 61). This benefit to the searcher cannot be provided without an integration of the structure with its collection.

5.2 Dataset Acquisition

Access to McGill University Library and Laval University Library online catalogue databases was requested from their respective managements. The following raw data were sought:

- LCSH terms and descriptions,

- The relations between the LCSH terms (i.e., broader, narrower, subdivision, cross-references, etc.),
- The library holdings to which each LCSH term has been assigned.

This research required a sizable subset of the collections of these institutions. A sizable collection was defined as one that contained over 100,000 unique bibliographic records. Although this number may be considered small in comparison to modern Web collections it places this research amongst the top 10% of largest collections reported in IV testing literature (Julien, et al., 2008).

The science and engineering domain was selected because it offered a high concept hierarchy complexity; in other words, the breadth and depth of the concept hierarchy found in this domain would serve as a difficult trial for this research.

Management from McGill and Laval University Libraries returned prompt and positive responses stating they were willing to support the research. These datasets were provided once the researcher had signed an agreement stating the data would not be distributed or used beyond the research needs. Both datasets contained raw MARC data of bibliographic and authority records (i.e., the books and the list of subjects assigned to them). Both libraries provided the subset of their collection associated with science and engineering which contained at least 100,000 bibliographic records per library.

The McGill dataset was chosen rather than that from Laval because its predominantly English-language collection eliminated the need for multi-lingual features required by the Laval dataset. The McGill dataset contained 130,940 bibliographic records and 280,435 authority records. These raw data served as the basis for the ensuing data treatment.

5.3 Raw Data Import

Libraries can efficiently export raw MARC which is a "widely used yet not always understood"[1] data format standard developed in the 1960's to facilitate data exchanges between libraries. The MARC format is a comprehensive standard which defines tagged containers, known as fields and subfields, for all possible information associated with

[1] http://www.loc.gov/marc/makrbrkr.html

various types of bibliographic items (e.g., monographs, journal articles, audio recordings, sheet music, etc.). This data structure entails that most records contain many empty fields since they may not be applicable to the current type of bibliographic record; thus, fields likely to be non-empty and useful for a specific collection must be chosen carefully.

Raw MARC data are not legible and must first be converted. This conversion is referred to as *MARCBreaking*[1] and can be performed using the graphical interface offered by the MarcEdit software (Reese, 2000). MarcEdit also provides exporting capabilities to a number of commonly used character encoding formats (e.g., UTF, Unicode, etc.).

The descriptions and data examples of each MARC standard[1] fields were read and selected for semantic content rather than physical characteristics or shelving/ordering information. Extracted from the MARC records were all fields which could potentially contain semantic information; for example, title, subjects, series, publication. Not included were fields containing purely numerical information. Initial data collection from the MARC records aimed to gather as much useful information as possible, and the fields listed in Table 5.1 were chosen because they were likely to contain useful bibliographic and authority data:

[1] http://www.loc.gov/marc/

Bibliographic Fields		Authority Fields	
Code	**Description**	**Code**	**Description**
020	International standard book number	010	Library of Congress Control Number
022	International standard serial number	100	Heading - Personal Name
035	System control number	110	Heading - Corporate Name
100	Main entry - Personal name	111	Heading - Meeting Name
110	Main entry - Corporate name	130	Heading - Uniform Title
111	Main entry - Meeting name	148	Heading - Chronological Term
130	Main entry - Uniform title	150	Heading - Topical Term
245	Title statement	151	Heading - Geographic Name
250	Edition statement	155	Heading - Genre/Form Term
260	Publication, distribution, etc. (Imprint)	180	Heading - General Subdivision
300	Physical description	181	Heading - Geographic Subdivision
362	Dates of publication and/or sequential designation	182	Heading - Chronological Subdivision
490	Series statement	185	Heading - Form Subdivision
500	General note	400	See From Tracing - Personal Name
502	Dissertation note	410	See From Tracing - Corporate Name
600	Subject added entry - personal name	411	See From Tracing - Meeting Name
610	Subject added entry - corporate name	430	See From Tracing - Uniform Title
611	Subject added entry - meeting name	448	See From Tracing - Chronological Term
630	Subject added entry - uniform title	450	See From Tracing - Topical Term
648	Subject added entry - chronological term	451	See From Tracing - Geographic Name
650	Subject added entry - topical term	455	See From Tracing - Genre/Form Term
651	Subject added entry - geographic name	480	See From Tracing - General Subdivision
700	Added entry - personal name	481	See From Tracing - Geographic Subdivision
710	Added entry - corporate name	482	See From Tracing - Chronological Subdivision
711	Added entry - meeting name	485	See From Tracing - Form Subdivision
730	Added entry - uniform title	500	See Also From Tracing - Personal Name
740	Added entry - uncontrolled related/analytical title	510	See Also From Tracing - Corporate Name
830	Series added entry - uniform title	511	See Also From Tracing - Meeting Name
		530	See Also From Tracing - Uniform Title
		548	See Also From Tracing - Chronological Term
		550	See Also From Tracing - Topical Term
		551	See Also From Tracing - Geographic Name
		555	See Also From Tracing - Genre/Form Term
		580	See Also From Tracing - General Subdivision
		581	See Also From Tracing - Geographic Subdivision
		582	See Also From Tracing - Chronological Subdivision
		585	See Also From Tracing - Form Subdivision
		680	Public General Note
		781	Subdivision Linking Entry - Geographic Subdivision

Table 5.1: MARC fields chosen for export

Table 5.1 shows that a broad set of fields was chosen in order to extract as much information as possible from the raw data. Visual inspection of the data table showed *data scarcity*; very few fields contained the vast majority of the semantic content of each record, most fields were empty for most of the records. The main useful fields were title (field 245) and subject added entries (fields 600-651); some records also offered a semantically salient series information (field 490). Beyond these fields, the bibliographic records contained low semantic content. The same procedure was applied to the subject authority records which also showed data scarcity.

5.4 Data Treatment Application

A custom-developed Microsoft Access application performed the remaining data treatments. Microsoft Access is a rapid development environment specifically designed for data applications. An Access database is a highly portable data store which can be accessed by other applications. Applications which do not required more than five concurrent users and have relatively small tables (i.e., less than a million records) are particularly well suited to Microsoft Access. Data access performance is equivalent to industrial strength data servers, such as SQL Server, since they use the same kind of query processing engine.

An overview of the data treatment process is presented in Figure 5.1. It shows the successive steps to extract, format and reconstruct the LCSH hierarchy of the collection.

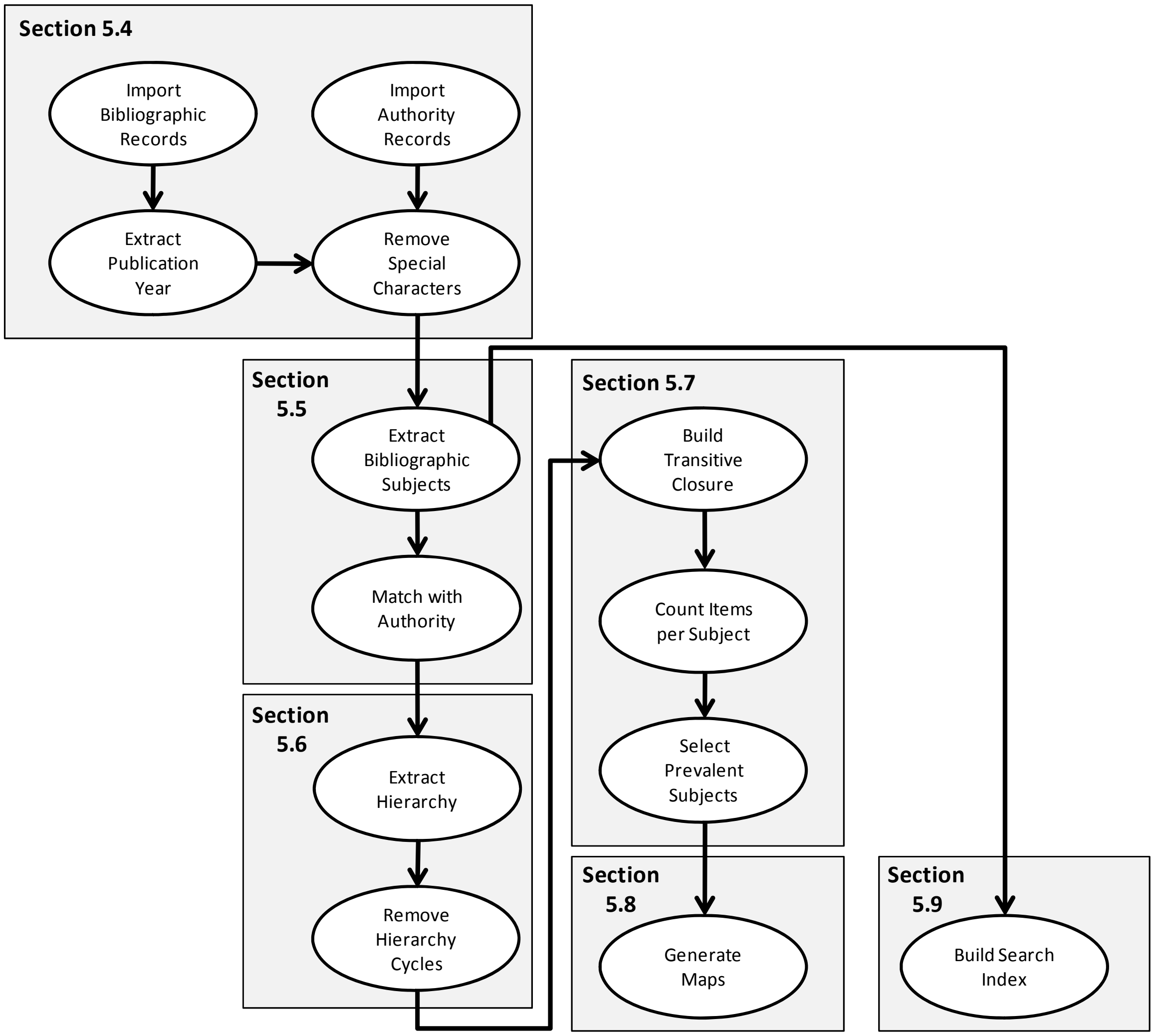


Figure 5.1: Overview of Data Treatment Process

This data treatment process was time consuming but was performed once after initial receipt of the data. This process served as an in-depth analysis of the LCSH relations and their subject assignments for a sizable real world collection. The details and stepwise effects of the data treatment process are described in the following sections referenced in Figure 5.1.

5.5 Import, Date Extraction and Clean-Up

The MARC data in CSV format was imported into two Microsoft Access data tables; one for bibliographic records and one for authority records. Both import routines accepted all records supplied: 130,940 bibliographic records and 280,435 authority records. From this point, only these two relational database tables were used as data sources—the original MARC data files were no longer required.

The MARC standard is not a normalized data storage schema; fields can contain multiples and various types of information. For example, there is no specific field for the publication year of an item; this information can be found as part of a number of other fields. Specifically, the last four characters of field 260 (i.e., Publication, distribution, etc.) should contain the publication year for monographs, and field 362 (i.e., Dates of publication and/or sequential designation) contains the range of available publications for periodicals. Based on these patterns, publication dates were extracted for 99.14% of the bibliographic records. The discrepancy was due to the fact that some bibliographic records did not contain a publication date in either field or the dates were obviously erroneous (i.e., far into the future).

Figure 5.2 presents the cumulative distribution of publication dates. It shows that the bulk of the collection were items published after 1940. The average publication date was 1976 and the median was between 1985 and 1986.

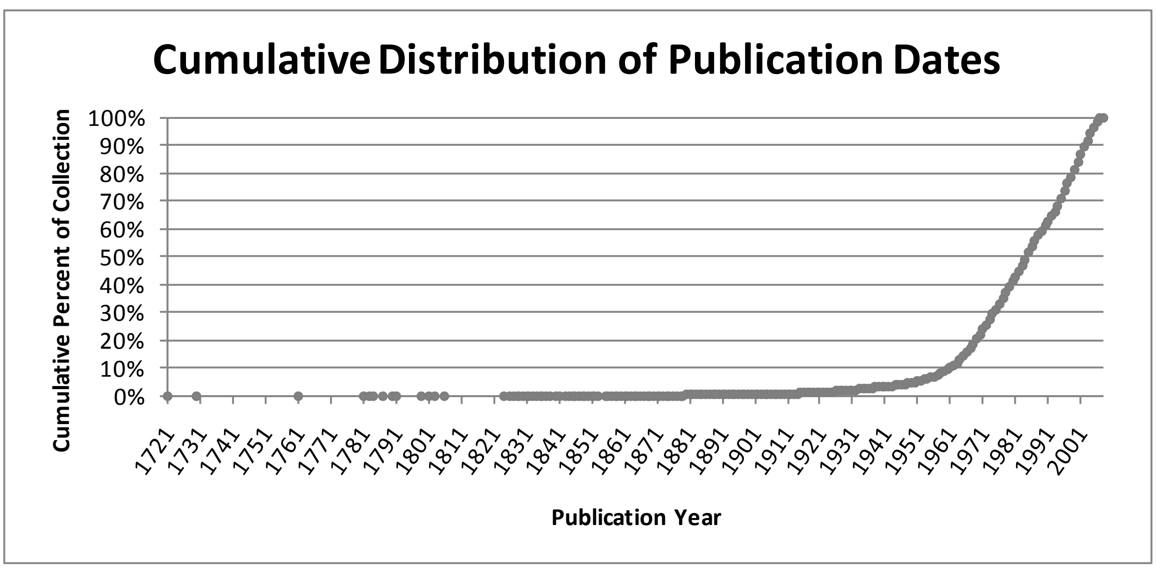


Figure 5.2: Cumulative Distribution of Publication Dates in Collection

MARC fields contain subfields identified by special combinations of characters (e.g., "$", "$ga", etc.). These characters identify various kinds of information contained in each field which often makes them illegible to untrained eyes; as a result, they were removed for usability. This was critical for the presentation of the bibliographic title (field 245), series (field 490), subject description (field 150), and related subjects (field 450).

5.6 Bibliographic Subject Extraction and Matching

Each bibliographic record has a field 650 (Subject added entry - topical term) which may contain one or multiple topical subject terms assigned to the item. This information is a string of characters where each subject term is delimited by a semi-colon (;) and should, at least in part, match an authority record. Field 650 contained information (i.e., non empty) for 122,393 bibliographic records or 93.47% of bibliographic items. This data were parsed and normalized into two additional tables containing the subjects found in the bibliographic records and their assignments to the records (see Figure 5.3).

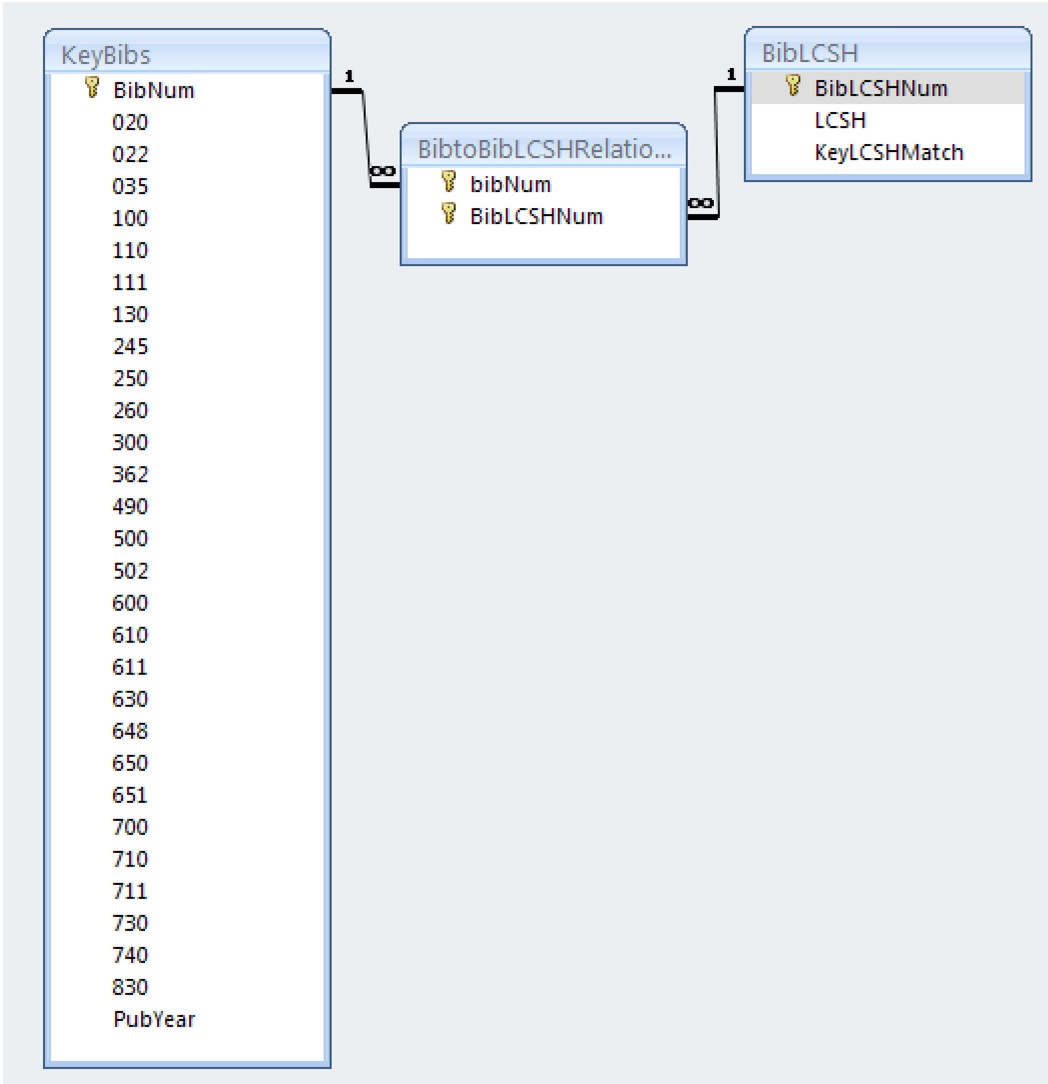


Figure 5.3: Normalization of Subject Assignments to Bibliographic Records

Figure 5.3 shows the bibliographic records table *KeyBibs*, the subjects extracted from field 650 in table *BibLCSH*, and the assignment table *BibtoBibLCSHRelations*. There were 63,515 distinct subjects extracted from bibliographic records. Each subject was assigned on average to at least two records for a total of 239,572 subject assignments, and there were almost two assignments (1.96 on average) per bibliographic record. Figure 5.4 presents the distribution of the number of subject assignments per bibliographic record. A power law distribution was obvious with 92%[1] of bibliographic items receiving three or fewer subject assignments and very few items receiving four assignments or more.

[1](49719 + 42366 + 20525)/ 122,393

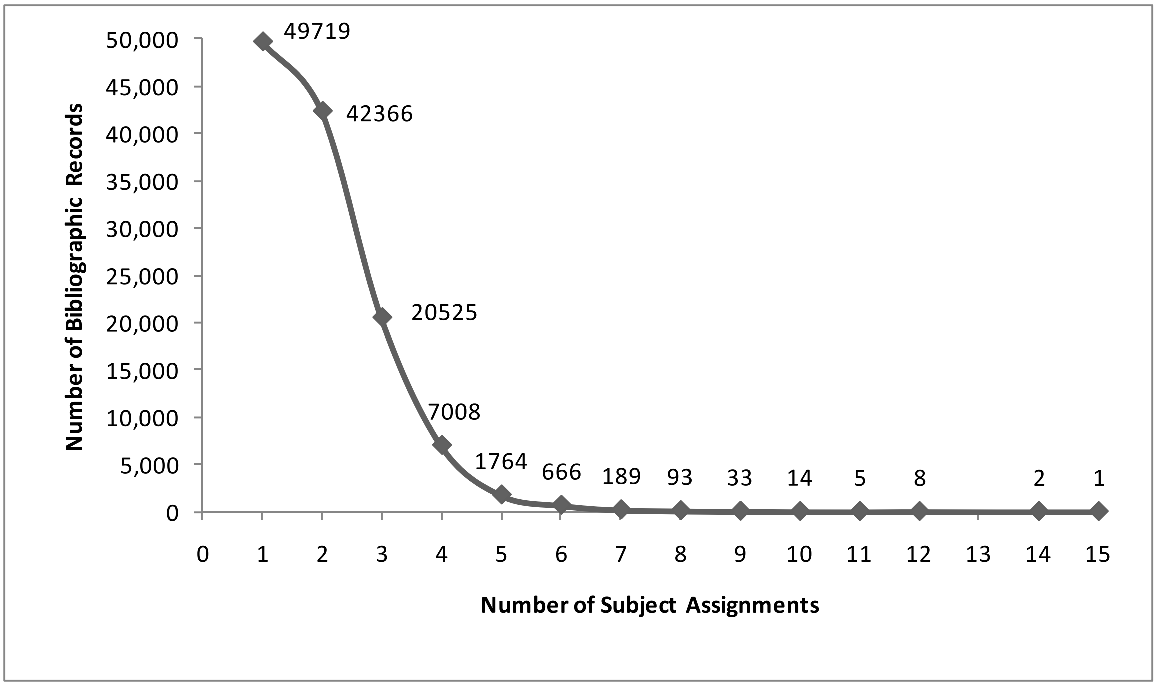


Figure 5.4: Distribution of Number of Subject Assignments per Bibliographic Records

As stated earlier, each subject term extracted from the bibliographic records should either exactly or partly match a record from the authority table. This match was critical because only the authority file defined the broad to narrow subject hierarchy.

Assuming correct spelling, subject terms which did not exactly match the authority file had likely been modified by the addition of one or many optional subdivisions (e.g., period, geographical region, topical, etc.). This meant that incremental removal of optional subdivisions could eventually allow exact matching of the original subject term to an authority record. Finding the authority match for each subject term found in the bibliographic records was performed using the following steps:

1. Search for the exact term in the authority records, if found then stop; else, continue to next step.
2. Remove the right most subdivision (identified by a double dash "--").
3. Search for the shortened subject term, if found then stop; else, if there are remaining subdivisions return to step 2.

Of the 63,515 subject terms found in the collection bibliographic records, 97.25% (61,768) were matched to an authority record. Figure 5.5 presents the distribution of successful matching methods. It shows that without removing any subdivisions, only 23% of assigned subjects matched an authority record. For the vast majority of assigned subjects (74%), incremental removal of subdivisions produced a match (labelled *Subdivision Relaxation* in Figure 5.5).

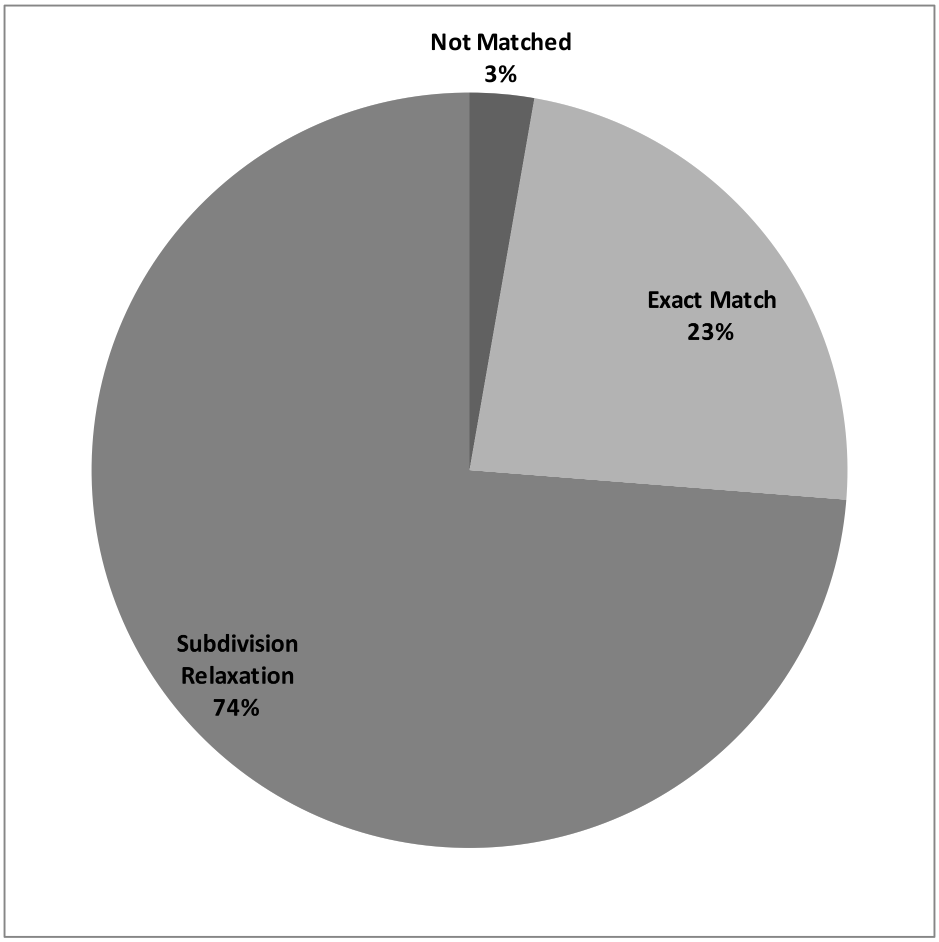


Figure 5.5: Subject Term Matching Methods

Removing subdivisions to match the authority table does not mean information was lost; on the contrary, the original subject terms (i.e., those with optional subdivisions) were kept with their bibliographic records. Subdivision relaxation was used strictly as a technique to find the matching authority record; indeed, it proved essential to placing each item in the LCSH hierarchy.

Less than 3% of subjects found in the bibliographic records could not be matched with an authority record. It is possible some of them were once valid subject headings which have since been removed from the authority records, or that some may have been part of copy cataloguing records which contained subject headings not part of local authority records.

A single bibliographic record had one or multiple subject headings which could match a subject authority record. A match meant the bibliographic record could be placed in the subject hierarchy. A bibliographic record without a single matching authority record meant a loss of access to that item since it could not be placed in the subject hierarchy. For this study, over 97% of bibliographic subjects were matched with authority records. This provided access to 99.87% of bibliographic records (122,228) with at least one subject assignment.

The subdivision relaxation process meant that some distinct subdivided terms were matched to their common root subject. For example, there were 48 subdivided versions of the subject "Acid rain" such as:

- Acid rain--Québec (Province)
- Acid rain--Environmental aspects--Congresses
- Acid rain--Ontario
- Acid rain--Québec (Province)--Charlevoix

None of these subdivided subjects had a corresponding authority record but they all matched the unmodified term "Acid rain". The practical impact was that items assigned to subdivided subjects were placed with their unmodified (i.e., without optional subdivisions) root subject term in the subject hierarchy.

Subdivision relaxation for term matching also reduced the complexity of the subject hierarchy. In effect, the number of distinct subject terms found in the bibliographic records was reduced to the number of distinct unmodified terms. In this case, this meant that the 61,768 bibliographic subjects were matched to 18,359 unmodified authority records. This amounted to a <u>70% reduction in hierarchy complexity</u> without loss of information or access to the collection.

The data treatment process had so far imported and cleaned the MARC data, extracted the subjects terms found in the bibliographic records, and matched these to the subject authority table. This list strictly contained the subjects specifically assigned to items in the collection; however, these did not provide the broad to narrow subject hierarchy which defined the collection domain coverage. This information could be found in the authority records and its extraction is described in the following section.

5.7 Hierarchy Extraction and Clean-Up

When used, authority field 550 (See Also From Tracing - Topical Term) is meant to contain one or many broader subject term(s) that should match another authority record. Using this information, searchers can effectively *crawl* up the subject hierarchy until no additional broader terms are specified. For the current research, the subject hierarchy was built by following broader terms starting from the set of 18,359 unique bibliographic subjects matched with the authority table. This process generated 31,708 broader relations with 9,670 unique broader terms.

Just over 2% (373/18359) of subjects extracted from the bibliographic records were conceptual islands; in other words, the subjects had no relations with other subjects. These subjects could not be linked to the subject hierarchy and were ignored. Their small number meant that the remaining subject hierarchy still provided access to 98.9% (121,102 / 122,393) of bibliographic items having at least one subject heading.

At this point in the process it was possible to traverse and inspect the subject hierarchy relevant to the collection at hand. There was a large number of broadest terms (i.e., those without a broader term). A visual inspection of these terms showed many seemed to be narrower subjects; for example, "Building, Wooden" had no specified broader term but "Building" is a valid subject and could, arguably, be a conceptual broader term. It seemed a comma could act as a subdivision identifier. These *comma delimited broader terms* also included the following examples:

- "Chemistry, Analytic" and "Chemistry, Organic" might both belong to "Chemistry"

- "Functions, Continuous", "Functions, Meromorphic", "Functions, Meromorphic", "Functions, Theta", and "Functions, Zeta" could all belong to "Functions"
- "Steel, Heat resistant", "Steel, Stainless", "Steel, Structural" may all belong to "Steel"

This pattern was exploited and 134 subjects which were previously thought to be broadest terms were inserted into the existing subject hierarchy. These *derived* broader relations were not explicitly specified in field 550; nonetheless, the relations seemed defensible and further reduced the complexity of the subject hierarchy.

Other attempts were made to reduce the number of broadest terms; however, these created obvious false relations and were not used. For example, for some subject terms, a subset of their words or characters was in fact another valid subject heading which could be argued to be a conceptual broader term. Although this could be spotted by the human eye, this was difficult to exploit in a controlled manner without *a priori* knowledge of which subset of the term should be matched.

5.7.1 Eliminating LCSH Cycles

Inspecting the succession of broader terms revealed that the subject hierarchy contained conceptual traps. For example, following the succession of broader terms as specified by the LCSH data:

- "Safety appliances" belongs to "Industrial safety",
- "Industrial safety" belongs to "Accidents--Prevention",
- "Accidents--Prevention" belongs to "Safety appliances"
- "Safety appliances" belongs to "Industrial safety",
- "Industrial safety" belongs to "Accidents--Prevention",
- "Accidents--Prevention" belongs to "Safety appliances"
- "Safety appliances" belongs to "Industrial safety",
- *ad vitam aeternam*; an endless conceptual loop.

These conceptual cycles are usability issues. It may be difficult for users to grasp how "Safety appliances" is more general than "Accidents--Prevention" as well as being more specific than "Accidents--Prevention" via "Industrial Safety". This also makes it very difficult to navigate the subject hierarchy.

The chosen solution was to delete the least number of relations in order to eliminate the conceptual cycle; in the example above, "Accidents--Prevention" cannot belong to "Safety appliances". A routine was developed to find and eliminate 26 such conceptual traps.

5.7.2 Other Relations Between Subjects

Authority field 550 can also contain related subject term references. These types of relations are not hierarchical; they are explicit and direct paths between branches of the hierarchy. Related terms could act as *short-circuit* or *warp* links between subjects allowing searchers to quickly jump from one part of the hierarchy to another. These are interesting because SW ontologies also contain different kinds of relations between concepts (Le Grand & Soto, 2006).

The related terms defined for subjects used by the collection were extracted and inspected. There were 4840 unique related terms assigned to 4138 subjects; however, the simplified subject hierarchy contained only 152 of these relations. They could have been visually represented in the SE-3D structure but this was beyond the scope of the current research. Visually representing various kinds of relationships between concepts will be necessary as these types of applications proceed through their development process.

5.8 LCSH Hierarchy Reconstruction

Reconstructing the LCSH subject hierarchy required an analysis of item distribution throughout the subject hierarchy. This would reveal the few subject branches that contained most of the collection. For a large hierarchy such as LCSH, this type of operation demands too much processing time for interactive software.

Data hierarchies are also called trees or directed acyclic graphs (DAG). Efficiently processing these often large data structures is an active problem in computer science (see for example Bender, Farach-Colton, Pemmasani, Skiena & Sumazin, 2005; Czumaj, Kowaluk & Lingas, 2007). Trißl *et al.* (2005) list common tasks performed on DAGs:

- **Reachability**: does a subject eventually belong to another or is it part of completely separate branches? The number of subjects travelled between the two is irrelevant.
- **Ancestors/Successors**: given a subject, what are all its broader terms, no matter how far removed? What are its more specific terms, no matter the distance?
- **Least common ancestor**: given two subjects, do they eventually belong to a common parent? If so, find the most specific or lowest common parent.

Solving these types of questions is classically done by recursively crawling up and down the subject hierarchy; however, this quickly becomes prohibitively slow as the size of the structure increases. Performance can be greatly increased by first computing the transitive closure of the hierarchy.

5.8.1 Transitive Closure

The transitive closure of a subject hierarchy is the list of subject pairs that are linked, irrespective of the number of subjects between them. Generating the transitive closure from the subjects and their relations can be done in several ways. Trißl *et al.* (2005) recommend the *Logarithmic* algorithm found in Valduriez *et al.* (1986), and they provide an SQL[1] implementation. Based on this algorithm, the transitive closure of the collection-specific LCSH hierarchy was generated.

The resulting transitive closure table contained linked subject pairs and the number of subjects between them. For example, Table 5.2 contains transitive closure records which show that the subject "Airports" had nine direct children (e.g., Airports--Planning, Access to airports, etc.) found at a distance of a single subject away.

[1] Structure Query Language

From Subject	To Subject	Number of Subjects Between
Airports	Airports--Planning	1
Airports	Access to airports	1
Airports	Airports--Maintenance and repair	1
Airports	Airports--Baggage handling	1
Airports	Airport buildings	1
Airports	Airports--Environmental aspects	1
Airports	Airports--Management	1
Airports	Airports--Cold weather conditions	1
Airports	Runways (Aeronautics)	1
Airports	Roads--Snow and ice control	2
Airports	Runways (Aeronautics)--Maintenance and repair	2
Airports	Airport terminals	2
Airports	Airports--Access roads	2
Airports	Slush on pavements, runways, etc.	2
Airports	Bridges--Snow and ice control	2
Airports	Runways (Aeronautics)--Snow and ice control	3

Table 5.2: Transitive Closure Data Example

Table 5.2 also shows that, via some of its direct children, "Airports" was also linked to six subjects at a distance of two subjects; finally, "Runways (Aeronautics)--Snow and ice control" was also accessible from "Airports" if the user was willing to travel three subjects away. This is a simple case and a broad subject such as "Algebra" was linked to over a thousand other subjects at a distance of up to eleven subjects.

The transitive closure was useful to efficiently answer questions such as how many bibliographic items were accessible via "Airports"? Accessibility meant more than finding the number of items which were directly assigned to that subject; it included items assigned to any of its more specific subjects, regardless of the distance. This fact quickly became a significant processing issue when considering that, for example, a single item assigned to "Hydraulic presses" was accessible from over 300 other subjects. This meant the number of items for each of these 300 subjects had to be incremented by the number of items directly assigned to "Hydraulic presses". Multiplied by approximately a million links between subjects, the processor demands for this operation became impractical.

The transitive closure provided fast information concerning all relations between subjects used to efficiently increment their number of accessible items. Any item assigned to multiple subjects was duplicated; thus, multiple subject access points for a single bibliographic record were accurately represented throughout the hierarchy, at every level of abstraction. This practice is not unique: Crop Circles (Wang & Parsia, 2006) also added duplicated subtrees "due to multiple inheritance" (p. 702).

5.8.2 Cumulative Count of Subject Assignments

Inspection of direct subject assignments confirmed a Bradford's Law distribution as shown in Figure 5.6.

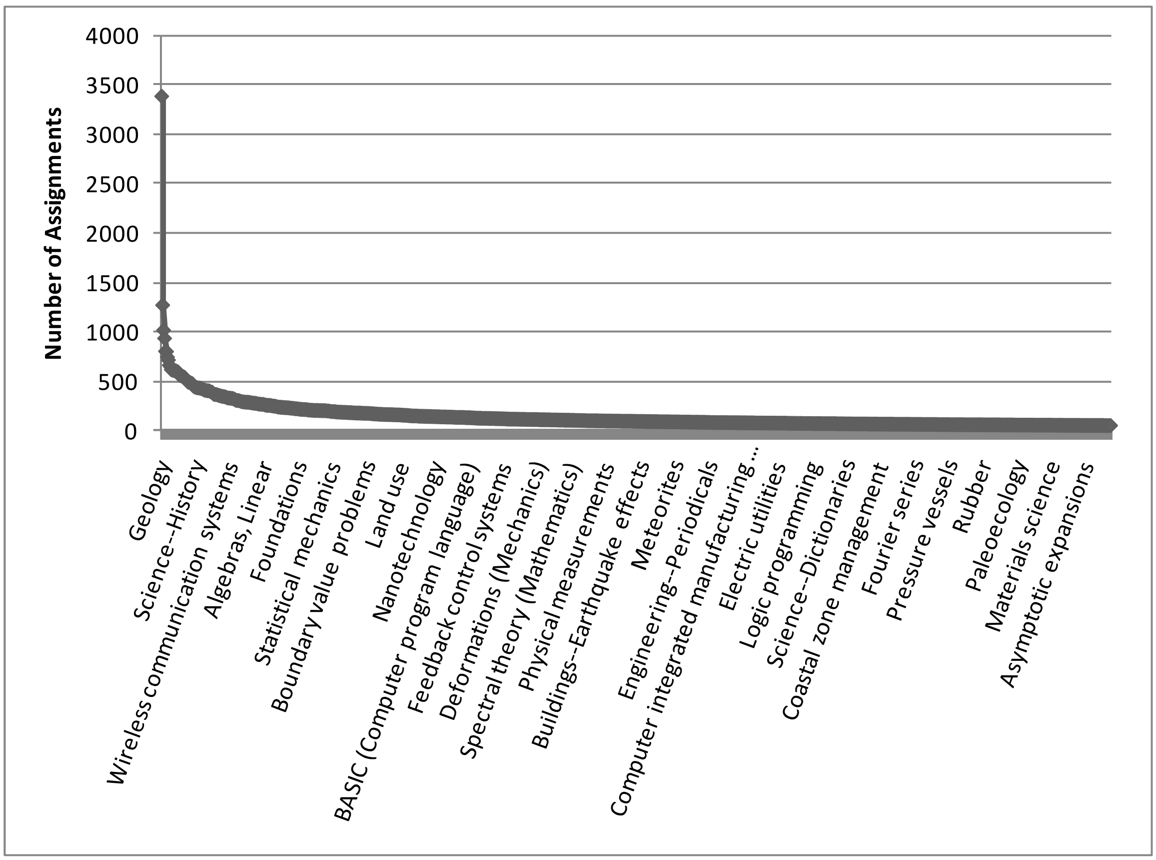


Figure 5.6: Distribution of Direct Subject Assignments

Figure 5.6 includes only the first 1000 most assigned subjects but clearly shows that a small number of the subjects contain many items while most subjects contain very few. The next data treatment step was to propagate these assignments throughout the subject hierarchy. For example, the transitive closure data showed that the 3400+ items assigned to "Geology" were accessible from "Science", "Earth sciences", "Environmental sciences", and "Physical Sciences". Propagating the direct subject assignments throughout the subject hierarchy would allow pruning of the LCSH hierarchy based on item distribution within the specific collection.

5.8.3 Prevalent Subject Selection

Simplifying the subject hierarchy began with its inspection and two useful phenomena were exploited to significantly reduce its complexity. The two data phenomenon exploited were *single* and *near single* child abstraction levels as described in the following section.

Single Child Abstraction Levels: the subset of the LCSH hierarchy assigned to the collection contained multiple instances where a subject acted as the lone narrower term providing no division of its broader parent subject. For example, "Spring Ecology" had a single more specific subject "Hot Spring Ecology" which in turn had multiple children. This implied that any subject more specific than "Spring Ecology" was also about "Hot Spring Ecology". This intermediate single class could be removed and concatenated with all its children, and the hierarchy was simplified by pruning an abstraction level without loss of information. There were 1476 such occurrences.

Near Single Child Abstraction Levels: inspection of the distribution of items throughout the subject hierarchy showed that, in many instances, at least 95% of the items under a subject belonged to a single child subject. This threshold was arbitrarily chosen as the quantity beyond which, based on subject assignments, a subject effectively had a single more specific subject. These cases were then treated in the same manner as a single child abstraction level previously described.

The exploitation of these two item distribution data phenomenon significantly reduced the complexity of the subject hierarchy. Specifically, the depth of the hierarchy was reduced to 22 levels from 25, and the number of transitive closure links was reduced by over 28% from over 830,000 to less than 600,000.

The subject hierarchy also became more conceptually intuitive and usable. This is illustrated by the case of "Engineering". According to the unmodified LCSH structure, this broad and contextually highly relevant subject is a more specific subject of both "Technology" and "Industrial Arts" as illustrated in Figure 5.7.

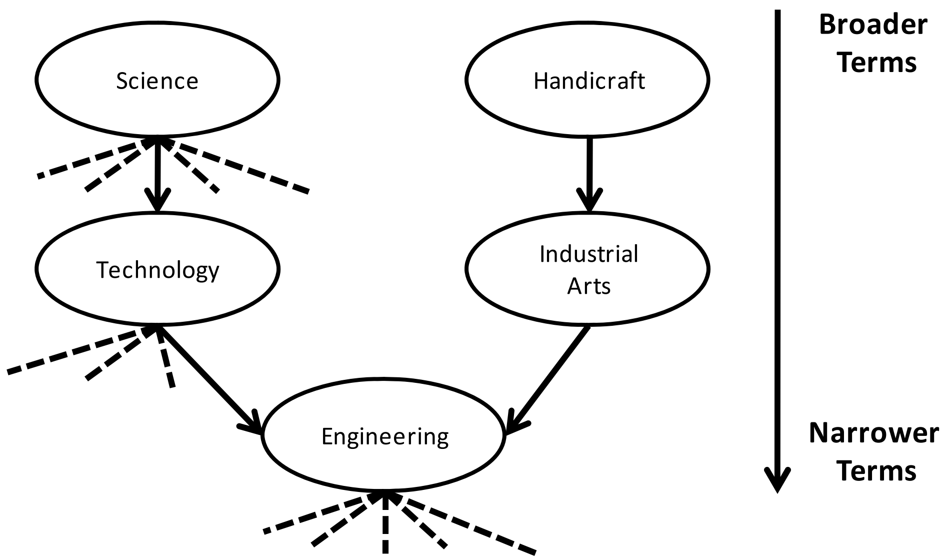


Figure 5.7: Local Hierarchy of Engineering Subject before Selection

As Figure 5.7 shows, there was a conceptual relation between "Engineering", "Industrial Arts" and "Handicraft", but the former was a single child abstraction level, and "Engineering" accounted for more than 95% of the contents of "Handicraft". These data facts rendered the relations irrelevant and reduced the complexity of the subject hierarchy as shown in Figure 5.8.

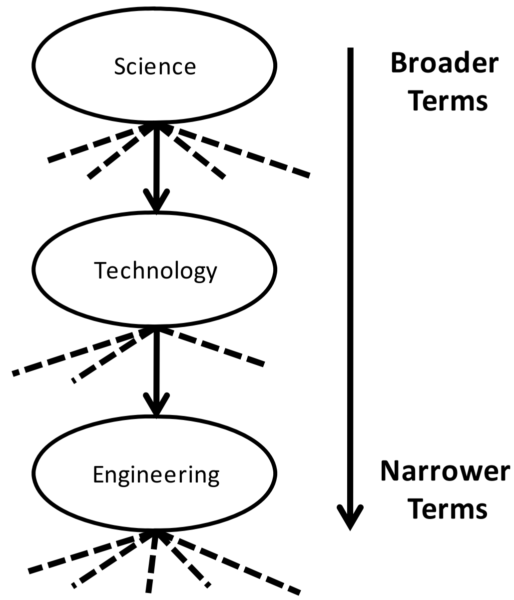


Figure 5.8: Local Hierarchy of Engineering Subject after Selection

Arguably, the hierarchy shown in Figure 5.8 better represents the contents of a science and engineering collection as opposed to Figure 5.7. Applied throughout the large LCSH subject hierarchy, these simplifications isolated subject branches and produced a clear subject hierarchy relevant for the collection at hand.

The data treatment application had so far imported and treated the raw MARC data, extracted subjects assigned to bibliographic records and matched them with the authority records, and extracted and simplified the subject hierarchy used by the collection.

5.9 Generate the Subject Maps

SE-3D had to be interactive and responsive to user actions; consequently, real-time processor intensive tasks had to be optimized through pre-processing. Generating the transitive closure (see section 5.8.1) is an example of pre-processing which facilitated subject hierarchy operations; in addition, SE-3D performance could be improved by *a priori* generation of the exact definition of the subject hierarchy visual representation. This allowed SE-3D to perform fewer real-time calculations to update the visual subject hierarchy in response to user actions.

The definition of the visual subject hierarchy included the following data:

- **Subject Maps**: represents a subject with narrower child subject. Each subject map was defined by the parent subject whose children it contained, and the total number of bibliographic items to which it provided access.
- **Map Contents**: each subject map was associated with the child subjects it contained. Map content definition included the relation between the parent and child subjects, child positions on the map, and their circle size proportionate to the number of items accessible via the child subject.

Section 5.8.3 showed how the complexity of the subject hierarchy was significantly reduced based on collection distribution; in addition, the visual representation was further refined to provide a clear representation of the prevalent collection subject branches. This was essential since many subjects contain very few items. These numerous but seldom assigned subjects were not representative of the collection distribution pattern; in fact, their sheer number could overwhelm the visual representation of the subject hierarchy.

Reduction of subject hierarchy noise was based on a set of heuristics which were iteratively developed and tested with the collection. Specifically, starting from the top of the hierarchy (i.e., subjects without broader terms), for all levels of abstraction, subjects containing the most items were considered first and the decision to include a subject in the visual representation was made based on the following rules:

- **At least 86% of contents must be shown**: the child subjects visually represented must contain the vast majority of the bibliographic items accessible from a set of direct children (i.e., at least 86%). This lower bound was iteratively determined to produce a good compromise between data distribution representation and reduction of visual noise. These types of threshold values are not uncommon and were also used by Fowler, Fowler & Wilson (1991).
- **A maximum of 99% of contents should be shown**: this upper bound ensured that subjects were not cluttered with too many very small child subjects (i.e., child subject with very few bibliographic items assigned). This was essential in cases where dozens of very small subjects could be added to the visual structure while never reaching exactly 100% of items. This is explained by the long tail of the Bradford's Law distribution (see section 5.8.2); in other words, it is practically impossible to include 100% of the items without including the multitude of subjects assigned to almost no items at all.

- **At least 1% of total child items**: a child subject was considered noise if it contained less than 1% of the total number of bibliographic items contained by all children of the immediate parent.
- **A maximum of eight subdivided subjects**: for any subject with multiple child subjects not considered noise by the previous rules, a maximum of eight subjects could lead to their own subject map. This restriction was devised to limit the visual breadth of the hierarchy since each map required a fixed amount of volume in a limited space; eight child maps could be efficiently spaced out in a grid around the parent map. In practice, this rule was seldom applied because very few subjects had more than eight children which were themselves parents; in fact, there were only 18 such cases ranging from nine to 14 sub maps. In those few cases, the remaining children were displayed as undivided subjects; furthermore, these undivided subjects were necessarily the least significant in terms of the number of bibliographic items they contained.

The complete data treatment including the above heuristics produced a clear pattern of collection distribution within the LCSH subject hierarchy. The resulting structure contained 1828 predominant subjects which provide access to a total of 13,997 subjects or only 22% of the subjects extracted from the bibliographic records (see section 5.6). Yet this relatively small number of subjects provided access to 92.81% of bibliographic items which contained at least one subject assignment; in other words, a 78% reduction in subject hierarchy complexity resulted in a loss of access to only 7.19% of collection items. This dramatic reduction of complexity was suggested by Simon (1996) stating that "most of the complex structures found in the world are enormously redundant, and we can use this redundancy to simplify their description" (p. 215). Fowler *et al.* (1991) also report this type of phenomenon in their associative thesaurus application.

The 7.19% loss of collection access was incurred throughout the data treatment process. For example, unmatched subjects between bibliographic items and subject authority resulted in a 0.13% loss (see section 5.6), the conceptual islands described in section 5.7 produced a further 0.89% loss, and the remainder was lost during the hierarchy simplification process. No matter how small, loss of access to a part of the collection is unfortunate. Inspection of the subject assignments to these items may provide some clues to mitigate this issue, but this is beyond the scope of this research.

Visually showing 1828 subjects while providing access to a total of 13,997 subjects meant that 12,169 (13,997−1828) subjects were implicitly part of the subject hierarchy. The contents of these *invisible* subjects were part of their nearest parent subject whose number of accessible items was significant. This meant that most subjects with few items were eventually represented by a broader parent although the bibliographic item still contained the original more specific subject. This had the benefit of omitting relatively small groupings of items belonging to highly specific or ill-covered subjects, while still providing access to their bibliographic items via a significantly large broader parent.

5.10 Build Search Index

The current research did not aim to produce an innovative search engine; the innovation was the interactive information visualization interface integrated with a classic search engine. The search index generation process follows established practice in natural-language statistical analysis and information retrieval systems design.

5.10.1 Search Index Choice

A search index is the list of terms known to a search engine. These terms are extracted from the collection to be searched, sorted alphabetically and linked to each document that contains them. For this study, terms were extracted from the following fields:

- Bibliographic field 245–Title statement
- Bibliographic field 490–Series statement
- Authority field 150–Heading - Topical Term
- Authority field 450–See From Tracing - Topical Term

The search index representation of each bibliographic record was represented by its title, series statement if available, the subjects to which it belonged and the synonyms of those subjects. These fields were selected because inspection of the data showed they contained semantic information relevant to the testing performed for this research. Other fields contained in the MARC structure were either mostly empty and/or contained information which pertained to physical characteristics of the item as opposed to its semantic contents.

5.10.2 Stemming

Terms found in the collection and query terms entered by users are usually stemmed; they are reduced to their common base form (e.g., organize, organizes, organizing). Porter's algorithm (Porter, 1980) is "the most common algorithm for stemming English" (Manning, Raghavan & Schütze, 2008, p. 31). Manning (2008) states that *Porter stemming* "has repeatedly been shown to be empirically very effective" (p. 31). Porter stemming implementations are freely available in various programming languages and Porter (Jan. 2006) offered a Visual Basic implementation produced by Christos Attikos. This code was ported to Access Visual Basic for Applications for the purposes of this research.

5.10.3 Relevance Ranking

Relevance ranking of results has become the norm with standard search engines. The vector space model and tf-idf weighting is an established and reliable method which provides good results ranking (Manning, et al., 2008 Chapter 6). The vector space models documents and queries as lists of terms; the more similar the two lists, the more relevant the document for the query. Term weighting assigns more or less value to a term based on its prevalence in documents and the collection as a whole. The classic weighting scheme tf-idf means two types of weights:

- **term frequency (tf)**: each term found in each document is assigned a weight proportional to the number of occurrences in that document. The weight is higher if the term occurs more often in the document and is normalized so longer documents are not systematically favoured.
- **inverse document frequency (idf)**: query terms entered by the user are given a weight inversely proportional to the number of occurrences of the term within the whole collection. The weight is high for terms that seldom appear in the collection.

Search results can be easily ranked using these weights. The importance of each document for a specific query term will be highest if the term is not found often in the collection but is found often in the document; in other words, a highly discriminate query term with many occurrences in a document will move that document nearer the top of the result list.

5.10.4 Search Index Characteristics

The selected data fields were parsed, and over 1.8 million terms were extracted producing a stemmed dictionary containing over 87,000 unique terms. A list of 38 stop words (see Annex 6) was used to filter words of little value for IR such as conjunctions, pronouns and articles. The science and engineering domain of the collection showed that single letter terms could be significant (e.g., "C" is a programming language), this was also the case for punctuation marks (e.g., "C#" is a programming language different from "C"). Inspection of the term dictionary reflected the domain of the collection. Stemmed terms from variants of "engineering", "science", "computer" and "system" had the highest collection frequency.

5.11 Corpus Analysis Conclusion

Data patterns can be found in the organized information collection; for example, the corpus analysis showed the LCSH structure was in large part hierarchical, and just a few subjects at every level of abstraction contained most of the collection. This provided a way to trim and reconstruct the subject hierarchy for the purposes of information retrieval. By reconstructing the LCSH hierarchy, this study explored to what extent the complexity of the structure could be reduced while maintaining access to the highest proportion of the collection. The assumption was that a simpler yet representative subject structure could provide an advantage to information searchers.

Findings made during the corpus analysis are compared with other reports of analysis of large collections organized by subjects. They add to research concerning subject assignment patterns, characteristics and reconstruction of information organization structures.

5.11.1 Headings per Bibliographic Record

There were on average 1.96 subject assignments per bibliographic record in the collection (see section 5.6) which compares with Frost *et al.* (1988) who found 1.8 subject headings per record. These numbers are high compared with McClure (1976) who estimated a stable average of 1.3 subject headings per item in the National Union Catalogue for the period between 1950 to 1973, and O'Neill *et al.* (1981) who found 1.41 subject headings per record in their study of OCLC cataloguing. Tonta (1992) found 3.44 headings per item which "seems to indicate policy change at LC for the number of headings to be assigned" (Fischer, 2005, p. 75). The findings of this research would support the idea that the number of assigned subject headings is increasing but they would cast doubt on Tonta's findings.

The dramatic difference between the finding of this research (average 1.96 headings per records) and Tonta's results (3.44) can be partly explained by the methodology used by the latter. The author chose a narrow sampling of 237 titles published strictly in 1987 in the field of LIS (Tonta, 1992, p. 15); as a result, the reported average number of subject assignments reflects the LC subject assignment policies of that year for that domain. Comparatively, the tested collection of science and engineering works contained 2,545 works published in 1987; they contained an average of 2.2 subject assignments per item. This difference suggests the works in the multi-disciplinary field of LIS may receive more subject headings than works in the broader test collection used in this study. Further research is necessary to verify the effect of knowledge domain on the number of subject assignments.

This research analysed the complete collection of physical items found in a science and engineering library (over 100K items) with publications from the middle of the 18th century through 2008 (see section 5.5). This would dampen the effects of more recent subject heading assignment practices. The numbers presented in this section suggest the tested collection contains a typical distribution of subject assignments per item and could be representative of other collections organized using LCSH.

5.11.2 Subject Headings Matching Authority

The proportion of subject terms assigned to bibliographic records having an exact match with an authority record was just above 23% (see section 5.6). Frost *et al.* (1988) reported that, in a sample from the University of Michigan Library collection, 44% of topical and geographic headings were an exact match with the LCSH authority. Ludy *et al.* (1985) studied part of the Ohio State University collection and found only 10% exact matches between 410K subject headings found in the bibliographic records and the official authority records from LC.

These low proportions of exact matches with authority records are explained by the systematic practice of adding free-floating subdivisions based on the perception of local needs by each library (see for example Hoffman, 1999; Smith & Cochrane, 1999; Wilk, Rotenberg, Schacham, Hoffman & Liebman, 2001). Specifically, since 1974, the free-floating subdivisions have produced a system where individual libraries often modify LC assigned subject headings with free-floating subdivisions (Shubert, 1992, p. 63). These have long been a part of LCSH system "although today they are more abundant in our catalogs and more difficult to apply than was probably ever imagined" (Conway, 1993, p. 47).

Considering this pattern of local additions of free-floating subdivision, unmatched subject headings found in bibliographic records were traced back to an authority record by removing the optional subdivisions. This proved highly effective for this research; 74% of subject headings found in bibliographic records were matched using the technique. Chan *et al.* (2000) suggested separating non-topical elements from subject heading strings, and Frank *et al.* (2004) removed all subdivisions to match the LC classification. Removing the free-floating subdivisions may benefit the user since "ultimately, the meaning of these strings is lost on most cataloger users and some catalogers" (Drabenstott, Simcox & Fenton, 1999, p. 384).

There is a clear benefit of removing LCSH optional subdivisions for the purposes of subject heading matching with authority records. This allowed a quasi complete correspondence between bibliographic records and the subject authority file. This process simplifies the information organization which may benefit the searcher in understanding and using the subject headings.

5.11.3 Revealing the LCSH Hierarchy

Working with large hierarchies is an ambitious endeavor and few researchers are willing to work with these complex structures. The largest hierarchy found is reported by Liu *et* *al.* (2005) who attempted an automatic text classification (TC) algorithm using the Yahoo! Directory structure which contained 246,279 categories organized into a 16-level hierarchy. A somewhat comparable collection with this research, Wang *et al.* (2007) attempted to extract the hierarchical structure from the DDC using 88,440 bibliographic records from Science and Technology (BDS&T) from OCLC WorldCat. Their structure contained 18,462 categories forming a hierarchy as deep as 23; "the deepest taxonomy ever tested in TC studies" (p. 139). The tested collection contained over 122,000 bibliographic records and over 63,000 subject headings with 25 hierarchy levels. This places the current research amongst the largest studied subject hierarchies.

Yi *et al.* (2010) analyzed the syndetic structure of the pure LCSH (i.e., the official list from LC) and found skewed distributions of relationships; in other words, as the number of relations increases (e.g., broader terms, narrower terms, etc.), the frequency of subjects decreases. The authors also found that LCSH offered a few large groups of interconnected subjects, but most subjects were connected to few. They found that the largest hierarchical group was associated with "Science". This suggests SE-3D may be less useful in less structured subject domains but this would have to be verified in future research.

Lack of comparable literature suggests this study is one of few to provide an automated and demonstrated method addressing the known syndetic structure issues associated with the very large LCSH hierarchy. Finally, this research is broader and more ambitious that the two cited TC studies (Wang *et al.* 2007; Liu *et al.* 2005) since the resulting hierarchy of subjects is integrated with the collection in a fully functional and testable IR system.

5.11.4 Semantic Structure Reconstruction

Wang *et al.* (2007) is the only study found which can be partly compared with the subject hierarchy data treatment portion of this research. The authors attempted to extract the hierarchical structure contained in the DDC. They encountered severe data sparseness; for example, "more than 60% of the categories have just 1 document" (Wang & Lee, 2007, p. 140). For the tested collection, almost 65% of subject headings contained only one document.

In agreement with observations made during this study, they report that it "is clear that the document quantity per category follows the power law distribution" (p. 140). To illustrate this skewed distribution of subject assignments they report that the top 1% most populous categories contained 25% of all their documents. In comparison, over 58% of the tested collection was contained in the top 1% most populous LCSH terms found in the bibliographic records.

As suggested by Wang *et al.* (2007), the information searcher may require a path different from the human classifier—there was no other research found that takes this approach for a functional IR tool. The approach assumes that a searcher would browse a different subject hierarchy from the classifier. This is arguably intuitive since both types of tasks and users have different objectives: 1) the searcher is looking for information, and 2) the classifier is looking for an appropriate subject class for a piece of information in hand. These objectives are very different and may require different tools; even so, the classifiers' hierarchy is currently the only one available.

For this research, highly effective hierarchy noise reduction measures were single and near single child abstraction level simplification (see section 5.8.3); moreover, like Bradford's Law power distribution, these data phenomena are likely present in other collections. These techniques could likely be applied to other data sets in order to reveal the distribution pattern of a collection within other subject ontology. This generalization would require further testing which is beyond the scope of the current study.

5.12 Limitations

This study analyzed a complete bibliographic collection and determined its distribution within a subject hierarchy in the domain of science and engineering. The findings and conclusions are specific to the collection analyzed and cannot be generalized to other organized information collections. The results may vary depending on the:

- information organization scheme
    - How is the collection organized?
    - What are the assigned subject terms and relationships between them?

- knowledge domain
  - What are the subject areas covered by the collection?
  - How broad or specific is the collection?
- availability of the collection
  - How much of the collection can be analysed?
  - How much sampling is necessary to reveal distribution patterns of the collection?

Assuming an organized collection is available (e.g., bibliographic or Semantic Web collections), the real-time updating of the collection distribution eventually becomes impractical. This is especially true for quasi-infinite online collections. A random sampling might reveal the same patterns without requiring access to the complete collection. This would require additional research concerning improved bridging between information organization and information collections.

Real time updating may not be necessary since incremental additions to the collection are not likely to significantly change the overall pattern of subject distribution. Simon (2006) specifically describes the Library of Congress (LC) as a growing organism with structure (i.e., LCSH). The growth of an LCSH organized collection does not imply an increase in LCSH complexity; indeed, an "increase in the weight of a steer...does not" (p. 99) imply the structure of the animal has changed. Analogically, trees fall and grow all the time but, baring catastrophic events, the forest changes very slowly.

The nature and effect of the information organization scheme may well depend on the knowledge domain covered by the collection. For example, a controlled subject vocabulary such as MeSH is limited to medical information, and Yi & Chan (2010) showed the broad LCSH contained highly connected and hierarchical subject areas (e.g., science). LCSH coverage of a different domain (e.g., social science) collection would produce a different subject structure. Future research could explore the specific impacts of the knowledge domain on the resulting subject structure.

Chapter 6: 3D IV APPLICATION DEVELOPMENT

A novel virtual reality IV subject structure and search engine was developed. The Subject Explorer in 3D (SE-3D) offers a visual representation of the subject hierarchy, works associated with each subject can be shown in a textual list, and keyword searching highlights matching areas within the visual subject structure. Innovative solutions were developed to address issues of object placement, label occlusion management, and specific effects of keyword searching on the visible parts of the subject hierarchy. Results of formative testing guided the choice of navigation metaphor and suggested adding animated sequences of movements to simplify movement in the 3D space.

6.1 Introduction

This development project aimed to facilitate information retrieval by capitalizing on existing information organization structures. Trained information professionals use these controlled vocabulary structures but few lay searchers do (see section 1.2.1); this suggests there is value in the information organization data but extraction is hampered by inadequate usability of the online access tools.

SE-3D aimed to use the existing subject hierarchies offered by large organized collections as a visual metaphor for interacting with the information collection. The application visually represents the LCSH subject hierarchy in virtual reality integrated with keyword searching. Novel human-computer interaction techniques were developed and tested in this VR information retrieval environment. There are few reports of this kind (see section 4.2); however, design is informed by knowledge stemming from virtual reality world design and navigation, graphic arts, human-computer interaction, and information retrieval.

6.1.1 3D Objects as Information Representations

Visually representing a subject hierarchy is an information visualization design problem. There is no real world counterpart on which to base the design of the HCI; however, there were prior works from which design can draw inspiration. As stated in section 4.3, SE-3D was inspired by of mix Cone trees and spatial displays. Cone trees use basic shapes connected via lines to suggest parent-child relations. Spatial displays are often basic shapes (e.g., circles or squares) placed on a 2D plane or background.

As stated in section 3.2.1.3.1 (Level of Realism), there are indications that a simplistic representational set of objects provides more effective visual communication as opposed to photorealistic renderings. Perhaps the strongest evidence of this principle is the controlled experiment from Modjeska and Waterworth (2000) cited in section 4.2.2.3.2 (Information Islands). Their results showed that basic geometrical shapes resulted in higher efficiency as opposed to more 'naturalistic' renderings. Inspection of the multiple IV examples provided in section 4.2 reveals that all existing IV prototypes and commercial applications were designed using representational shapes (combinations of circles, rectangles and lines).

6.1.2 Object Positioning

Positioning large numbers of objects in a finite 3D space is difficult. One objective is to fit and show as many objects as the space permits. As the number of objects increases, so does the number of hidden or occluded objects. Design must strike a balance between number of objects and their inevitable occlusions. This balance should respect the visual stability of the structure; objects should always be found at the same position in space in order to support way-finding.

The aim was to support users acquiring a mental 3D map of the collection akin to acquiring a 2D mental map of an urban center. Lynch (1960) showed the importance of stable and conspicuous landmarks for orienting in urban spaces. Interactive object layout algorithms are known to suffer from instability issues (see review by Herman, Melancon & Marshall, 2000); in other words, the visual placement of the objects changes over time. This is especially troubling because "stability is a very important aspect of interactive layout algorithms" (Lee, *et al.*, 2006, p. 1416).

The issue of efficient usage of screen space concerns visual layout to maximize the number of objects shown in a usable and pleasing manner. This issue was reported by Robertson *et al.* (1991) stating that Cone trees became ineffective when the branching factor reached 30; in other words, when parents can have 30 child nodes or more. The branching factor (i.e., the number of children per subject) is a strong indicator of hierarchy complexity (i.e., breadth and depth): a high branching factor reduces the number of levels which can be drawn in a finite space. Carriere *et al.* (1995) stated Cone trees tended to lose their efficacy for hierarchies exceeding 1000 nodes, they provided methods to increase this upper limit and lessen occlusion of node labels. Fekete *et al.* (2003) specifically addressed the scalability of IV applications by visualizing up to a million units of information.

6.1.3 Object Occlusions

Reducing visual occlusion (i.e., partially hidden objects) is especially critical for the legibility of textual labels. Placing multiple objects in 3D space necessarily produces overlaps which must be managed (Shneiderman & Plaisant, 2005, p. 241). Wang et al. (2006) reported link occlusions by labels in OntoTrack (Liebig & Noppens, 2004), Kleiberg *et al.* (2001) suggested occlusion in 3D environments were an "inherent problem" (p. 91), and Chalmers (1993) stated that some "strongly 3D structures" suffered from "many occlusions and obstructions of view" (p. 384).

Occluded objects "appear invisible to the user" (Teyseyre & Campo, 2009, p. 90) and the information these objects carry cannot easily reach their audience. Dense information displays can overwhelm users (Ware, Purchase, Colpoys & McGill, 2002) because visual objects "may be very close together or occlude each other, and the links may cross one another" (Lee, et al., 2006, p. 1414). These issues have been reported by others (Akrivi, *et al.*, 2007; Freitas, *et al.*, 2002; Shneiderman & Plaisant, 2005) but solutions are contextual; the specific shape of the hierarchy (i.e., breadth and depth) favors specific approaches.

Occlusion reduction in IV interfaces often means reducing overlaps in order to display more objects in a single view; for example, fsvis (Carriere & Kazman, 1995) is a variation of Cone trees "which can visualise about 5000 nodes of a hierarchy without occlusion" (Song, *et al.*, 2004, p. 21). This number is significant but these nodes were nameless abstract visual objects; they had no descriptive textual labels. Interestingly, most occlusion reduction techniques "attempt to show the entire overview of the graph (...) labels are usually ignored" (Lee, et al., 2006, p. 1414).

Placing labels in 2D or 3D environments is an unresolved issue. There has been much attention paid to labelling and label placement in 2D environments (Ali, Hartmann & Strothotte, 2005; Edmondson, Christensen, Marks & Shieber, 1996; Fekete & Plaisant, 1999; Foote & Thomas, 2005). There is work with dynamic labelling of interactive 3D medical illustrations (Hartmann, Ali & Strothotte, 2004; Ritter, Sonnet, Hartmann & Strothotte, 2003). Much of this research centers around displaying the appropriate label for the objects most likely to be of interest for a specific viewpoint, or placing as many labels as possible with minimal overlaps while maintaining a clear link between a label and the object to which it belongs.

Label placement in 2D IV has been an ongoing research question since the late 1990s' (for example: Fekete & Plaisant, 1999; Li, Plaisant & Shneiderman, 1998). This type of research aims to maximize the number of visible labels while minimizing their overlap. For example, Fekete *et al.* (1999) developed a technique which dynamically revealed the labels found in the neighbourhood of the mouse cursor. This allowed showing fewer labels by default and more around the areas the user specifically wished to inspect. An interactive 3D environment inevitably contains label overlaps which are in fact useful to convey relative label position using stereoscopy.

Research concerning dynamic labelling for interactive 3D environments has received some attention for applications in medical imaging (Hartmann, et al., 2004; Ritter, et al., 2003), and general 3D scene and object labelling (Cmolík & Bittner; Maass & Döllner, 2006; Stein & Décoret, 2008). These studies are informative but they do not yet scale to issues presented by SE-3D containing hundreds of densely overlapping labels; in fact, all of them study small numbers of labels (20-30 at most). Higher numbers of labels are studied in the field of digital cartography (Peterson, S. D., Axholt, M., Cooper, M. & Ellis, S., 2009a; Peterson, S. D., Axholt, M., Cooper, M. & Ellis, S. R., 2009b) but, contrary to SE-3D, their labelled objects are essentially contained on a single 2D plane.

6.1.4 Initial Viewpoint, Colors

3D application memorability is dependent on the viewpoints provided to the user since "spatial memory for scenes is viewpoint specific" (Ware, 2008, p. 109). Users of interactive 3D applications can theoretically choose any viewpoint; however, design can ensure certain viewpoints are easily attained and refined to be pleasing and memorable.

The aesthetic quality of SE-3D viewpoints was influenced by the choice of colors. Color theory is a vast domain (Bleicher, 2005; *Colour Source Book*, 2006; Holzschlag, 2001) often associated with graphics design and fine arts. These domains are beyond the scope of this research; however, design is a broad and necessarily applied discipline which requires an informed choice of color.

6.1.5 Navigation Controls

SE-3D was a simulated 3D world where "navigation is often the primary task (...) and refers to the activity of moving through it" (Teyseyre & Campo, 2009, p. 95). Essential to facilitating navigation are *afforded actions* (Raubal & Worboys, 1999) which suggest actions that can be performed with the 3D objects. Users were meant to explore the 3D space and "interact with data. However, they may get lost in the virtual world" (Le Grand & Soto, 2006, p. 76). Design of navigation features should include predefined navigation paths and multiple levels of detail. These must be tested and refined with test users.

6.1.6 Integration of Browse and Search

A review of existing IV applications (see section 4.2) showed that most existing hierarchy applications provided visual interaction with the organizing subject structure with little or no integration with the information itself. A loose integration of subject structure browsing and keyboard searching features creates two largely independent tools. As shown in section 3.1.2 (Integrate Search and Browsing), these tactics are not performed independently—searchers often switch between the two in a single search episode. A search tool which facilitates switching between tactics may be beneficial to information searchers.

6.1.7 Formative Testing

Usable software is produced by testing with users early and often throughout the development lifecycle (Shneiderman & Plaisant, 2005). Nielsen (1993) states that formative evaluation "is done in order to help improve the interface as part of an iterative design process" (p. 170). A completely novel software such as SE-3D required adequate development for test users to learn and practice using a functional prototype; indeed, it was difficult to imagine how users could test paper mock-ups of a 3D interactive abstract world. Nielsen (1993) showed that three test users produce the greatest number of usability issues found for the least cost.

An interactive data-centric 3D application development project contains multiple parts. These include technology constraints, definitions of the 3D world or scene, the objects placed within the world and their behaviours, the colors and materials of the objects, the navigation controls, and the integration of keyword searching.

6.2 Technology constraints

SE-3D was developed as a desktop virtual reality application. It ran on a PC without special visual equipment (e.g., 3D eyewear); however, depth is simulated on a 2D screen and must be inferred by the user. This is the same type of technology used by current 3D games which formed the technological framework for SE-3D development. This framework included hardware graphics capabilities, logic programming (i.e., programming language) and choice of graphics creation and manipulation library (e.g., DirectX, OpenGL).

SE-3D was built over a period of 18 months using Microsoft Visual Studio 2008 and the TrueVision3D[1] (TV3D) game development library. TV3D is an abstraction layer specifically designed for 3D game development based on the Microsoft DirectX graphics library. Abstraction layers facilitate and accelerate the development process and TV3D was chosen because it had a significant usage history, a broad and active user community, and a low purchase price. This technology framework provided a rapid development

[1] http://www.truevision3d.com/ Last Accessed: 17/09/10

platform, and a large user base with ample support. These choices meant that SE-3D could only be deployed as a client on a Microsoft Windows based PC.

A platform independent Web based application is always the favoured deployment method but the technology is not currently available for interactive 3D applications. This is due to the asynchronous nature of online communications which does not allow fluid interactive motion; as a result, online multi-user 3D games (e.g., Unreal[1]) require players to first install a local client, and Web based 3D requires a third party browser plug-in (e.g., Cosmo Player VRML Plugin[2]).

6.3 Objects

In terms of specific design tasks, SE-3D was a world populated with three kinds of objects:

1. **Subject Maps**: containers for direct narrower subjects of a parent subject.
2. **Subjects**: containers for information items whose semantic content has been labelled by a subject term
3. **Relations**: links between a parent subject and its narrower subjects.

Based on existing IV prototypes, basic geometrical shapes were associated with each of these objects. Cone trees used a 2D rectangle as the representation of a hierarchy node; likewise, SE-3D subject maps were designed as 2D squares. Each of these 2D squares could act as a 2D spatial display for a collection of child subject circles (see for example Grokker, Fluit, et al., 2003; Modjeska & Waterworth, 2000) with meaningful placement. Hierarchical links between objects were represented as lines in Cone trees, FSN (Tesler & Strasnick, 1992) and MeSHBrowse (Korn & Shneiderman, 1995).

This visual structure had little meaning without textual subject labels representing the semantic content of the collection. Similar to Wang *et al.* (2006), subject circles were assigned a floating label and their radii was proportional to the relative number of items

[1] http://www.unrealtournament.com/ Last Accessed: 23/09/10
[2] http://cic.nist.gov/vrml/cosmoplayer.html Last Accessed: 17/09/10

assigned to the subject. Each subject map offered a floating label containing the broad parent whose narrower children could be found on its 2D plane.

As shown in Figure 6.1 and Figure 6.2, viewed from a distance, SE-3D was a combination of small squares linked via lines, spreading wider as the semantic contents become more specific i.e. towards the bottom. The individual subject circles placed on each subject map were discernable only when the user approached a subject map of interest—an implicit overview and zoom-in functionality.

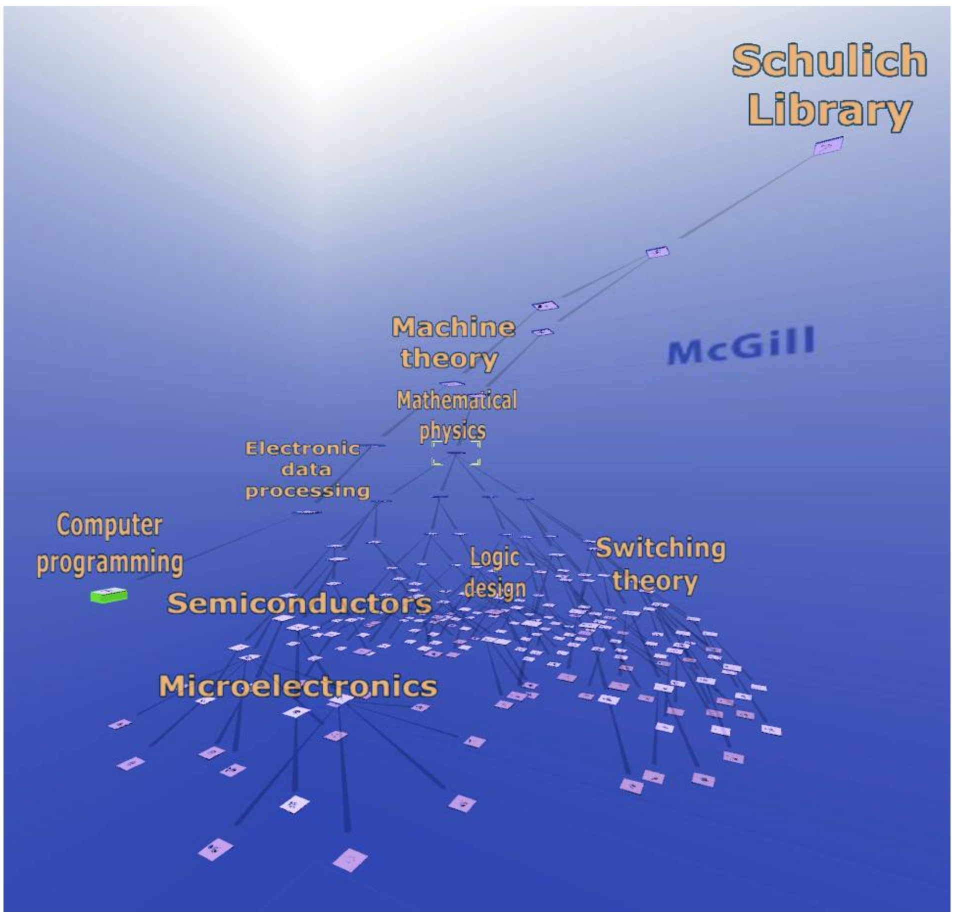


Figure 6.1: SE-3D structure overview example 1

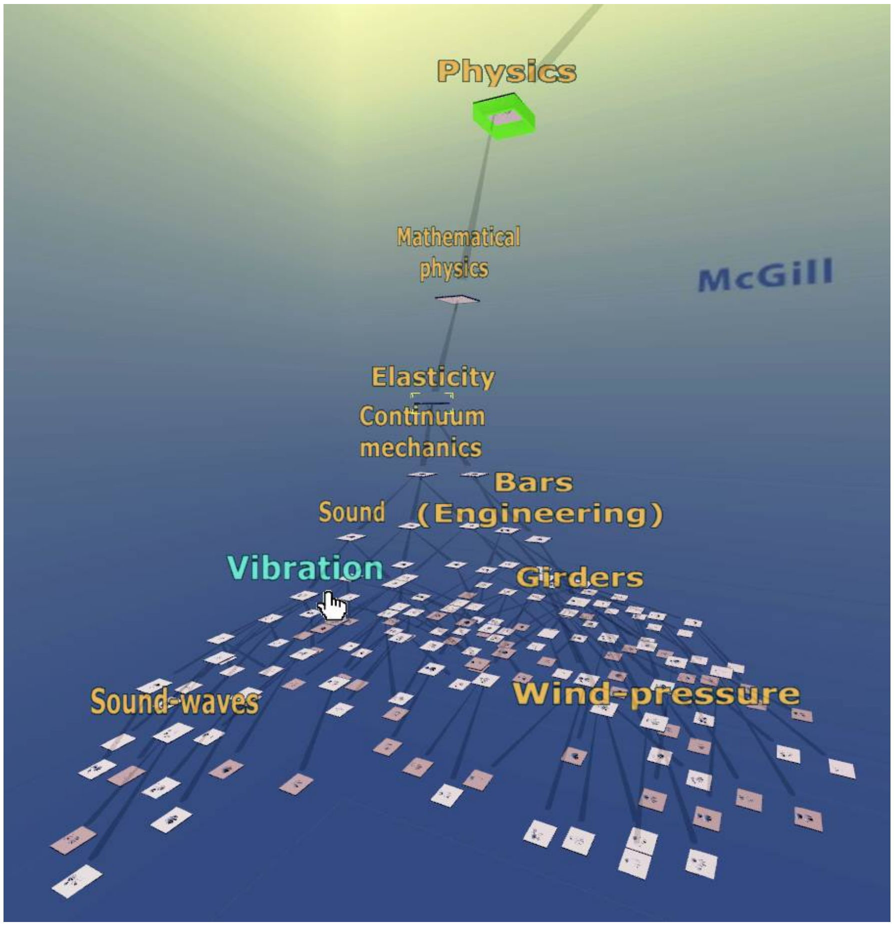


Figure 6.2: SE-3D structure overview example 2

Each subject map and its label was always placed at the same position in space; therefore, known subject labels acted as landmarks "to help users to orient in a 3D world" (Teyseyre & Campo, 2009, p. 90). Controlled vocabulary terms were also used as visual landmarks by Fowler *et al.* (1991) to "supply information about both the content and structure of the database" (p. 147).

6.3.1 Object Positions

The basic elements of the SE-3D world were placed incrementally to represent a larger proportion of the subject hierarchy; as a result, the object positioning logic became critical. The number of objects grew exponentially even in this simplified and reconstructed LCSH hierarchy containing over 1200 maps, the links between them, and more than 4300 subject circles. It was impossible to display all these objects in a limited amount of 3D space. Too many objects would inevitably hide others and conflicting visual elements created an unpleasant jumbled visual mess.

SE-3D aimed for a meaningful placement of objects and the following object positioning objectives and constraints were developed:

- show as many subject maps as possible in space
- attach meaning to the placement of the subject maps in space
- attach meaning to the placement of subject circles on each map
- ensure subject labels legibility by reducing label occlusions
- ensure the stability of the visual layout over time

The development of software features to meet these requirements is described in the following section.

6.3.1.1 Map Positioning

Cone trees and its descendents positioned child nodes around the parent in a circular layout using an arbitrary order (e.g., alphabetical), and the radii of the layout circle was proportional to the number of children to be placed. This produced a stable but meaningless layout—the user was given no indication of the relative importance of each child and had to inspect them all to choose between them.

The branching factor provides an indication of the size and complexity of a hierarchy to be displayed. The tested collection offered a reconstructed LCSH hierarchy with a total of 622 prevalent subjects having at least one child subject. Figure 6.3 shows the cumulative distribution of the number of child subjects per broader parent. It reveals that the vast majority (83%) of parent subjects had four children or less, and that limiting the branching factor to eight would still accurately represent 97% the vast majority of child relationships.

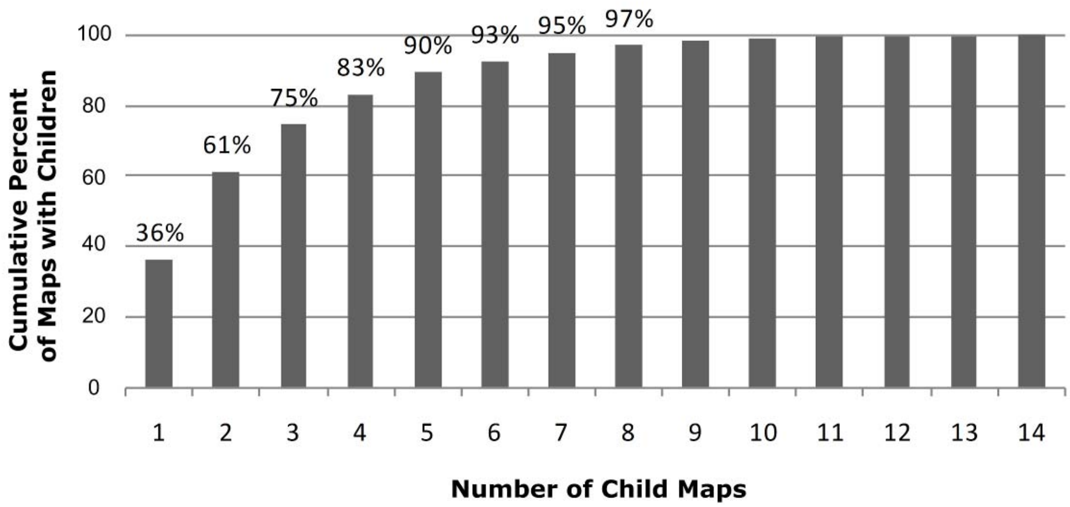


Figure 6.3: Cumulative distribution of child subjects per parent

SE-3D placement of child subjects suggested the relative importance of each sub branch based on the number of assigned bibliographic items. For each parent subject with at least one child map, the most prominent sub branch was always placed diagonally downwards and west, the second most prominent was then placed east, followed by south and north. Any remaining child branches were placed between the first four in a fixed order. In this manner, travelling towards the west wall of the world (i.e., the left wall) always meant going towards the most prominent branches, the east wall was always the direction of the second most prominent branch, and so forth.

This layout logic provided static and meaningful layout; in addition, it had the benefit of clearly isolating the largest subject branches. Potential branch collisions were greatly reduced since the two largest branches were spaced furthest apart from each other, and, thanks to the Bradford's Law distribution, each additional branch was exponentially smaller.

Applying meaning to object placement did not necessarily mean users would perceive or understand it; however, it was a step beyond arbitrary. The hope was that with training and experience some users might learn to exploit the information communicated via the placement; this would require further testing beyond the scope of this research. Ultimately, advanced users might want to choose the placement logic.

6.3.1.2 Circle positioning

SE-3D used one dimension to represent father-child relationships between subjects (the vertical or up/down dimension of the world) which left two dimensions to communicate other relations. Many prototypes and commercial applications are strictly 2D (see section 3.2.6–Spatial Displays); many could be applied to the surface of each subject map in SE-3D. This meant each set of children could have its own 2D spatial display. This section refers strictly to the issue of placement of subject circles on each subject map as shown in Figure 6.9.

As stated in section 4.3.2, SE-3D aimed to reveal the connections between child subjects based on term co-occurrence; in other words, the strength of the connection between two subject terms would be proportional to the number of times the terms had been assigned as a pair to the same bibliographic record. This approach was suggested by Bates (2003) and applied by Fowler *et al.* (1991) in their innovative query system.

Given the strength of the connection between each subject, various methods are available for the placement of the objects on the 2D map. Minimum Spanning Tree (MST) (Graham & Hell, 1985) is a method which chooses to represent only connections beyond a certain strength or threshold value. Pathfinder Networks (PFNETs) (Schvaneveldt, Dearholt & Durso, 1988) liberally prunes all connections except the one having the highest strength (White, 2003). Self-Organizing Maps (SOMs) (Kohonen, 1997) use spatial proximity to communicate the strength of the relation between terms. Each of these methods has its strength and weaknesses; for example,

- Buzydlowski (2003) found that SOMs outperformed PFNETs in representing "the mental models of 20 experts in selected fields of the humanities" (White, et al., 2004, p. 5301)
- Fowler *et al.* (1991) described PFNETs as an alternative to MST threshold networks (p. 145)
- Cribbin *et al.* (2001) found no significant difference when comparing MST and PFNETs for IR tasks. Both were outperformed by a text only list interface.

An alternative is the force-directed algorithm (Di Battista, 1999) which is very popular because it is simple and easy to understand (Lee, et al., 2006, p. 1415). These types of algorithms are slow and produce a different visual layout solution each time they are invoked (Julien & Cole, 2009; Lee, et al., 2006), and labelled nodes suffer from occlusions (Gansner & North, 1998). These shortcomings were not deciding factors in the case of SE-3D since 1) processing of the solutions would be done in batch pre-processing with no impact on the performance of the production system, 2) each solution would be generated once and stored resulting in a static visual representation, and 3) labels in 3D float above their objects and do not require space on the 2D surface

A force-directed layout was chosen because it is relatively easy for end users to understand—the closer the subject circles, the more they are *connected*. The exact nature of that connection may not be known *a priori* by the user but there may be value in suggesting which subjects are more likely to be read together within the same information items (see section 4.3.4–SE-3D Support in the Searching Process).

Using the procedure reported in Julien *et al.* (2009), force-directed layouts were tested for a subset of the subject maps. The results were mixed. Maps with few subjects tended to produce pleasing results with a clear portrayal of connection strengths between subjects; on the other hand, maps with many subjects, perhaps more than 20, often resembled a chaotic mess of randomly placed subject circles.

These dense subject maps were difficult to force-direct as the available 2D surface area became scarce causing multiple collisions. The messy maps were usability concerns and the algorithm would require additional research and testing beyond the scope of this research; as a result, force-directed layout was abandoned.

6.3.1.2.1. Alternative Subject Circle Placement

A replacement subject circle placement strategy focused on 1) clear visual communication of relative subject importance, 2) consistent layout between subject maps, and 3) uniform usage of available 2D space. Specifically, the subject circle layout algorithm followed the following steps: for each subject map, for each subject circle in decreasing order of prominence (i.e., from the subject containing the most bibliographic records down to the least):

1. Place the first and largest subject (i.e., the circle with the largest radius) in the middle of the 2D subject map
2. Place the second to the left of the first with a fixed space between them
3. Fan out the rest of the circles counter clock-wise along radius equal to the space between the centers of the first two subject circles.

This simple algorithm consistently produced subject map layouts as shown in Figure 6.9. The layouts were deemed relatively pleasing and easy to understand; however, they did not convey the subject co-occurrence frequency. As shown in this section, the issue of effective 2D layout algorithms is an open one—it would be an interesting subject for further SE-3D development and testing.

6.3.1.3 Object Occlusion

The SE-3D label placement issue was related to dynamic labelling of interactive 3D worlds which has received little attention (notable exceptions: Maass & Döllner, 2006; Stein & Décoret, 2008). The simplified subject structure contained thousands of subjects labels—too many for both human and hardware capacity. Label filtering mechanisms had to be developed.

6.3.1.3.1. Limiting the number of visible objects

Ware (2008) stated that minimizing occlusions usually "means that overall depth in the scene should be limited" (p. 95). This was applied to SE-3D by constraining the depth of the hierarchy shown. This had the added benefit of reducing hardware requirements; in fact, graphics hardware capabilities imposed an upper limit of objects the virtual world could contain. Failure to respect hardware capabilities would result in a loss of movement fluidity (i.e., choppiness) as the user navigated through the 3D world.

Performance tests showed that even with some of the most powerful hardware available no more than five levels of the hierarchy could be fully expanded while maintaining fluid motion. This depth limit is likely contextual depending on the specific breadth and depth of the hierarchy used.

The SE-3D branch expansion required a user-selected subject map and performed the following actions:

- **Indicate the current center map**: a floating yellow frame was drawn around the current center map (i.e., last subject map selected by the user) as shown around "Physics" in Figure 6.4.

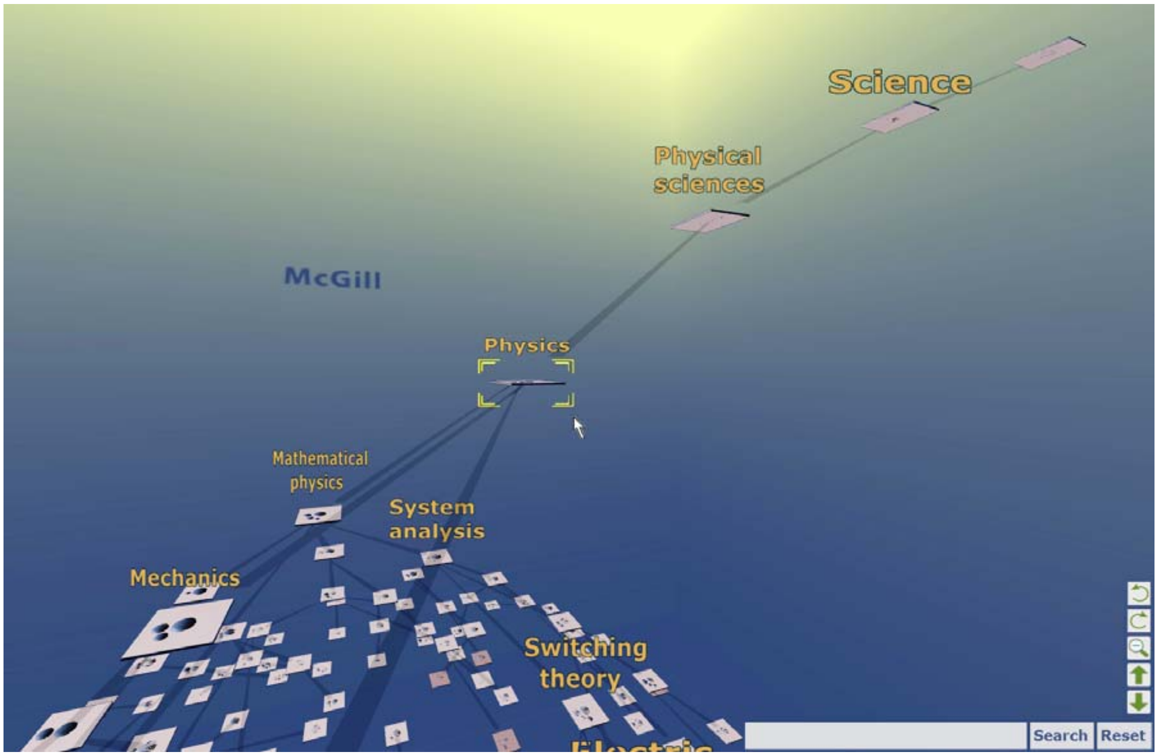


Figure 6.4: Yellow frame indicating current center map

- **Reveal path to top**: starting from the center map, show the broader subject maps until the top of subject hierarchy is reached. This provided a consistent path towards broader subjects and the collection overview they provide. This path leading back to the *top* also acted as a way-finding feature.
- **Expand all branches five levels down**: starting from the center map, expand all branches for no more than five levels away from the center map level. Depending on the breadth and depth of the subject hierarchy in a specific area, this could mean the structure was fully expanded.

Fully expanded branches terminated with their most specific subject (often referred to as a *leaf* node in graph drawing). These end points provided users with the assurance that the structure was not infinite and a searcher could fully explore an entire subject area. As shown in Figure 6.5, these semantic end points or leaf nodes were visually indicated by a darker shade of gray.

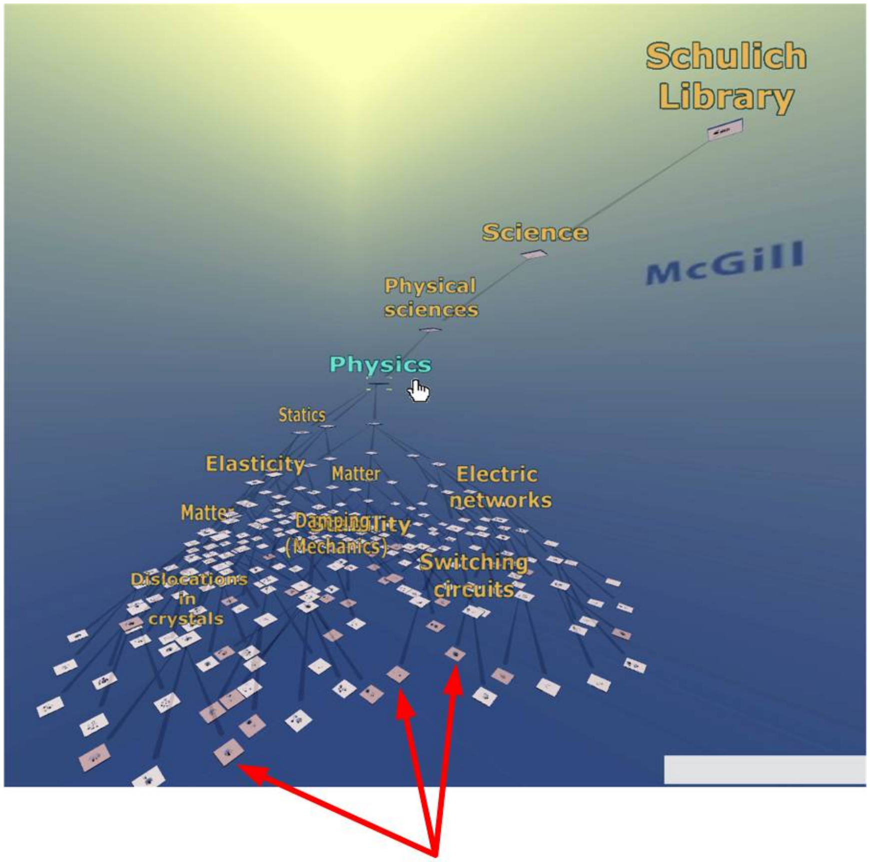


Figure 6.5: Darker subject maps indicate most specific subjects

By iteratively selecting subject maps of interest, the searcher drilled down towards smaller and more specific portions of the collection as shown in Figure 6.6. It shows that the user has selected "Elasticity" and its child branches are almost all fully expanded.

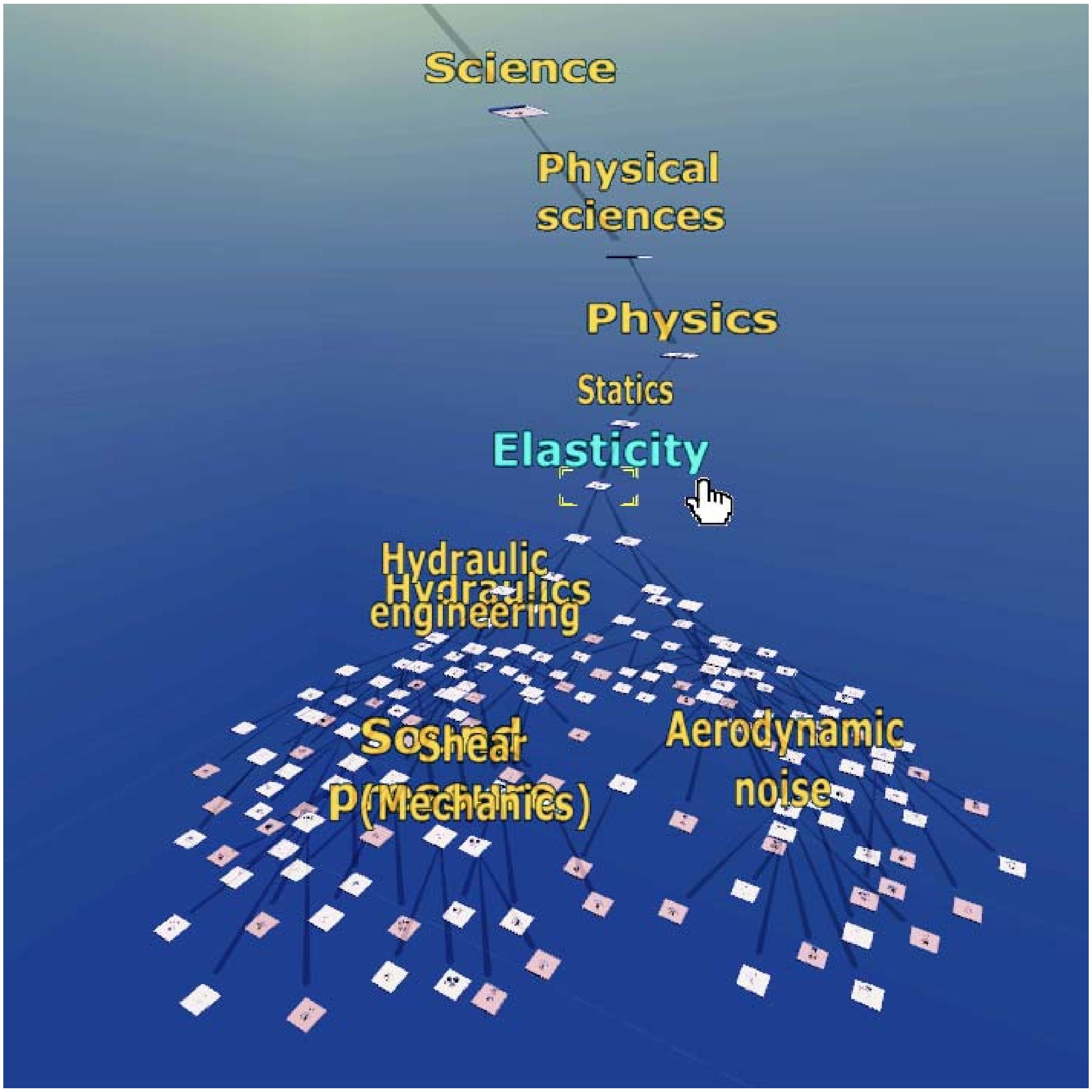


Figure 6.6: Progressive selection of more specific subjects

By dynamically limiting the visible depth of the subject hierarchy, the design established a compromise between showing as much of the hierarchy as possible and providing fluid navigation; however, there were too many subject labels to be shown and their occlusion was still a problem.

6.3.1.3.2. Dynamic Labels

The numerous overlapping subject maps shown in Figure 6.5 and Figure 6.6 hint at the difficulty of presenting their individual labels; indeed, showing all labels produced massive cloud of illegible words. Kleiberg *et al.* (2001) suggest that occlusion in 3D worlds can be "overcome by interaction on the model" (p. 91). As the user navigates to change his/her viewpoint, legible labels can dynamically appear while occluded ones disappear. For a specific viewpoint, SE-3D would draw only the fully visible non-occluded labels. This behaviour is illustrated in Figure 6.7 which shows the subject structure centered on "Cybernetics".

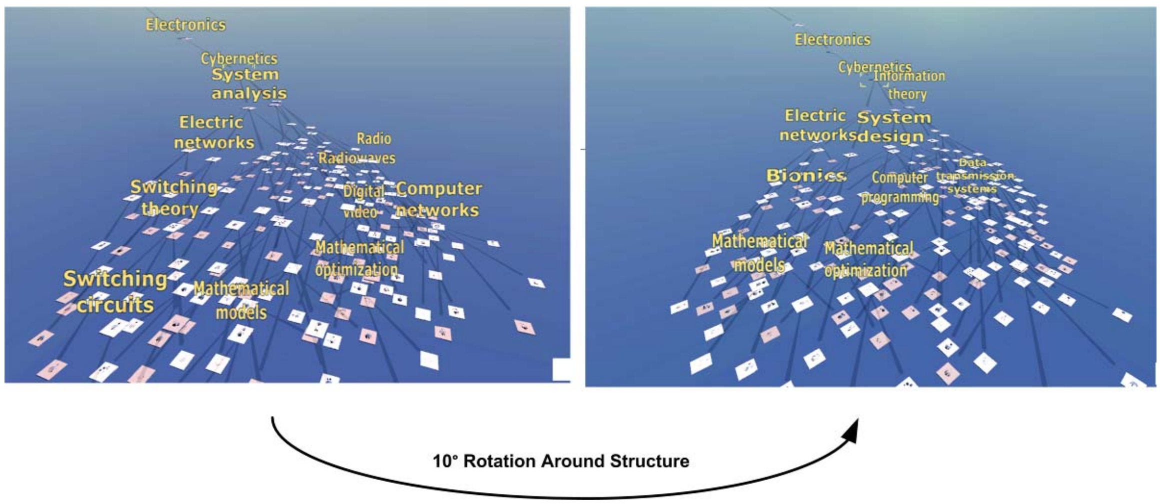


Figure 6.7: Dynamic labeling as the user changes the viewpoint

Figure 6.7 shows labels for an initial viewpoint on the left. A slight 10° counter-clockwise rotation around the structure is performed. This effectively hides some labels (e.g., "Switching Theory" is no longer visible on the right) and reveals labels newly visible from the modified viewpoint (e.g., "System Design" is now visible on the right).

The label management approach developed for SE-3D is novel and was the best of the tested alternatives; however, it had drawbacks. Firstly, labels placed deep inside the center of a large and broad structure may never be visible as the user navigated around the whole structure. These labels were effectively occluded from every viewpoint; however, they could be revealed by manually selecting and isolating the subject branch, or navigating within the structure. This is an extra interaction many users may never bother to choose. Secondly, Peterson, Axholt, Cooper & Ellis (2009a) reported that label movement disturbed their test users, and the unstable nature of the dynamic labelling could be distracting. SE-3D labels could appear to flicker on and off as the user quickly moved around the structure; this sometimes created a distracting *blinking label affect*.

Mitigation of this issue could take the form of an animated fade-in/fade-out of the labels, but this would be the subject of future development and testing.

6.4 World View

The first impressions created by a 3D application depends on the initial view the user is given. In the case of SE-3D, the colors, lighting, and interactive objects shown in the world aimed to produce a pleasing and calm atmosphere conducive to exploration. Figure 6.8 presents the initial *establishing shot* (Ware 2008 p. 142) presented to the user which establishes *gist* and the overall positions of the objects.

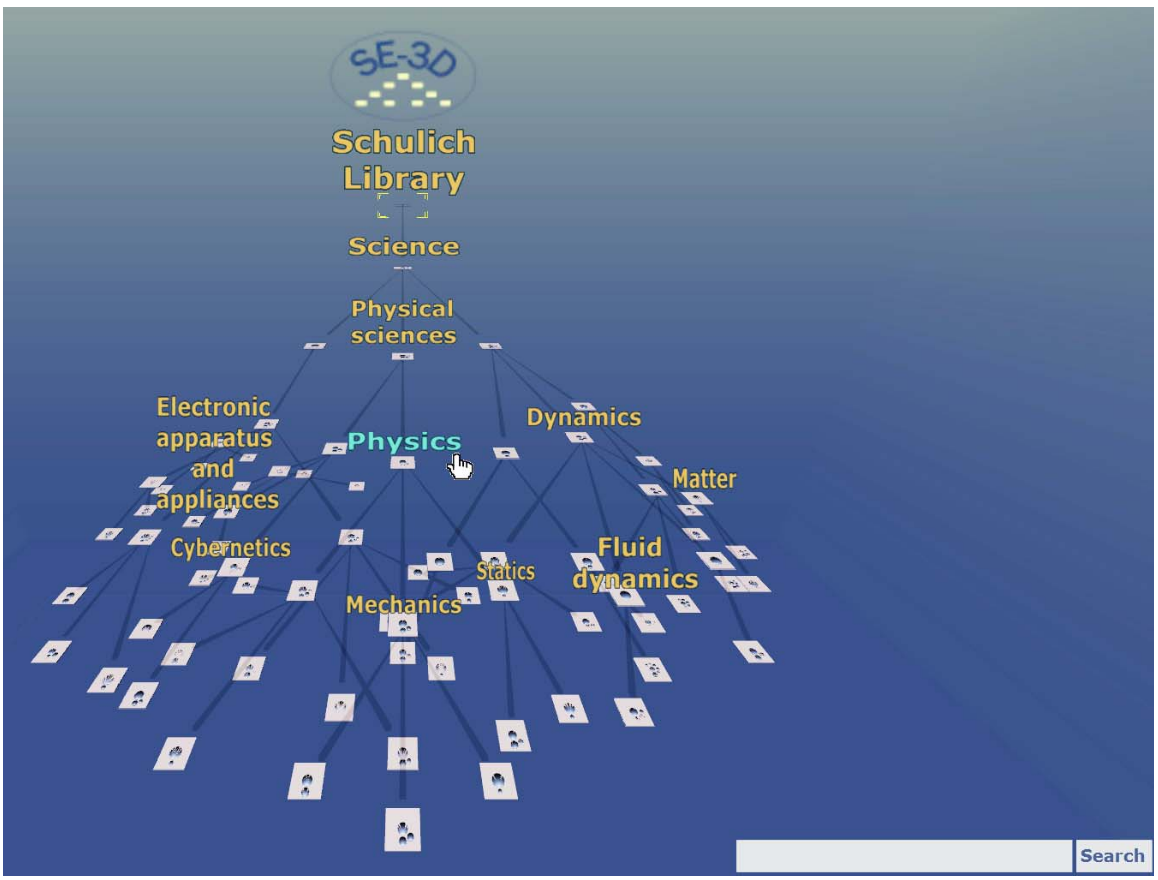


Figure 6.8: SE-3D initial world view

This initial view set the tone for the ensuing interaction. It offered at a glance all the types of objects that existed in the world: the subject labels hovering over objects which led to other subjects. A keyword search box was always visible at the bottom right of the display.

A light source acting as the sun stemmed from the top at a 70° angle based on top-down lighting bias (Lidwell, et al., 2003, p. 196). Light should cast shadows and their absence is conspicuous. Shadows also provide additional depth cues which make them valuable in a simulated 3D as shown around the circles in Figure 6.9.

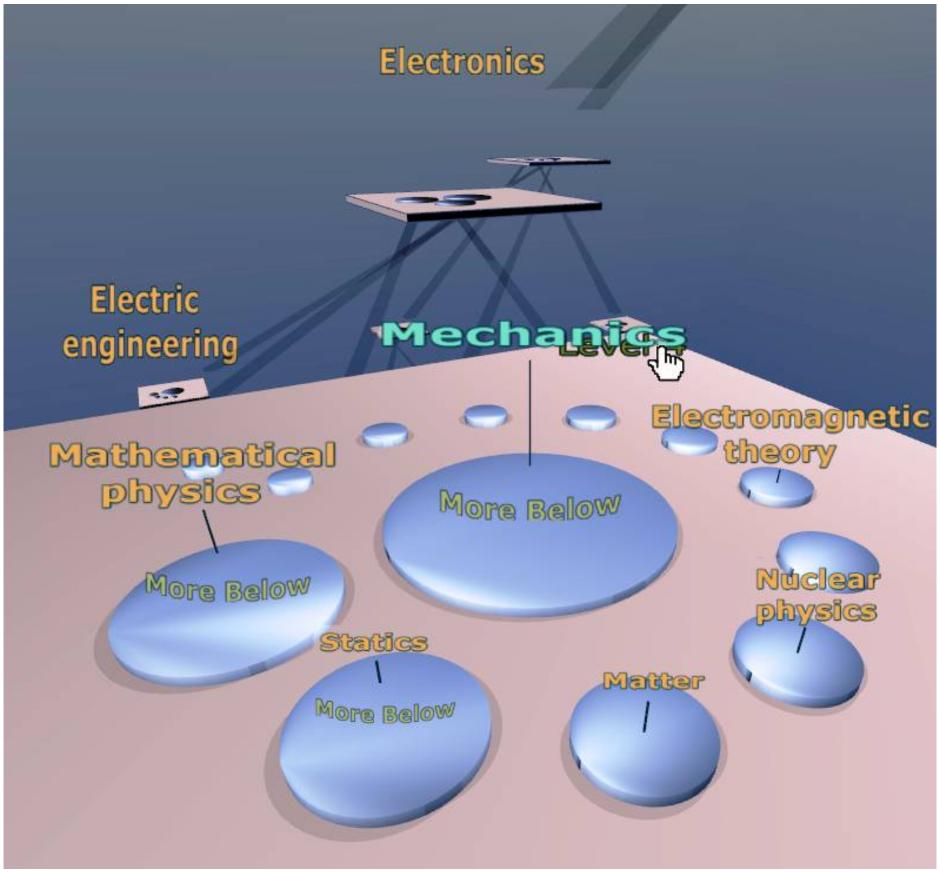


Figure 6.9: Shadows in SE-3D

6.4.1 Colors

There are few design guidelines to support color choice; however, there are many examples from which to draw inspiration. Color palettes were collected from existing coloured visual graphics designs (Bleicher, 2005; Carter, 2006; *Colour Source Book*, 2006; Holzschlag, 2001). The palettes had to convey a desired mood while offering an adequate number of individual colors. SE-3D required distinct colors for the following objects:

1. The square subject maps
2. The individual round subjects

3. The subject labels
4. The links between subjects
5. Interface components
    - buttons, search box, and results list
6. World background
7. Status colors
    - subject matching a search, hyperlink in result list, hovered subject, previously visited subjects, last clicked subject.

The number of these color requirements was acceptable based on a review by Ware (2008) showing "the number of colors recommended is always between 6 and 12" (p. 77).

Carter (2006) provided a large number of examples from the world of advertizing and publishing. One of these presented an illustrated paper calendar (p. 305) which contained eight distinct colors. This palette was chosen since, like SE-3D, a calendar is an information presentation artefact which contains textual elements. The palette mood was warm and comforting based on shades of red, orange, yellow, blue and a neutral pinkish gray.

The neutral gray was assigned to the subject maps as they would be the predominant reoccurring object in the structure and they were containers for other objects (i.e., the subject circles). A soothing light blue was chosen for individual subject circles and a shine was applied to their surface to make them clearly stand out from the neutral gray on which they were placed. Subject labels were assigned a light orange-yellow color to maximize contrast. Interface buttons were drawn as light green on neutral gray, and text displays were high contrast as white on very dark blue background.

The links between the subjects were initially drawn as translucent or semi-transparent cone shaped tunnels stemming from a parent subject circle extending towards the child maps. Translucency effects were difficult to manage by graphics cards and these links produced much visual overlap and occlusion. Repeated trials showed that a much thinner cone almost resembling a long thin branch produced an appropriate suggestion of relationship.

6.4.1.1 World Background

The choice of world background colors was not obvious and changed numerous times. As with most 3D computer designs, the world was constrained inside a six face cube: one for the top or sky, one for the bottom or ground, and four faces for each of the sides (i.e., left, front, right, back). A different color or image could have been used for each face but the transition between them (i.e., the edges where the faces meet) had to be perfect otherwise the eye would be automatically drawn to this visual dissonance. Applying a black background created a strange scene of outer space without stars which provided no indication of directionality. Users could not tell if they were travelling up, down, forward or back because there was no conspicuous spatial referent (see section 3.2.7 Virtual Reality for IV). This issue has been reported by Chalmers (1993) yet is still found in recent IV prototypes (Perez & de Antonio, 2004; Westerman & Cribbin, 2000).

The top face of the world was intuitively a sky from which light would be emitted; this suggested a bright color such as sky blue or warm yellow. The bottom face would suggest progression towards deeper and more specific levels of the semantic hierarchy; a dark almost black color. These choices provided users with referents when travelling up and down the structure.

The sides of the world should provide directional reference without adding conflicting visual elements with the semantic hierarchy. Early attempts were made with various landscape scenes (e.g., mountains on the left, plains on the right, etc.) but this was option was discarded because of visual overload and difficult transitions between scenes. The visual neutrality of the world sides was important; thus, a gradient transition from the dark bottom color to the top sky color was applied to all faces. The resulting world shown in Figure 6.8 was made up of a very dark blue bottom, a bright yellow-green top sky and gradient sides acting as transitions between those two colors.

6.4.1.2 Status Changes

Interactive software is necessarily dynamic and objects change with user commands and resulting system actions. Changes of object color or shape act as action confirmation from the system and remind the user of actions performed in the past. A classic example is the Web page hyperlink which changes color once it has been clicked; this reminds Web surfers that they have previously followed that specific link to another Web page.

Five status changes were signified to the user in the following manner:

- **A subject matches a search**: subjects that contained works matching a keyword search were visually identified. This was done by isolating matching subject labels by not drawing non-matching subjects. The remaining matching labels were coloured either in bright orange or red to signify two levels of relevance density. The matching subject circle color was also changed to a bright orange.
- **Previously visited subject map**: a subject map that has been closely inspected for more the three seconds was judged to be visited. This triggered the appearance of a conspicuous bright green border around its square outer edge.
- **Last selected subject map**: the LCSH subject structure was too large to be shown completely; the user had to choose branches to explore in further detail. The last branch selected by the user was identified by a floating bright yellow frame which changed position to match each new branch selection.
- **Hovered subject**: when multiple overlapping labels were shown, inspection of a specific label was facilitated by a visual confirmation of the currently selected subject label. The subject label situated immediately underneath the mouse cursor was redrawn bright turquoise to quickly communicate its currently active status. The system would take no further action unless the user clicked on the label.

Hyperlink in result list: the search result list contained white text on a dark background except for subjects assigned to each work. These were hyperlinks triggering a flight to the specific position of that subject in the visual hierarchy. Guided by the ubiquitous Web navigation paradigm, these textual subject hyperlinks were drawn light blue and underlined.

6.5 Navigation Controls

SE-3D links between subject maps acted as afforded navigation paths. Gestalt theory (see section 3.2.1.2) states that lines between visual objects create a relation between them. Users intuitively follow these lines from one object to the other as they would follow a road between two cities on a travel map. Figure 6.1 and Figure 6.2 show how the dark lines connecting subject maps guide searchers to valid navigation paths.

6.5.1 Levels of Details

There were three levels of detail in SE-3D: 1) the semantic structure overview where only subject maps were visible, 2) the subject map view where all child subjects were visible for a single parent, and 3) a detailed classic ranked textual list of information items. Searchers could inspect either a broad overview of the subject structure based on prominent subjects and their narrower subjects; alternatively, travelling close to a subject map of interest could reveal additional narrower subjects which had no child maps of their own.

The first two levels of detail were representational of the collection distribution within the semantic hierarchy. The third level was the most detailed and allowed on demand inspection of individual information items (see top left of Figure 6.10).

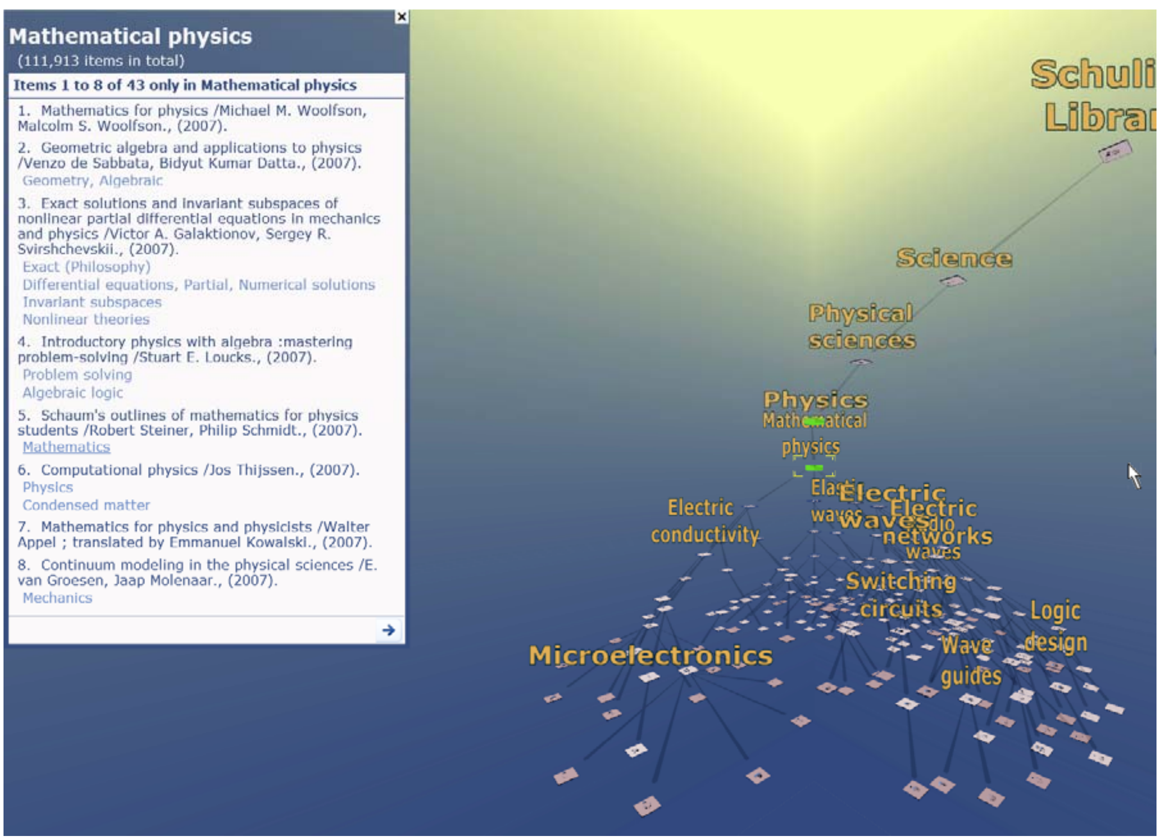


Figure 6.10: Detailed list of individual items in SE-3D

6.5.1.1 Dynamic Semantic Hierarchy Overview

Users could expand a subject branch by single-clicking on a subject map or a specific subject circle. This removed the branches that were not direct descendent of the chosen subject, and revealed additional structure depth, if any, belonging to the selected subject. This action also updated the textual result list with items assigned to the subject. In this manner, the user could interactively navigate up and down the hierarchy while simultaneously inspecting bibliographic items. The interactive branch expansion feature found in SE-3D is illustrated in Figure 6.11 and Figure 6.12.

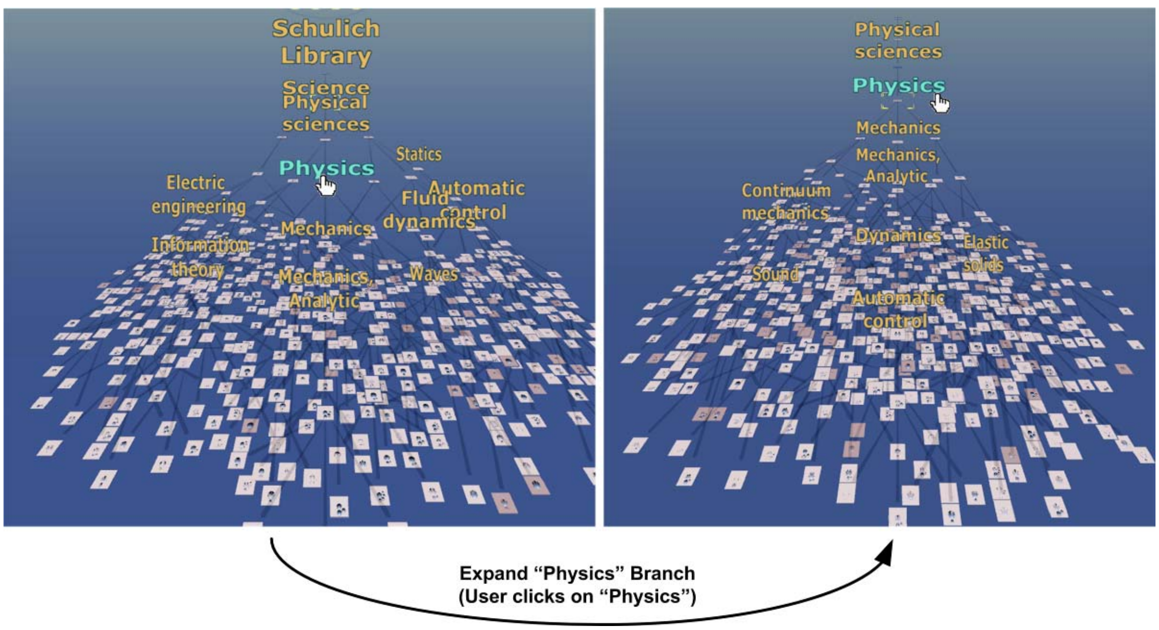


Figure 6.11: Expanding the "Physics" subject branch in SE-3D

The left side of Figure 6.11 shows the user hovering over "Physics", the user clicks the left mouse button to signify the physics branch is of interest and should be further expanded. The result is shown on the right side of Figure 6.11. It shows that branches not belonging to "Physics" have disappeared (e.g., "Statics" and "Electric engineering") and the very broad subject of physics has been further expanded.

This type of interaction becomes more dramatic as the specificity of the expanded branch increases. Continuing the interaction started in Figure 6.11 above, Figure 6.12 shows the user recognizes "Dynamics" as a subject of interest and decides to expand the branch.

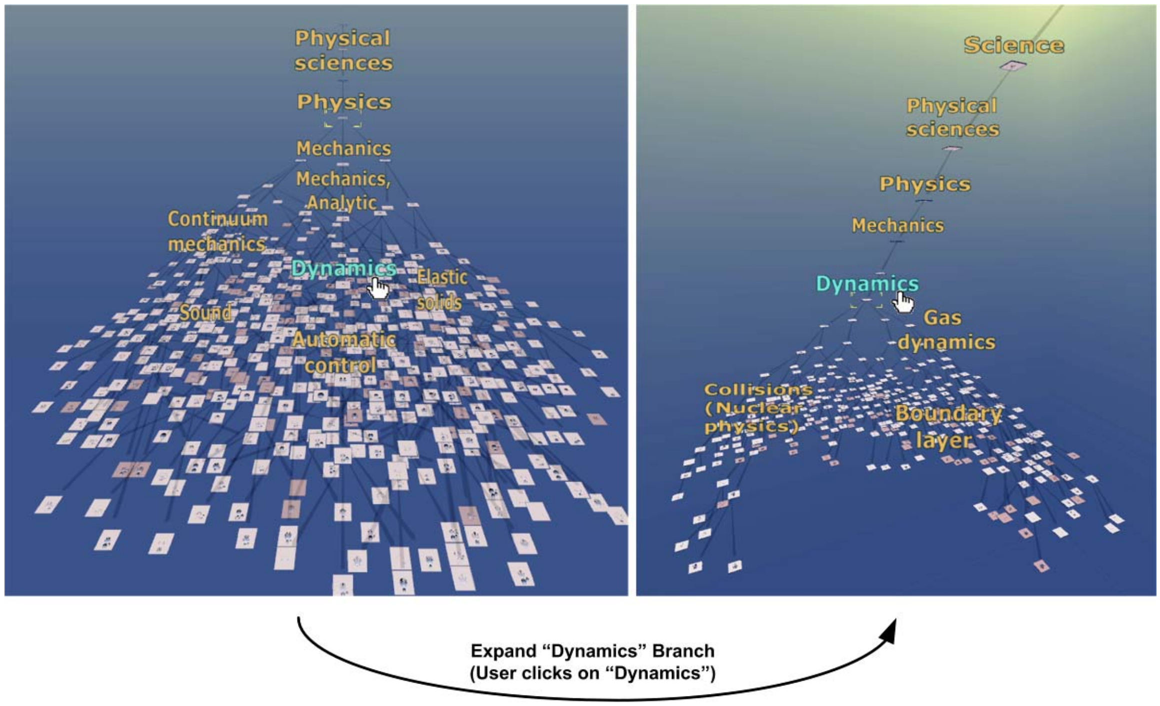


Figure 6.12: Expanding the "Dynamics" subject branch in SE-3D

Figure 6.12 shows a dramatic difference in the shape of the expanded physics branch on the left and the expanded dynamics branch on the right. The latter is obviously more specific and contains a reduced set of child branches which are partly fully expanded (as indicated by their darker shade of gray). The searcher is offered interactive visual cues as to his/her gradual descent into more specific subjects, their proportional narrowing and eventual most specific subject.

Figure 6.1 and Figure 6.2 show examples of broad overviews of the semantic structure. Rotation around these structures reveals their shape in terms of breadth and depth of the subject area. These shapes may allow browsers to remember and recognize known subject areas and hone into unexplored semantic space. Dynamic subject labels (see section 6.3.1.3) automatically associate semantic meaning to these otherwise abstract semantic structure shapes.

Semantic overview shapes may offer an aesthetic memorability which could foster curiosity. Figure 6.13 presents two examples of this visual effect. It shows that a broad subject such as "Mathematics" (left side of Figure 6.13) spreads out into a broad set of deep branches forming a specific shape.

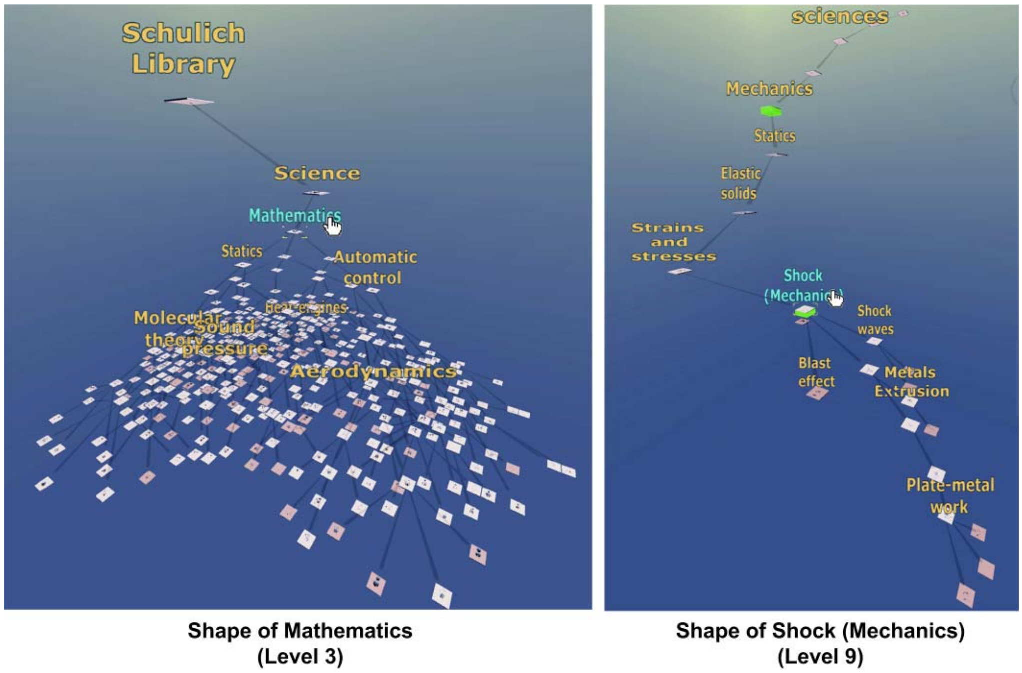


Figure 6.13: Aesthetically pleasing and memorable structure overviews in SE-3D

A more specific subject such as "Shock (Mechanics)" (right side of Figure 6.13) offers a narrow and shallow shape with fully expanded branches. The shape of these known subject branches may eventually be recognized explicitly. Users may also be pleasantly surprised at the pleasing or unusual shapes of newly discovered subject branches. These effects are of interest but beyond the scope of the current research.

The right side of Figure 6.13 also demonstrates the implementation of previously visited subjects. Subject maps which have been closely inspected for at least three seconds were drawn with a strong green border. This allowed users to quickly see where they had been and the remaining unexplored territory. Figure 6.13 shows that "Shock Mechanics" and "Mechanics" have both been closely inspected.

6.5.1.2 Subject Details

Figure 6.14 presents a close up of the "Physics" subject map. It shows that the first three most prominent subjects (e.g., "Mechanics" and "Mathematical physics") lead to more specific subject maps of their own; this is identified by the "More Below" label hovering above them. These three most prominent subjects contained at least 86% of all items assigned to any child branch, and the remainder of the subjects contained an exponentially smaller portion (see section 5.9–Generate the Subject Maps).

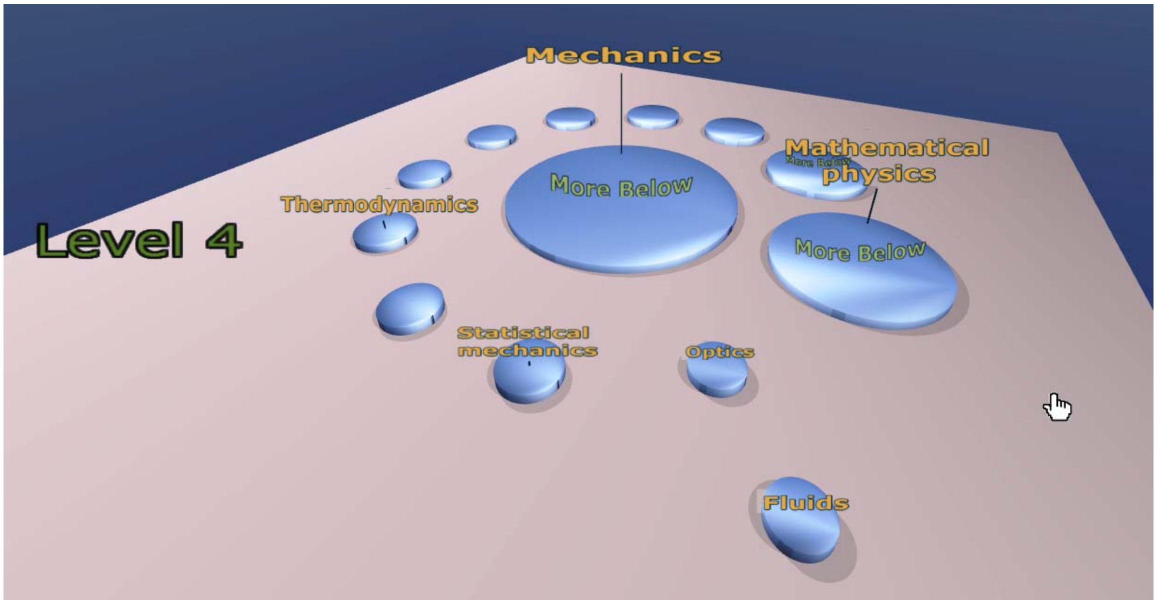


Figure 6.14: Close up of individual subject map in SE-3D

The undivided subjects belonging to physics (e.g., "Optics", "Fluids" shown in Figure 6.14) could be selected to show a list of assigned items but they did not lead to their own narrower subject map. These subjects may have narrower subjects as per LCSH authority; however, for this collection, their relative importance for "Physics" did not warrant further subdivision.

Figure 6.14 also shows that labels floating above the subject circles were placed at different heights and visually linked to their blue circle via a thin black line. The height of the label was proportional to the relative prominence of the subject on the subject map; the more items a subject circle contained, the higher its label was placed. This served as an additional cue to the relative importance of subjects and reduced the chance of label occlusion. The thin black lines clearly anchored each label to its subject circle.

Each subject map also showed a hierarchy depth label (e.g., "Level 4" on the left side of Figure 6.14). All subject maps on a particular level of the hierarchy had the same label. This informed the relative specificity of each subject; for example, a subject on level thirteen would arguably be more specific than a level four subject.

6.5.2 Keyboard and Mouse Commands

Bridging the *gulf of execution* (Norman, 1988) depends on a correct translation of user intentions to a language the system understands. Crossing this HCI hurdle can be facilitated by effective designs of navigational controls (e.g., keyboard commands, screen buttons, mouse interactions, etc.). As shown in section 3.2.1.3.1 (Level of Realism), the mapping choices between keyboard/mouse input and associated system response are often arbitrary since some concepts have no obvious visual representation. They must be learned and practiced.

Movement in a 3D space means two types of often parallel actions: 1) changing the position of the viewer and 2) changing the gaze of the viewer. The analogy of a camera on a tripod is often used in 3D world development; the tripod can be moved in space, and the camera can be rotated on its tripod. Performed in parallel, these actions allow full mobility and visibility in a 3D space.

6.5.2.1 Viewpoint Position and Direction

Changing the camera position in SE-3D permitted movement forward, back, left, right, up and down. The first four were associated with the keyboard arrows which provided movement on a single plane in space. Up and down movement is less intuitive to flightless humans. This meant arbitrary keyboard keys were assigned to vertical movement. After some reflection, the letter "T" would control movement towards the top or up, and the letter "B" towards the bottom or down. Up/down movement was also possible using up and down arrow 2D screen buttons.

Providing intuitive controls for camera gaze direction proved to less obvious. Like a human, a camera can shift its gaze side to side, up and down. Gaze control can be independent from position control. SE-3D used a full 3D world where users were encouraged to explore multiple planes representing multiple depths of the semantic hierarchy.

A gaming analogy is MechWarrior[1] whose characters can independently change their position and shift their gaze (i.e., torso twist); in fact, both actions have a separate set of associated keyboard commands. Learning to integrate these two types of highly associated actions takes some time. SE-3D had to be simple to use and having users learn to use and integrate two sets of commands for movement and gaze shifting would add significant complexity. No solution was obvious.

The initial attempt to control gaze was based on mouse movements. The system would assume the user meant to look towards the direction of the movement of the mouse cursor on screen; for example, moving the cursor towards the right would steadily shift the gaze towards the right. The approach allowed gaze direction control without specific commands; unfortunately, it could produce chaotic gaze movements if the mouse cursor was rapidly moved around the screen. This was not perfect but further enhancements would await user testing.

6.5.2.2 Zooming in and out

Early developer testing suggested that a zoom functionality was necessary. This would be a quick way to move forward (i.e., zoom-in) or back (i.e., zoom-out) while maintaining a fixed gaze direction. These types of functionalities are animated transitions, a form of "automatic camera assistance during the transition phase from one focus object to the other" (Teyseyre & Campo, 2009, p. 91). Animated transitions (see section 3.2.1.3.2) attempt to clearly communicate the temporal change in system status; in this case, the change in camera position and resulting viewpoint.

Early navigation testing by the developer suggested a zoom-in would target a specific subject map, and a zoom-out would be a request for an overview of the currently displayed structure. A zoom towards a specific subject map required the user to 1) indicate the intended target subject map by hovering over it, and 2) press the space bar. This triggered an automatic animated flight towards the targeted subject map. In this manner, the user could successively *jump* from one map to the other by hovering and pressing space.

[1] see http://www.microsoft.com/games/mechwarrior4/

Zoom-in is illustrated by Figure 6.15. showing the user hovering over "Elasticity" (left side of Figure 6.15), the space bar is pressed and the viewpoint is automatically *flown* with an animated transition to a close-up viewpoint of the specific subject map (right side of Figure 6.15).

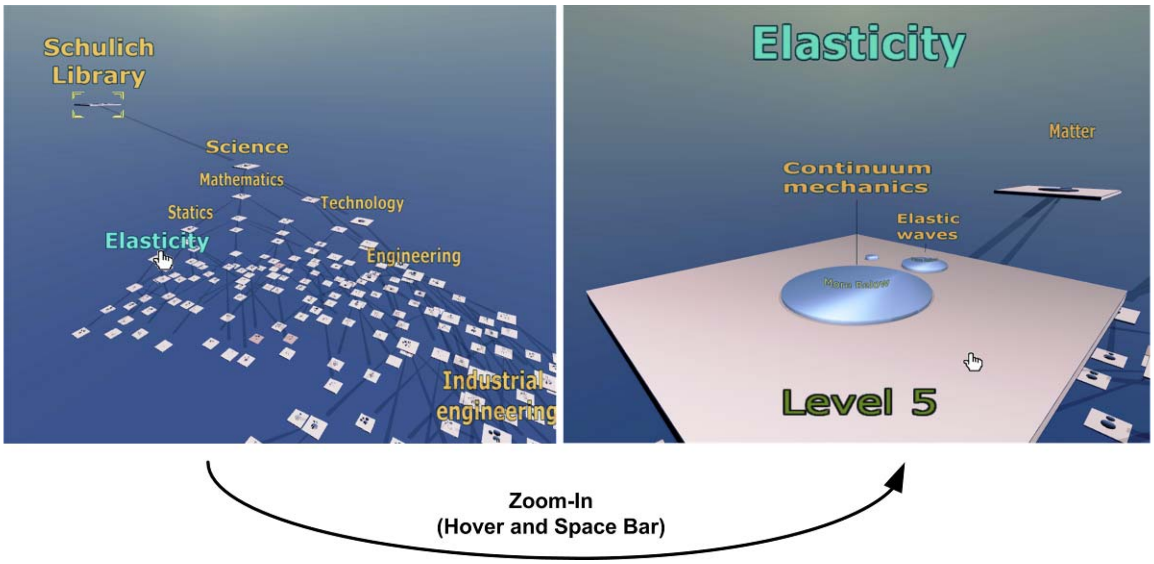


Figure 6.15: Zoom-In feature of SE-3D

A zoom-out would signal to the system that the user wanted an overview of the currently visible structure. This command was mapped to the "Z" key on the keyboard and a 2D magnifying glass icon on the screen. At any point in space, this feature created and animation during which the camera viewpoint rotated towards the closest subject map, and moved backwards until all of the structure was within the field of view.

A zoom-out was essentially an *overview on demand* (Shneiderman, 1996) as illustrated by Figure 6.16 showing the initial viewpoint is a close-up of a subject map (left side of Figure 6.16). The user presses the "Z" keyboard key, and using an animated transition, the system automatically moves the viewpoint backwards until the complete structure is visible.

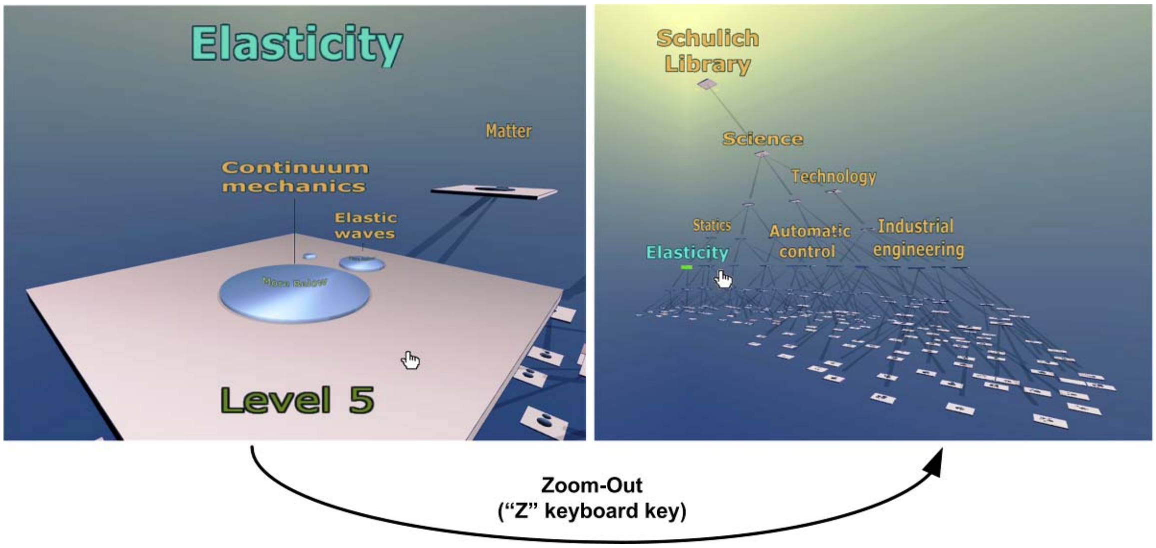


Figure 6.16: Zoom-out overview on demand from SE-3D

6.5.3 Constrained Movement in Space

Navigation in SE-3D involved moving around a virtual 3D space. Although some may have received training in this type of environment (e.g., gamers, airplane pilots), this would not be intuitive for many people. Given total freedom of movement, users could veer off towards pointless directions and end up in oblivion.

Movement can be facilitated by limiting the range of possible directions in space. These user navigation constraints "can be in many cases a helpful compromise that avoid user disorientation and enhances usability of 3D worlds" (Celentano & Pittarello, p. 276). Cubaud, Stokowski & Topol (2002) provided "constrained browsing within the collection in order to guide beginners" (p. 281), and Bladh *et al.* (2004) "limited the user's freedom" (p. 55) to reposition their 3D version of the classic Treemap (see section 4.2.2.4.2).

Users of SE-3D were not able to navigate too far away from the visible structure. Beyond a certain point, movement commands away from the structure were simply ignored. This maximum distance from the structure was iteratively tested and established as a factor of the dimensions of the currently visible structure. The factor was selected so as to provide just enough room to inspect the complete structure. This spatial constraint prevented users from travelling too far away from the information collection.

Speed of movement was controlled by the system using three intensities: 1) fastest when users were moving around the whole structure and distances were greater, 2) medium speed when moving within the structure, and 3) slowest speed when very close to a specific subject map and detail was at its greatest. The choice of speed was made automatically and periodically by the system based on the location of the user viewpoint relative to the semantic structure.

6.6 Interaction Between Search and Structure

SE-3D was meant to be a visual subject browsing tool integrated with the traditional search box and ranked list of results (see sections 4.1.3.7 and 4.3.4). Subjects containing bibliographic records matching a keyword search were isolated by removing all non-matching subject labels. The matching subject label color was also changed to reflect the number of matches contained by the subject (as defined in section 6.3.1.1.2). Changing label colors to reveal a search match was also used by Lee *et al.* (2006) in their hierarchy browsing tool. Non-matching subject labels would disappear leaving the subject map (i.e., the gray squares) as a visual placeholder for the subject so as to maintain visual continuity in the structure. In this manner, keyword searching acted as a dynamic subject label filter shown within the visual semantic hierarchy.

The keyword search feature returned results if all query terms matched at least one bibliographic record (i.e., title, author, series, subject headings). This is the same behavior as popular keyword Web search engines (e.g., Google). The absence of results was a failed search which would reset and redraw the visual subject structure to its initial top level view. Depending on characteristics of the failed search, one of the following system actions was triggered:

- **No index matches**: none of the user query terms were found in the search index. A user message suggested retyping or changing the query terms.
- **Some index matches**: one or more user query terms were found in the search index. A user message listed the term(s) not found in the search index and suggested removing them.

When results were found what should be the scope of search and its impact on the displayed parts of the subject hierarchy? A set of matching bibliographic items belonged to a set of LCSH terms, but these might not be currently visible in SE-3D. What would happen when there were no matching subjects in the currently visible parts of the structure? Two basic options were seriously considered:

- The search would reveal the relevant matches only in the currently visible parts of the structure. Matches in other parts of the structure would have to be found through additional expansion/retraction of subject branches.
- The search would prune the structure showing only the nodes relevant to the search. To prevent free-floating subjects without links to the structure, their parent subjects would be made visible until the top of the hierarchy was reached, even if none of them matched the search.

Initial developments produced the latter which restricted the visible structure to "nodes relevant to the search results" (Lee, et al., 2006, p. 1418). Developer testing showed the results of a drastic structure pruning could be disorienting; for example, a very specific and narrow search would eliminate most of the currently visible structure and reveal deep unknown branches. There could be little or no resemblance between the initial and post-search structures and the matching subject maps could be situated very far away from the initial viewpoint. It was feared searchers would lose sight of where they came from and their current location relative to known areas of the subject structure.

Section 4.3.4 states that SE-3D should allow users to progressively restrict searching to specific subject areas. This would mean that user-selected subject branches define a subset of the collection to be searched. This was consistent with the information patch concept (see section 3.3) which states that, depending on the task at hand, searchers may want to extract value for very large patches (i.e., broad sets of information items) or much smaller patches (i.e., more specific groups containing fewer items). This concept was also suggested by the “everything is a collection” property (Furnas & Rauch, 1998, p. 82) which states that a query can be specifically directed at any level of collection granularity.

A keyword search would reveal the matching subjects in the currently visible parts of the hierarchy, but this assumed there were such matches. In the absence of any matches in the currently visible structure, users would be left to *hunt* for the matches elsewhere in the structure. This might have been amusing for some but would likely be frustrating for others, and necessarily time consuming.

A combination of filtered searching and structure pruning was adopted as a comprise solution. Matching subjects were revealed in the current structure; if there were none, the structure was pruned to reveal the closest matching subjects not currently shown. A user message informed of this action.

The search impact on the subject structure is demonstrated by Figure 6.17. It shows the user has performed a keyword search for the keyword "Algorithms". The system automatically checks the spelling and returns a classic list of textual result (see top left of Figure 6.17).

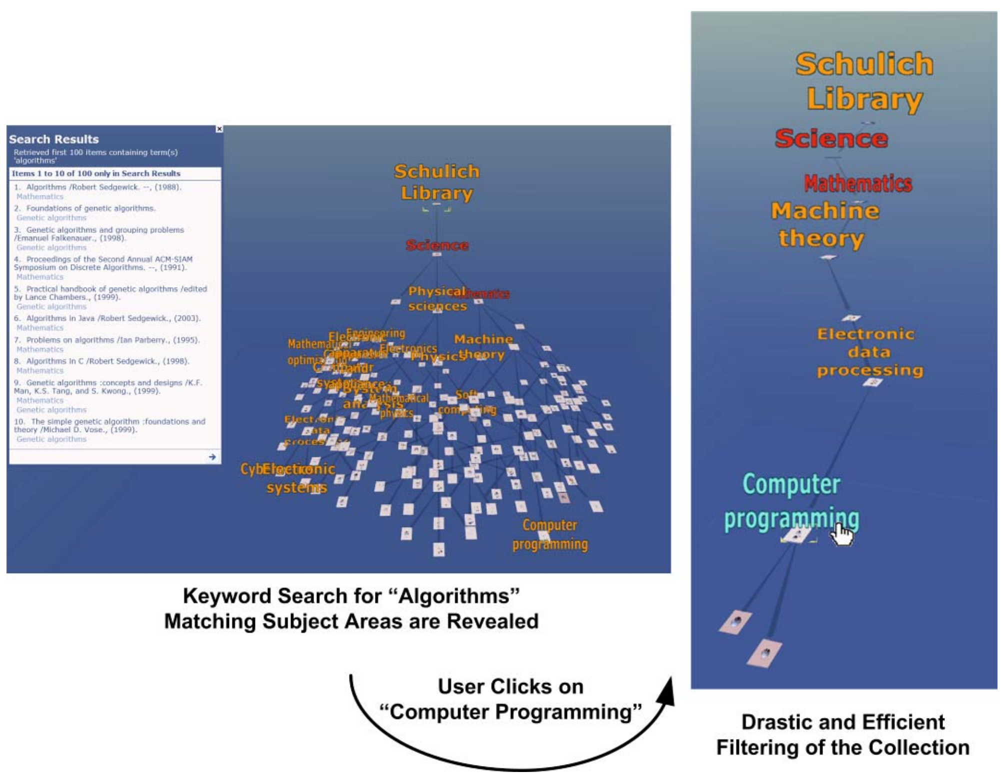
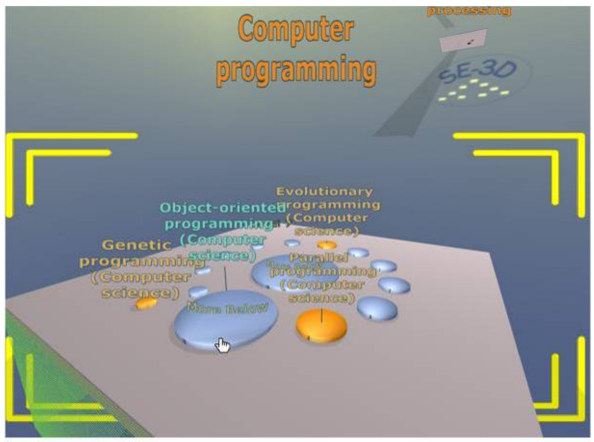


Figure 6.17: Search reveals matching parts of structure in SE-3D

The left side of Figure 6.17 also shows the matching subject labels have changed color from yellow to either orange (i.e., some matches) or red (i.e., many matches), all other non-matching subject labels have disappeared. The left side of the structure offers multiple matching subject areas for this broad query but the searcher recognizes "Computer programming" (bottom right of structure) as a promising related subject. He/she clicks on this label and the structure is drastically reduced to show the path to this subject (right side of Figure 6.17). Closer inspection of this subject is performed by zooming into the subject map as shown in Figure 6.18.



Zoom into "Computer Programming"
Matching subject in orange

Figure 6.18: Search reveals matching subject circles in SE-3D

Figure 6.18 shows child subjects belonging to computer programming; some in orange signifying they contain query matches, others in the original blue which do not. The searcher has narrowed the scope of the keyword searching.

From the interaction point shown in Figure 6.17, let us posit that the searcher dramatically changes his/her interest and chooses to search for the terms "heat exchange design". This narrow search has no matches in the currently visible part of the semantic hierarchy. As shown in Figure 6.19, SE-3D then redraws the structure from the top with the matching subject branches.

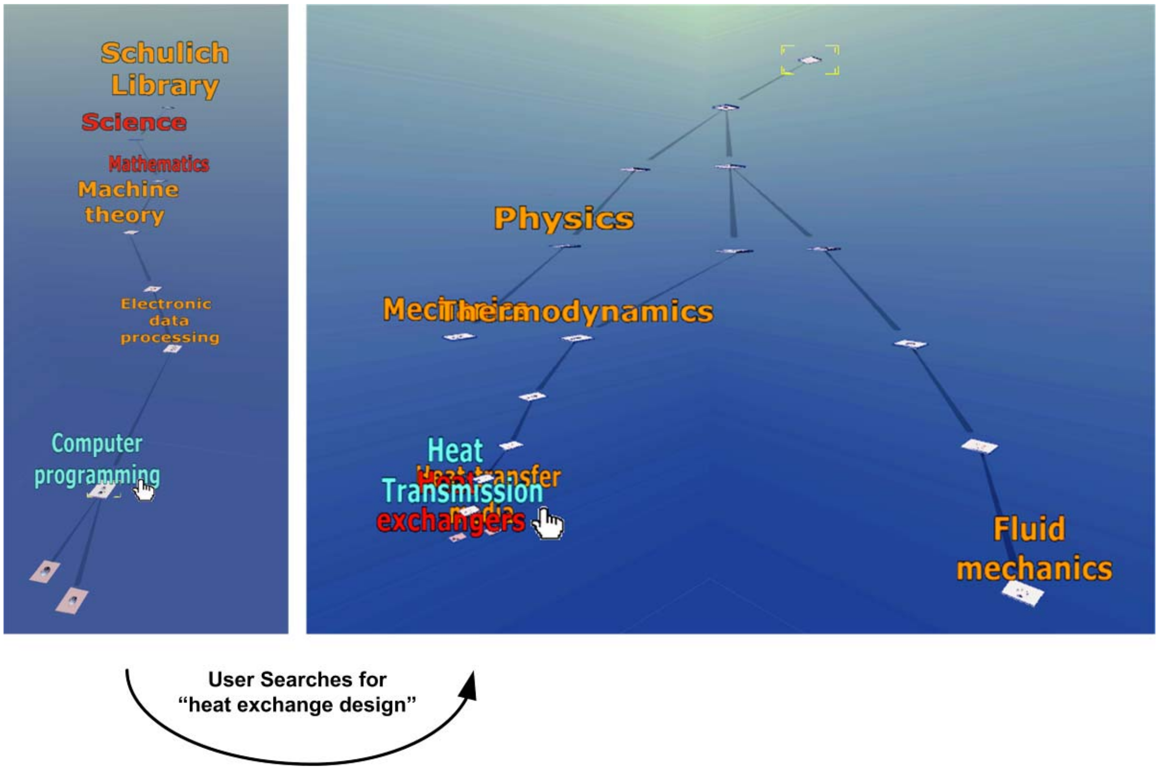


Figure 6.19: Visible structure without matches is drastically changed in SE-3D

Figure 6.19 shows that there is little resemblance between the initial visible parts of the subject structure on the left and the pruned filtered structure resulting from the keyword search terms. The seemingly unrelated query has removed the initial branches and replaced them with its own set of matching subject branches. Unintended disorientation resulting from this drastic change in the visual structure may be mitigated using an animated transition sequence. This would show the gradual disappearance of the initial branches and the appearance of the new set of matching ones. This would be beyond the scope of development for this version of SE-3D.

6.7 Formative Testing

SE-3D was iteratively tested informally with four different participants drawn from convenience sampling (i.e., family, friends, faculty). The initial design was refined until adequate stability and usability was achieved. Software is considered stable when it consistently delivers its intended functionality without crashing or producing unexpected behaviour. Initial formative tests asked participants to simply navigate in the environment and provide initial impressions. Later tests required them to use specific features while the developer/researcher recorded issues and their impressions.

Informal tests generated a wealth of issues simply by observing the behavior of the user and recording events which caused negative impressions, bugs, and navigation difficulties. These issues provided tangible directions for further development efforts. As the development process addressed each issue, new formative tests yielded fewer issues and the software was eventually ready for piloting.

User testing within the development cycle entails a tight coupling of testing and ensuing modifications. The ideal situation is a developer able to critically administer the user tests. This allows the developer to quickly hone in on problematic parts and even perform modifications in real time for immediate feedback and re-testing. In this manner, a single user test generated modifications which were integrated into the prototype and tested by the next pilot participant. This maximized the benefit of each test since each subject discovered new and unknown issues. These successive rounds of formative testing produced the following development priorities:

Basic navigation

Performed with a single user during two sessions lasting 90 and 120 minutes. Participant was asked to navigate from one visual object to the other while observations were noted. Potential solutions to problems encountered were discussed and the following issues became next development tasks.

- Integration of subject hierarchy and bibliographic data
- Choice between airplane vs. helicopter metaphor
- Gaze direction controls

Pilot 1: Completing tasks

Performed with a single new user during a session lasting 180 minutes. Participant was asked to complete a list of predefined IR tasks. This yielded requests for easier navigation and orientation in the 3D space. The following development tasks became a priority:

- Animated movements
- Wall mounted landmarks

Pilot 2: Efficiently Completing Tasks

Performed with a single new user during a session lasting 120 minutes. Participant was asked to complete a refined and shorter version of the first pilot task list. Further navigational features were suggested and implemented.

- Interaction between result list and visual structure
- Mouse wheel zooming

Pilot 3: Stable System

Performed with a single new user during a session lasting 90 minutes. Participant was asked to complete tasks as independently as possible. The system was stable and ran as expected. Tasks were completed without major obstacles.

These development priorities are further described in the following sections.

6.7.1 Subject term vs. Bibliographic information

SE-3D treated the complete bibliographic record, including its assigned subject terms and their synonyms, as a single information item. A keyword search could only match a bibliographic item—even in the unlikely event that the matched words stemmed strictly from controlled subject terms. This single index approach simplified the integration and usage of keyword searching; however, a search could not return strictly a subject term even if it exactly matched the user query. For example, Figure 6.20 shows that a keyword search for a specific LCSH term (e.g., "computer software") does not directly return the actual subject heading in the result list (left side of Figure 6.20); instead, it returns matching bibliographic items which may or may not have been assigned to computer software.

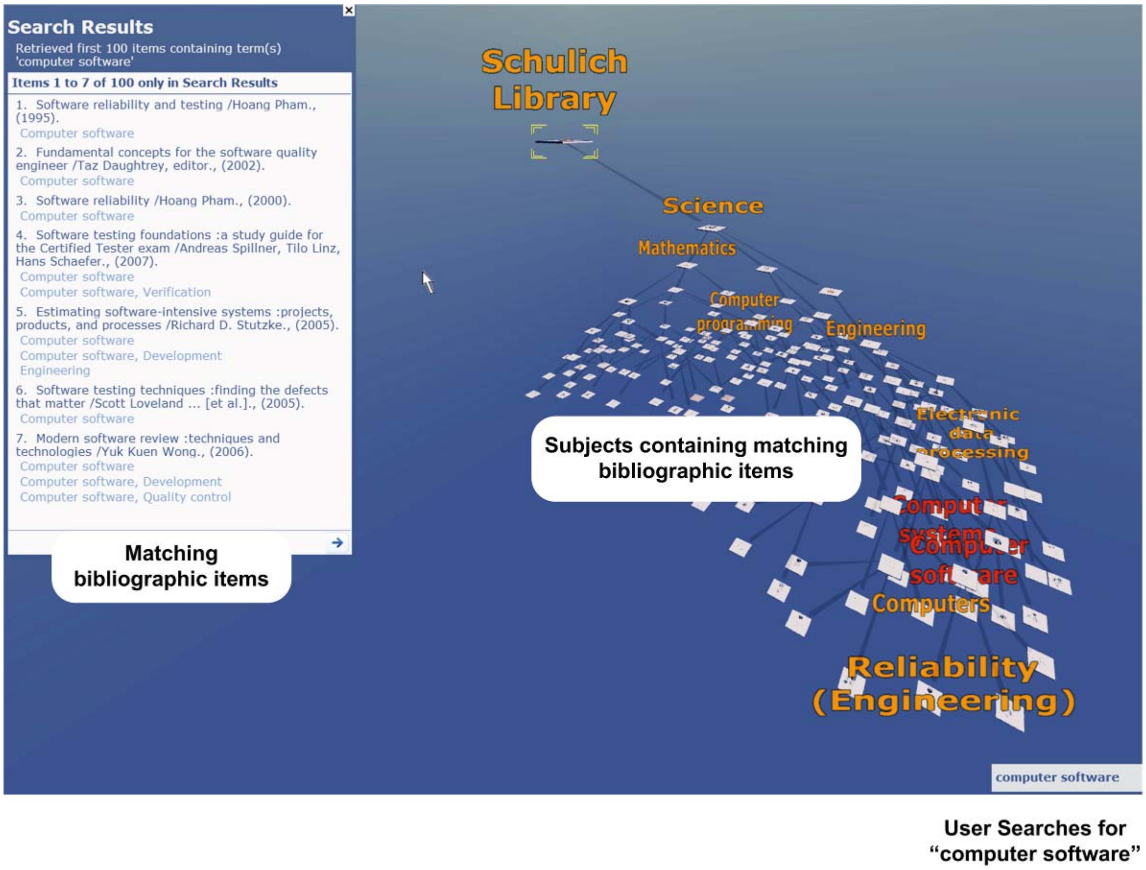


Figure 6.20: Effect of keyword searching in SE-3D

Figure 6.20 shows that specific subject terms are listed as part of the individual information items in the textual search results list. It also shows that subject labels attached to the semantic hierarchy do not necessarily match the query; in fact, the subjects labels are visible as soon as one or more of the assigned items match the query. This may be counter-intuitive to some searchers who might expect the "Computer software" label to be the only visible one since it is an exact match; however, the simplicity gained by the single index was deemed worth this potential issue.

6.7.2 Navigation Controls

> *The cost of getting a good viewpoint in 3D is almost always higher than clicking to follow a hypertext link or zooming in two dimensions. (Ware 2008 p. 97)*

As suggested above, a major issue with 3D navigation is providing a set of simple navigation controls for viewpoint selection. SE-3D formative tests confirmed this fact. There is no ubiquitously known set of commands for moving within an interactive 3D environment. Ease of navigation as a design priority was clearly indicated by early formative tests, and some of their mitigating suggestions were of high value.

6.7.2.1 Flying an Airplane vs. a Helicopter

The first issue concerned the basic flying vehicle metaphor. The initial navigation was based on airplane piloting controls. An airplane reaches a specific target in space if the pilot chooses an adequate approach trajectory. Specifically, a plane cannot move sideways; instead it must move forward while simultaneously rotating towards the target (i.e., follow an arcing trajectory). The airplane concept of forward movement is also difficult to grasp for novice users; it follows its nose which could be pointing in any direction including up or down. The airplane metaphor proved much too difficult for untrained users and something easier had to be found.

Most users could be expected to understand movement on a floor or 2D plane which is part of the everyday world. These movements could be intuitively mapped to the keyboard arrow keys. Forward, back, and side-to-side were understood as movement a 2D surface beneath ones' feet. The added ability to move up and down allowed the user to navigate to any other *floor* in space.

These controls resembled a simplified helicopter metaphor. The user could be standing still in space, move in any direction on the current horizontal plane, up or down. This was simpler set of navigation controls which capitalized on pre-existing human knowledge of movement on 2D surfaces.

6.7.2.2 Gaze Direction

The initial commands developed for gaze control (see section 6.1.5) were an obvious usability issue. Formative tests showed that users exploring this novel application would quickly glance at various points on screen and their mouse cursor would follow their eyes. Since gaze direction followed mouse movement direction, the result was a constantly shifting viewpoint difficult to control. A better solution had to be found without adding specific gaze control command keys.

A test user suggested the viewpoint should indeed follow the mouse cursor but only when it was hitting an edge of the screen. This would effectively be the user's way of signifying that he/she wished to see what was beyond the current viewpoint in the direction of the screen edge. This behaviour is illustrated in Figure 6.21 which shows the cursor hitting the top of the screen edge (left side of Figure 6.21) and right side of the screen (right side of Figure 6.21).

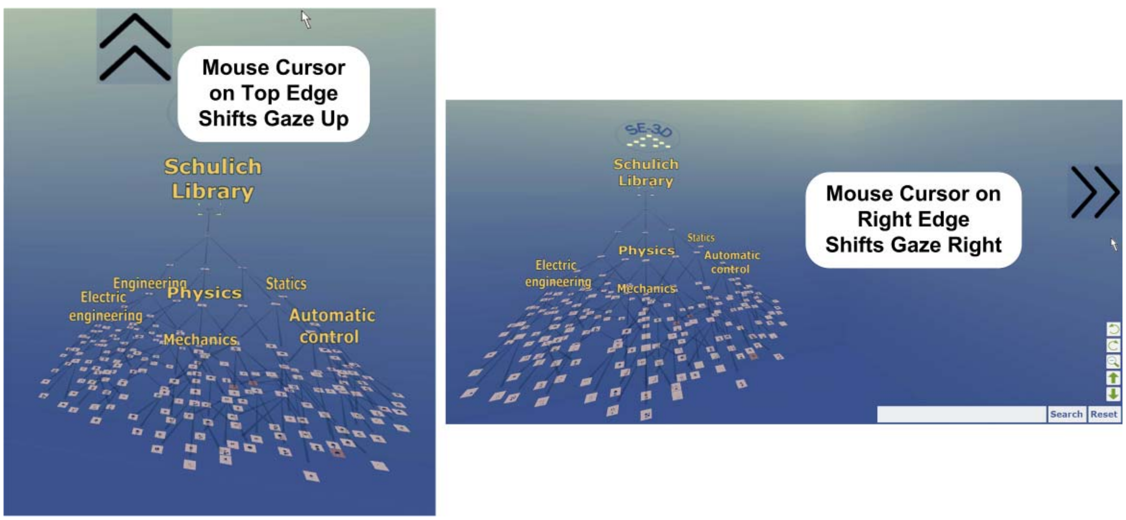


Figure 6.21: Mouse cursor on screen edges for gaze control in SE-3D

Figure 6.21 shows that the system confirms the execution of a gaze control shift by drawing double arrows pointing towards the gaze shift direction. This technique proved to be intuitive and unobtrusive, and is found in some commercial computer games (e.g., Unreal Tournament, Call of Duty).

6.7.2.3 Animated movements

Early formative tests revealed users often repeated sequences of movement and gaze direction shifts in order to inspect objects from different viewpoints. This was not surprising since "real time rotation (is) indispensable to understand 3D visualizations of trees" (Kleiberg, et al., 2001, p. 87). Smooth rotation was tedious and complicated for users to perform manually.

A test subject suggested there should be a way to automatically rotate the viewpoint in space. This was a good idea but rotation requires an axis and its location would not be explicitly specified by the user. To keep navigation controls simple this meant the system would have to *guess* the appropriate rotation axis for the current viewpoint.

When the user is very close to a subject map, a good vertical axis of rotation would likely intersect with its center. Rotation controls were mapped to the "F" and "G" keyboard keys and their use, when positioned close to a subject map, is illustrated by Figure 6.22.

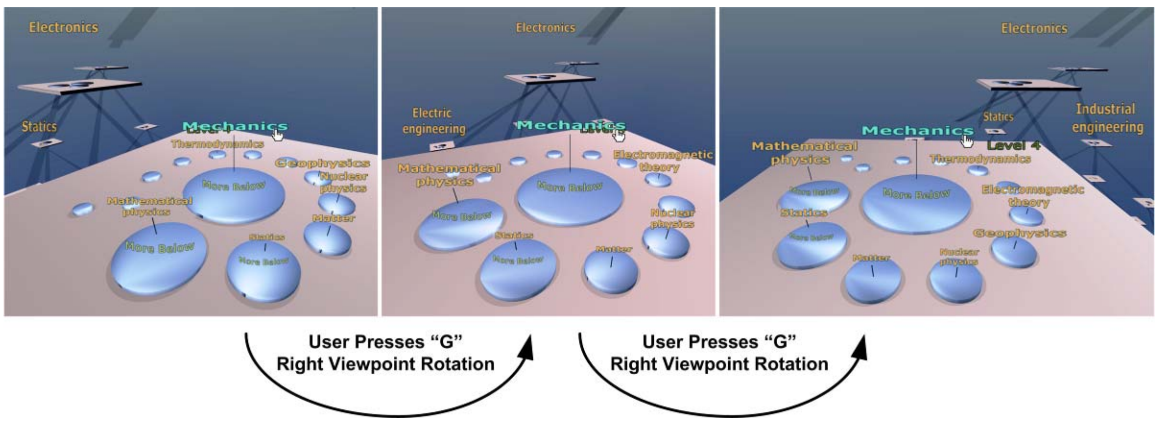


Figure 6.22: Automatic rotation around a subject map in SE-3D

Figure 6.22 shows the viewpoint is incrementally rotated counter-clockwise each time the user presses the "G" keyboard key. Users could perform a continuous rotation by keeping the key pressed as long as required. Pressing the "F" key produced the equivalent clockwise rotation.

Developing this type of interaction for rotating around the complete structure proved to be more difficult since there were various choices of rotation axis. The user might want to rotate around a specific part of the structure, around the center of mass, or around the current center map (i.e., last selected map).

The introduction of an additional control to specify the desired rotation axis was not considered since it would add complexity; thus, rotation around a user specified part of the structure would not be possible. Systematic rotation around the center of mass was not chosen because unbalanced structures would sometimes produce an axis of rotation positioned far from expectations.

The most promising option was to rotate around the current center map. This was the last subject map specifically selected by the user and its relative position was usually close to the structure center of mass. The resulting structure rotation around the current center map is illustrated by Figure 6.23.

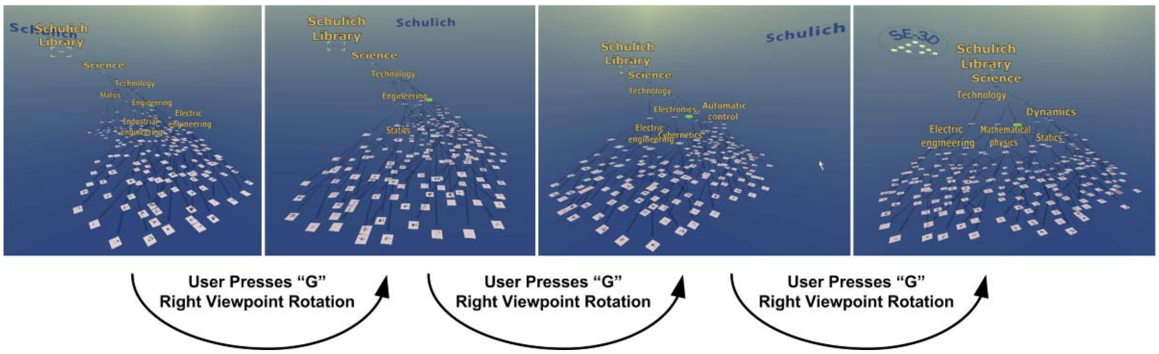


Figure 6.23: Automatic rotation around the whole structure in SE-3D

In the same manner as rotation around a subject map, Figure 6.23 shows that pressing the "G" key incrementally changed the viewpoint and allowed smooth rotation.

6.7.2.4 Wall Mounted Landmarks

Figure 6.23 also shows the implementation of wall mounted landmarks which provided directional cues. This was suggested by a test subject as he was rotating around the structure and wondered if he'd travelled completely around. He remarked it would be nice to have some unique identifying sign on each wall so he could quickly recognize in which direction he was looking at. The suggestion was consistent with information foraging theory (see section 4.1.3.3–Subjects as Patches in Static Space) and was quickly implemented with the following textual labels:

- Front wall showed a SE-3D logo as shown in right most pane of Figure 6.23
- Left wall showed the textual label "McGill"
- Right wall showed the textual label "Schulich"
- Back wall had no identifying feature which made it unique

Like cardinal directions, these wall mounted landmarks provided a fixed referential system to anchor the semantic hierarchy within a larger static world. These early graphic designs served the purpose of clearly and uniquely identifying the four sides of the world.

6.7.2.5 Interaction between Result List and Structure

Formative observations showed test subjects felt clicking on a piece of text should *bring* them somewhere. Consistent with Web hyperlinks, they would click on elements of the textual result list shown at the top left of Figure 6.20 and expect a corresponding system action. This seemed reasonable as there was indeed a place in space associated with each subject shown in the textual result list—a hyperlinked subject would automatically bring the user to that place in the semantic hierarchy.

The hyperlinked subject behavior is illustrated in Figure 6.24. It shows the user is curious about the subject of "Genetic algorithms" and decides to click on it in the textual result list (see left side of Figure 6.24). This triggers an automatic animated transition to the location of the subject within the semantic hierarchy.

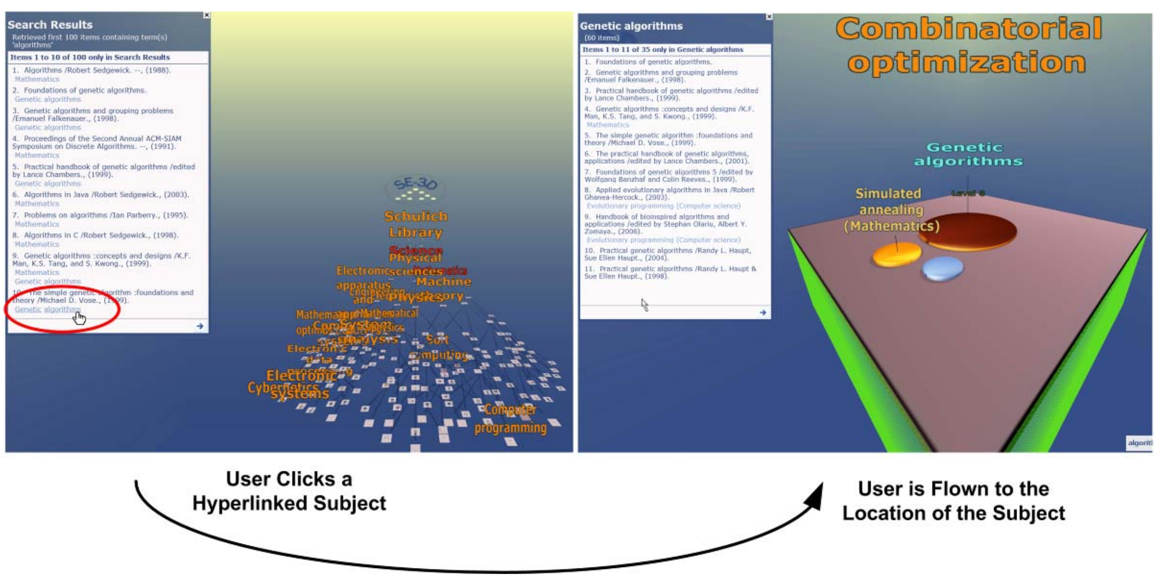


Figure 6.24: Hyperlinked LCSH subjects fly users to the visual semantic hierarchy

In this specific example, "Genetic algorithms" is the largest child subject of "Combinatorial optimization". This subject is a very specific subject map which ends a subject branch as indicated by the its darker shade of gray.

6.7.2.6 Mouse wheel zooming

Observations made during formative testing showed the zooming features were used sparingly. Inquiries with test subjects suggested the zoom commands (see section 6.5.2.2) were unintuitive. One highly pertinent suggestion was to replicate the mouse wheel zooming features of geographical information sites (e.g., Google Maps, Google Earth). This seemed a reasonable suggestion and it was the last implemented change in SE-3D before the start of controlled experimentation.

The initial implementation of mouse wheel zooming required a target subject for SE-3D to establish the direction of the movement. This is illustrated by Figure 6.25 which shows the user is indicating his/her subject of interest by hovering over its label (left pane of Figure 6.25) and rolls the mouse wheel forward to approach a fixed distance.

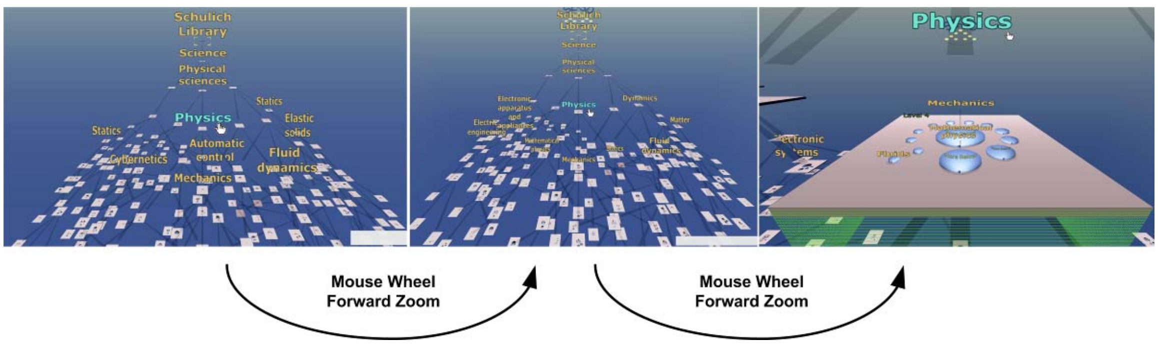


Figure 6.25: Mouse wheel zooming in SE-3D

Figure 6.25 shows that successive mouse wheel rolls would eventually carry the viewpoint to a close-up of the target subject map. The reverse was true for backward mouse wheel rolls; rolling the wheel backwards moved the viewpoint backwards a fixed distance until the whole of the structure was visible.

This feature was a late addition and its implementation was not perfect. It should not be necessary for users to indicate a target subject by hovering. A simpler solution would be to simply move forward and back along the current gaze direction.

6.7.2.7 Final Navigation Controls

The addition of navigation features stemming from formative testing to the initially developed set (described in section 6.5.2) produced the final set of navigation controls as illustrated in Figure 6.26. It shows simplistic renditions of a screen, keyboard and mouse along with SE-3D specific commands.

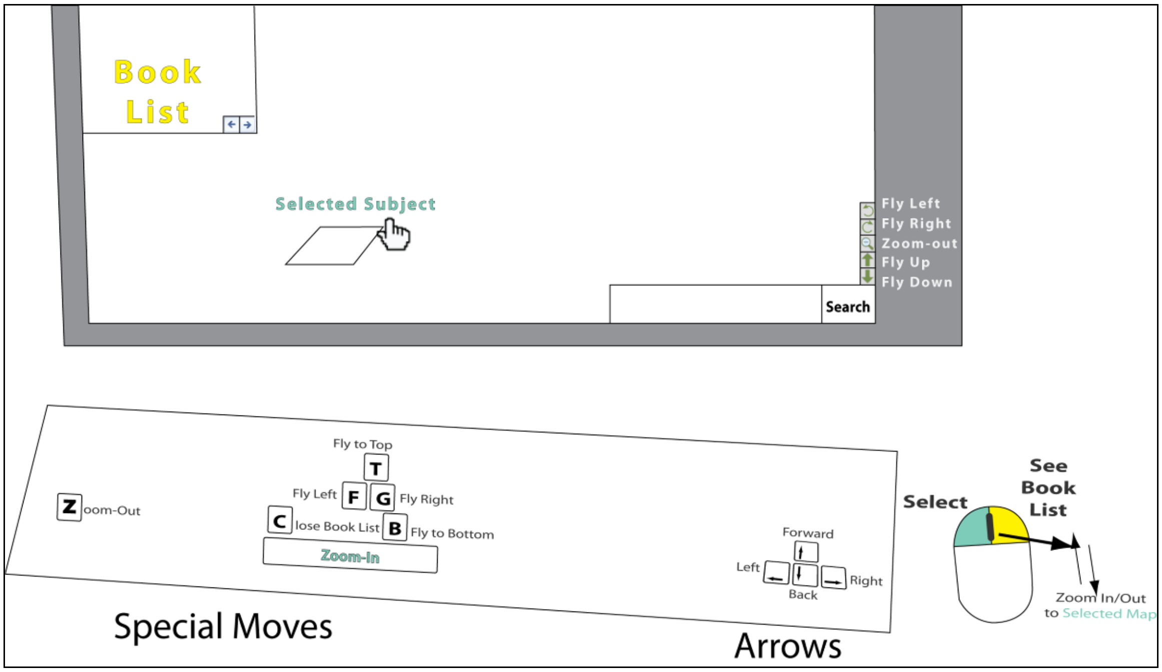


Figure 6.26: Final set of navigation controls used by SE-3D

As Figure 6.26 shows, the SE-3D controls were as follows:

User Action		System Response
Keyboard	**Mouse**	
"Z" key	Backward wheel roll Magnifying glass screen button	Zoom out
Space bar	Forward wheel roll	Zoom in
"T" key	Up arrow screen button	Movement straight up
"B" key	Down arrow screen button	Movement straight down
"F" key	Left turn screen button	Rotate clockwise around structure or map
"G" key	Right turn screen button	Rotate counter-clockwise around structure or map
Arrows		Move forward, back, and sideways on current plane
	Left click	Expand selected subject branch
	Right click	Show book list for selected subject
"C" key	Close book list button	Book list disappears

Table 6.1: Detailed list of navigation controls used by SE-3D

These controls could be mapped many other ways but they proved adequate during formative tests and piloting. They would be used during further testing.

6.8 Development Conclusion

This study has described the development and formative testing of a virtual reality information browsing and searching tool. SE-3D aimed to facilitate information retrieval by integrating visual browsing of the subject structure and keyword searching of the collection. This study is relevant to research in the fields of information visualization, bridges between information retrieval and information organization, and usability testing. Specifically, SE-3D includes a novel integration of keyword searching and IV browsing, a practical way to manage hundreds overlapping labels in 3D space, and a case study of IV for IR application design and formative testing.

The literature on IV for IR contains few comparable applications to SE-3D (see section 4.2). A descendant of Cone Trees (Robertson, *et al.*, 1991), SE-3D applied the technique to a bibliographic collection organized using LCSH, integrated with keyword searching. In this regard SE-3D builds on the work of Cat-a-Cone (Hearst & Karadi, 1997) described in section 4.2.4.2.2; however, there are critical differences between the two. Firstly, SE-3D used the larger and broader LCSH hierarchy while Cat-a-Cone was based on the more constrained MeSH taxonomy.

Beyond the broader subject hierarchy used by SE-3D, one of its defining characteristics is the integration of visual subject structure browsing and keyword searching. Cat-a-Cone showed where in the MeSH hierarchy a single item was situated. This is in contrast with SE-3D which reveals the distribution of keyword search matches throughout the currently visible hierarchy. SE-3D users can prune the subject hierarchy based on keyword searching or mouse selection. This may afford drilling down towards more specific relevant areas for the current task. This technique of highlighting matching items within a visual representation of the collection has been done with automatically generated 2D document maps (Davidson, et al., 1998; Seeling & Becks, 2004), but no existing report were found of its usage in 3D or with a predefined subject structure.

A preliminary search of the literature yielded few practical methods of dealing with hundreds of overlapping floating labels, fewer still in 3D interactive spaces. The approach used by this study capitalized on the intuitive notion that, from any viewpoint, labels hidden behind others are simply not legible. To reduce visual clutter these illegible labels are not drawn until the user changes his/her viewpoint. The technique proved to be usable but it could be refined.

The iterative testing used to develop and refine SE-3D is an applied case of IV design. Although many are willing to develop IV applications, few go on to test them (see section 3.2.8–IV User Evaluations). It is one thing to design and build an IV application; significant additional work is required to produce stable software lay searchers can use with minimal assistance and oversight (Brooks, 1995). This is exemplified by HCI problems during usability tests (Cribbin & Westerman, 1999, pp. 206-207; Lee*, et al.*, 2006, p. 1423; Sutcliffe*, et al.*, 2000a, p. 752).

Formative tests with pilot participants helped address critical navigation issues. These three early interactions with real users were necessary and efficient in terms of number of new issues found per test user. The resulting VR application for information browsing and searching is a kind of information retrieval environment. This type of application offers an integrated set of search tools aiming to save time and increase the accuracy of information searches.

6.9 Limitations

SE-3D was developed as an information access tool for a bibliographic collection organized using LCSH in the domain of science and engineering. The resulting software application may not be as usable with other collections whose subject organization offers too few hierarchal relations. Porting SE-3D to other collections and information organization schemes will be the subject of future research.

Dynamic labels demanded significant processing, and this became a performance bottleneck which could be mitigated by researching and implementing performance enhancements. Finally, SE-3D is necessarily a product of the many design decisions made during its development and testing. Design is partly a creative endeavour and other designers may have chosen different solutions.

Chapter 7: EVALUATION

Let us restate the specific question the controlled experiment hoped to address:

> *Are there differences between a 3D IV system and a text-only subject browser in terms of user performance and experience for undergraduate students performing IR tasks? Performance is measured by completion time and accuracy. Experience is measured by perceived speed, usefulness, ease of use and preference of the system.*

An attempt to answer this question was made using a controlled comparative experiment. The objective was to evaluate the effect of the search interface (i.e., textual Web baseline or SE-3D) on completion time, accuracy and affective reactions. The experiment was a repeated measures two-factor within subject design; the same group of test participants performed equivalent tasks on both system. Differences observed cannot therefore be attributed to individual differences but rather to the system, the task, or a combination of the two. This specific study aimed to assess novice user performance and impressions of SE-3D and verify that it "minimizes the amount of learning imposed on users" (Sutcliffe, 2003, p. 76).

A repeated measures design assumed independence of observations. This implied the tasks performed on each system were different enough to prevent learning and of equivalent difficulty. Observation independence was ensured by counterbalancing the order of systems and task sets using a Latin square design (see Table 7.1).

	Baseline	**SE-3D**
Task Set 1	5 users Task set 1 with baseline Task set 2 with SE-3D	5 users Task set 1 with SE-3D Task set 2 with Baseline
Task Set 2	5 users Task set 2 with baseline Task set 1 with SE-3D	5 users Task set 2 with SE-3D Task set 1 with Baseline

Table 7.1: Latin square design

The following sections describe the baseline system used for comparison, the test participants and selection procedure, their incentive, the training they receive with each system, and the tasks they perform. These are followed by the detailed experimental procedure and results.

7.1 Baseline System

A baseline Web-based interface was developed to reproduce hyperlinked ranked lists of results and subject browsing features found in a traditional OPAC exemplified by the McGill MUSE library catalogue[1]. The baseline system was representative of a traditional Web-based search box and OPAC LCSH structure browsing features (see review by Julien *et al.* (2008)). Comparison with a classical outline view (e.g., Windows File Explore) was ruled out because it was judged novice user conditions could not be created for this ubiquitous tool. The participants were filtered via pre-selection questionnaire (see Annex 2) to ensure they all lacked familiarity with traditional LCSH browsing features. Test participants had experience with hyperlinks and Web page navigation which might have put SE-3D at a disadvantage.

Users of the baseline system always had access to a search box and list of textual results which was identical in content and ordering to that offered in SE-3D. The search box offered by the Web baseline (see Figure 7.1) used the same query engine as the SE-3D test system. The only difference between the systems was the LCSH hierarchy navigation features.

[1] http://catalogue.mcgill.ca/ Last access: 17/09/10

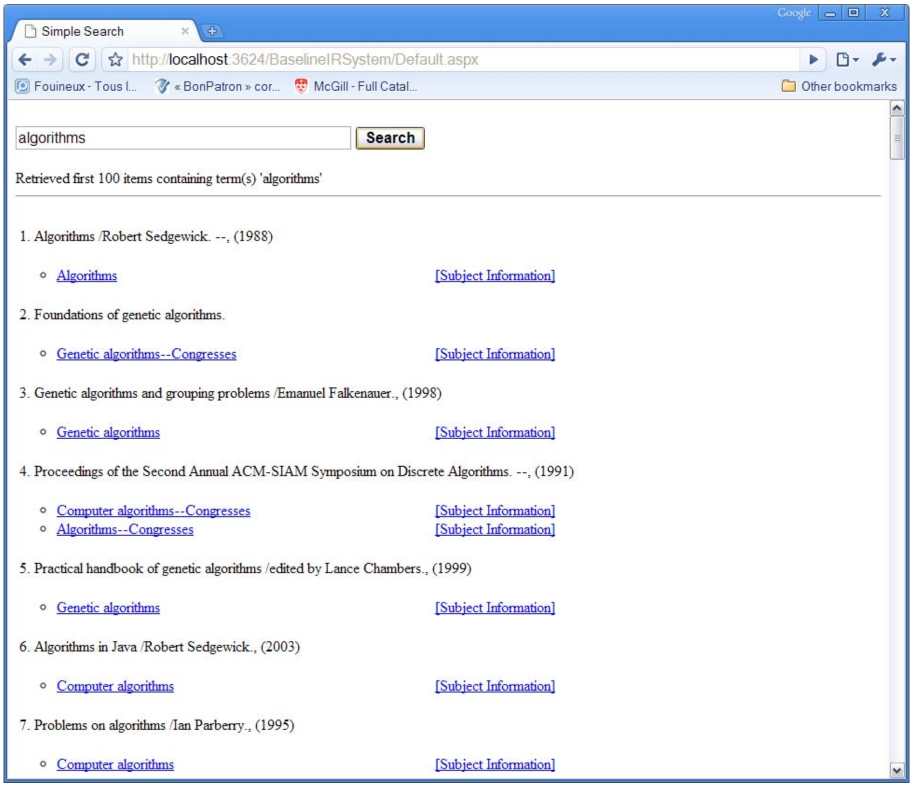


Figure 7.1: Baseline Web System Search Box and Results

When the user clicked on a "[Subject Information]" hyperlink shown in Figure 7.1, the system offered a list of broader and narrower subjects (see Figure 7.2).

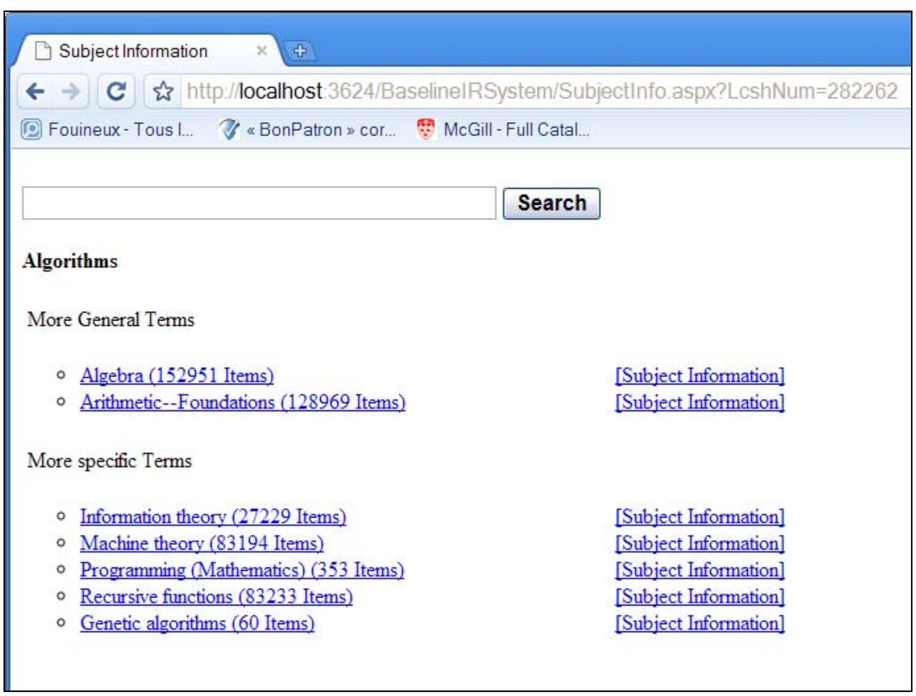


Figure 7.2: Baseline Web System Subject Information

Each broader or narrower subject could be selected to show its assigned items (i.e., equivalent result list shown in Figure 7.1) or show that subject's broader/narrower terms. This type of HCI based on successive lists (i.e., one for each subject term) of hyperlinked subjects terms is found in many traditional OPACs including the McGill University MUSE[1] catalogue and the latest Laval University Ariane2 catalogue[2].

7.2 Users

24 test participants were recruited from the McGill first year undergraduate engineering population. This number was chosen based on TREC guidelines which demand a minimum of 16 users per interface (Voorhees & Harman, 2005) and a review of information visualization usability studies (Julien, et al., 2008) suggesting more users may be necessary to isolate the effect of the interface amongst variability stemming from task types. First year students were chosen to ensure minimal prior knowledge of the McGill library collection.

The collection used during the experiments was a science and engineering collection of physical items (i.e., books and printed periodicals). The choice of engineering students seemed well advised since interest in the collection has been shown to be a source of participant motivation in search usability studies (Borland & Ingwersen, 1997). Subject interest may provide additional insurance of sustained “subject involvement” (Turetken & Sharda, 2005, p. 292) required for learning and using a novel visual system as compared to a more familiar textual baseline.

7.3 Recruitment, Selection and Compensation

Participant recruitment was initially performed using a bulletin board add (see Annex 1) in the science and engineering faculty building. The advertisement described the project, the type of participants sought, and asked interested students to signify their interest by sending an email to the researcher. This initially produced meagre results and more active recruitment was required.

[1] see http://newaleph.mcgill.ca/
[2] see http://ariane2.bibl.ulaval.ca/ariane/

By far the most productive recruitment activity were live presentations given by the researcher at the beginning of mandatory first year (i.e., U0 or U1) engineering courses (e.g., physics, chemistry) with permission from their respective instructors. In less than three minutes the researcher addressed these large groups of students (100+ per class), quickly described the research and, if they were interested, asked them to pick up a paper hand out version of the recruitment advertisement. It invited those interested in participating to send an email to the researcher.

These interested students were sent an electronic pre-selection questionnaire (see Annex 2). The potential participant completed the questionnaire online and returned it via e-mail. This allowed selection of participants based on the following criteria:

- No known form of color perception deficiency
- Age between 18 and 21
- Existing yet minimal prior usage of McGill MUSE catalogue
- Little or no knowledge of LCSH organization
- Little or no prior usage of MUSE LCSH browsing features
- Native speakers of English

Candidates not selected were notified via e-mail. Individual appointments were scheduled with selected candidates. Experiment participants received $20 once the session was completed.

7.4 Procedure

The following sections describe the training dispensed to participants, the tasks performed and the detailed experimental protocol.

7.4.1 Training

Participants were told they were testing ways to explore the library collection and would receive a self-paced presentation demonstrating the system features (see Annex 4. *Baseline Web System Training Slides*, and Annex 5. *SE-3D Training Slides*). They were able to use the respective systems during these presentations. This was followed by four training tasks described in the following section. The training session lasted until the user was satisfied he/she was ready or a maximum of 20 minutes per system.

7.4.2 Tasks

There is no consensus concerning which types of tasks are appropriate for controlled IV testing (see section 3.2.8.4 *Tasks*). As suggested by IV testing reports (see section 3.2.8.4 *Tasks*), a range of IR tasks was chosen. They included semantic hierarchy navigation, simple and complex IR.

7.4.2.1 Hierarchy Navigation Tasks

Testing hierarchy navigation tasks assumes that users will be better information seekers if they understand how the information is organized by easily navigating the organizational structure. Hierarchy navigation tasks have often been used in IR usability testing (Barlow & Neville, 2001; Bladh, et al., 2004) and include at least three sub-types (Plaisant, *et al.*, 2002):

1. finding a specified existing subject,
2. finding relations between subjects (e.g., common parent or children, relative prevalence of subjects), and
3. returning to a previously visited subject.

These three types of hierarchy navigation tasks were included in this experimental design.

7.4.2.2 Information Retrieval Tasks

Section 3.1 *Information Retrieval* categorized IR tasks as known-item searching, extended fact-finding, open-ended browsing and problem analysis, and exploration of availability (Shneiderman & Plaisant, 2005). This broad classification is the stated task model of some IV testing studies (Grun, et al., 2005; Turetken & Sharda, 2005), and very similar to task models used in others (Sutcliffe, *et al.*, 2000a; Zhang & Marchionini, 2005). To support comparability with existing research, this model also served as the basic IR task classification for this research. There are other IR task models (see section

3.1) but the Shneiderman & Plaisant (2005) classification was the most popular within IV evaluation literature.

Only a subset of the broad IR task classification by Shneiderman and Plaisant (2005) could be performed with the specific tested collection. Types of tasks which could be tested were limited to exploration of availability and known-item searching since the collection was comprised of bibliographic records without the full-text information. A bibliographic record or subject heading is a form of collection coverage and availability information. Finding a specific LCSH term or a bibliographic record containing a specific word is arguably a kind of known-item searching where the search constraints and success criteria are clearly defined before the start of the IR interaction.

Determining if a collection offers documents covering a specified subject can be simple or complex depending on the wording of the question (Pirolli, *et al.*, 2000, pp. 163-164; Yee, *et al.*, 2003, p. 406) and familiarity with the subject. Questions that contained vocabulary matching the target subject were defined as simple while the opposite (i.e., question vocabulary does not match target subject) were defined as complex.

Pirolli and Card (2000) specifically controlled participant familiarity with target subject terms. They asked participants to rate their familiarity with a subject term before locating the term in the hyperbolic browser (see section 4.2.4.1) as compared with a standard Windows Explorer outline interface (see Figure 1.1). They found that familiarity was highly correlated with target finding effectiveness. This study controlled for effects of participant familiarly with subjects by counter-balancing task set and system pairings, and random assignment of participants to these pairings.

7.4.2.3 Tasks Tested

The experimental design required two equivalent yet independent task sets. This meant both sets of tasks should be equally difficult to perform, and performing one set of tasks should not have an effect on the performance of the other set of tasks. Table 7.2 lists the two task sets and the two training task sets.

Question Type		Task Set 1	Task Set 2	Training Set 1	Training Set 2
Hierarchy Navigation	Target close to top of hierarchy	1. Find a subject which directly belongs to "Operations Research"?	Find a subject which directly belongs to "Mechanics"?	Find a subject which directly belongs to "Electric engineering"	Find a subject which directly belongs to "Mechanical Engineering"
		(2) Find the nearest common more general subject of "Steelwork" and "Mechanics".	Find the nearest common more general subject of "Electric motors, Linear" and "Nuclear Physics"	Find the nearest common more general subject of "Motors" and "Remote Control"	Find the nearest common more general subject of "Electric current converters" and "Standardization"
	Multiple Choice	(3) Which of these subjects contains more documents? "Matter, Properties" "Solution (Chemistry)" "Optical transducers" "Electric machinery, Synchronous"	Which of these subjects contains more documents? "Physical measurements" "Dynamics" "Jet engines" "Loudspeakers"	Which of these subjects contains more documents? "Mathematics" "Science"	Which of these subjects contains more documents? "Technology" "Science"
	Target close to bottom of hierarchy	(4) Find a subject which directly belongs to "Sand, Foundry".	Find a subject which directly belongs to "Cleaning compounds"		
	Multiple Choice	(5) Which of these subjects is more specific? "Power transmission" "DC-to-DC converters" "Artificial satellites, Control systems" "Set theory"	Which of these subjects is more specific? "Wave-motion, Theory of" "Electric generators, Alternating current" "Airplanes, Handling characteristics" "Machinery"		
		6. Return to "Operations Research" and find another one of its more specific subjects.	Return to "Mechanics" and find another one of its more specific subjects.		
Simple Retrieval		You are looking for information on Visual Basic programming. Find 2 promising books on the subject.	You are looking for information on internet security. Find 2 promising books on the subject.	Find what you think would be a promising book on land surveying	Find what you think would be a very promising book on bridge design
		You are looking for information on game design. Find 2 promising books on the subject.	You are looking for information on climate change. Find 2 promising books on the subject.		

Question Type		Task Set 1	Task Set 2	Training Set 1	Training Set 2
Complex Retrieval		How would you go about fixing a broken computer? Find 2 promising books for this question.	Why do concrete pipes eventually fall apart? Find 2 promising books for this question.		
		What kind of trees are used for construction of buildings? Find 2 promising books for this question.	How can you fix a leaky faucet? Find 2 promising books for this question.		

Table 7.2. IR Tasks Tested during Controlled Comparative Experiment

As shown in Table 7.2, there were six hierarchy navigation tasks, two simple and two complex IR tasks. Within hierarchy navigation, tasks 1 and 5 were searches for specific subjects situated respectively near the top and bottom of the LCSH hierarchy; this offered control over effects of the target subject depth within the semantic hierarchy.

Tasks 2, 3 and 4 were subject relations tasks, and task 6 was a return to a previously visited subject. Subject relations tasks 2, 3 and 4 involved finding two subjects and their common parent several broader terms away (task 2), judging the relative prevalence of subjects or branches of the hierarchy (task 3), and judging relative specificity of subjects (task 4).

Notice that the numbering format of these tasks reflects their presentation order during the test. Tasks 1 and 6 had an imposed order since they were necessarily dependant, tasks 2 to 5 are numbered within parentheses indicating their order was randomized before each test. The order of the two simple and the two complex IR tasks was also randomized. The order of task types (i.e., hierarchy navigation, simple and complex IR) was fixed and followed task type complexity. The system and task set order pairings were evenly distributed and counterbalanced (see Table 7.1).

Equivalency and independence of the task sets were ensured by choosing target subjects placed at equivalent depth and breadth of the semantic hierarchy but in different main branches. The objective was to have target subjects that were just as difficult to find but in different unrelated areas of the collection. The prevalence of two main branches of the test collection (pure sciences and engineering) provided different subject areas where equivalent target subjects could be found. The equivalency of the task sets was confirmed by pilot testing.

Paper and pencil were available to the participants throughout the session. It was expected they might record quantities of items in subjects or the level of a subject for later comparison. The papers were collected but their analysis was considered beyond the scope of this study.

7.4.3 Independent Variables

Independent variables in this case were participants, system and task; however the repeated measures design controls variability stemming from the participants leaving only the latter two. Tasks were averaged across their type (i.e., hierarchy navigation, simple and complex retrieval) producing three levels of task type.

This experiment was a repeated measures two factor within-subject design. The two within-subject factors were system (2 levels) and task types (3 levels). Each participant generated accuracy and time data for each combination of system and task type (i.e., 2 × 3 = 6). Table 7.3 illustrates the data collected for each subject.

	Baseline Web System	**SE-3D**
Hierarchy Navigation (6 tasks)	Accuracy, Times	Accuracy, Times
Simple Retrieval (2 tasks)	Accuracy, Times	Accuracy, Times
Complex Retrieval (2 tasks)	Accuracy, Times	Accuracy, Times

Table 7.3. Experiment Data Collected for Each Subject

7.4.4 Dependant Variables

Table 7.4 provides a list of dependant variables, their conceptual and operational definitions, and the specific measures collected.

Concept		Operational Definition	Specific Measure
Automated	Performance	Accuracy for the tasks	Number of correct answers
		Speed performing the tasks	Time between the moment the task is shown and the moment the participant presses the "Next" button
Post-test	Speed Perception	How fast is each system?	Two 11-point Likert scales (one per system). 0="Not at All", 10="Very Much" Free form text to explain answer
	Usefulness	How useful is each system?	Two 11-point Likert scales (one per system). 0="Useless", 10="Extremely Useful" Free form text to explain answer
	Ease of use	How easy is each system to use?	Two 11-point Likert scales (one per system). 0="Extremely Difficult", 10="Extremely Easy" Free form text to explain answer
	Preference	Which system do you prefer? Why?	Two 11-point Likert scales (one per system). 0=" Extremely Difficult ", 10=" Extremely Easy " Free form text to explain answer
	Familiarity	How familiar were you with Web based search engines?	11-point Likert scale. 0="Not at All", 10="Extremely Familiar"
		How familiar were you with 3D game like interfaces such as first person shooter games (e.g., Call of Duty, Unreal) or 2nd Life?	11-point Likert scale. 0="Not at All", 10="Extremely Familiar"

Table 7.4: Dependant variables, operation definitions and specific measures.

Details of the dependant measures shown above are provided in the following sections.

7.4.4.1 Information Foraging Measures

Dependant variables stemming from the IF model are time per question and within-path value. These IF objectives were operationally defined as combinations of time and accuracy:

- **Time**: reducing between-patch navigation should reduce time (Olston & Chi, 2003) taken for each question from the instant the question is shown until the subject moves on to the next question.

- **Accuracy**: represents within-patch value defined as the number of correct answers.

Accuracy described if a question was answered correctly and the assessment of right versus wrong answers could be partly subjective. Referring to Table 7.2, hierarchy navigation tasks offered finite constrained answers which made accuracy judgement straightforward. For example, task set 1, question 4, either the answered subject directly belonged to "Nitrides" or not.

Simple and complex retrieval tasks shown in Table 7.2 required a more flexible definition of accuracy. For a question that asked for two books, an answer with only one correct book would receive a score of 50% for that question. Considering the ill-constrained definition of relevance (Chowdhury, 1999; Chu, 2003; Meadow, et al., 2007), accuracy scores were assessed by the researcher and validated by a professional librarian.

7.4.4.2 Post-test

Performance considerations are one aspect likely to affect system preference; however, as Davis (1989) showed, perceived usefulness and ease of use may also affect users choosing the system they prefer. These self assessed concepts were measured using post-test questions with 11-point Likert scales. IV usability test participant familiarity with the tested tools may have significant effects on performance (see section 3.2.8.3). This study tested two types of IR interface interaction styles: 1) Web based search engines and 2) desktop virtual reality gaming; as a result, self-assessed familiarity with Web based search engines and 3D games was also recorded post-test.

7.4.5 Experimental Sessions and Testing Engine

The tests were administered via a custom Microsoft Access application which generated and balanced the task set and system pairings, guided the user through the test and recorded test and post-test measurements. The testing engine ran on a separate screen from the test system (see Figure 7.3).

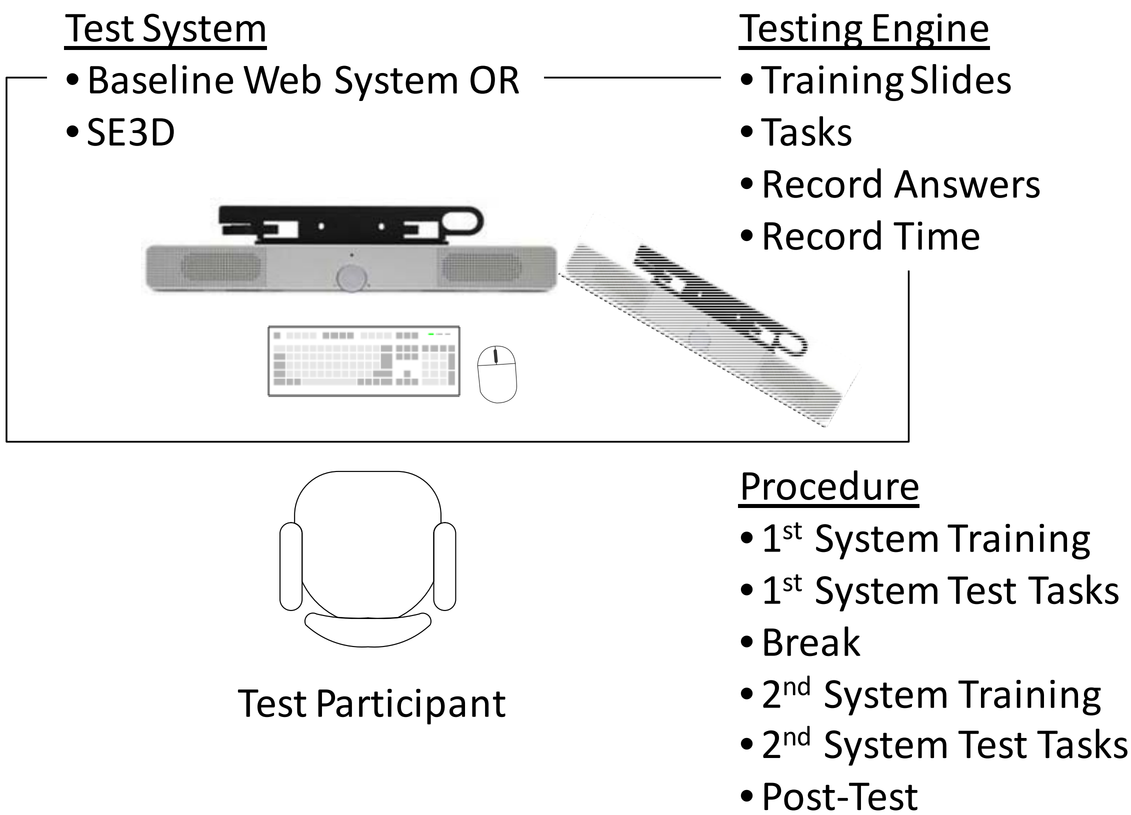


Figure 7.3. Experiment Setup

Three participants piloted the following experimental protocol:

1. **Greeting and Informed Consent**: Subject is greeted and offered refreshments and is asked to sign the informed consent form (see Annex 3).
2. **Testing Engine Access**: When the subject is ready he/she is given access to the testing engine (see Figure 7.4).

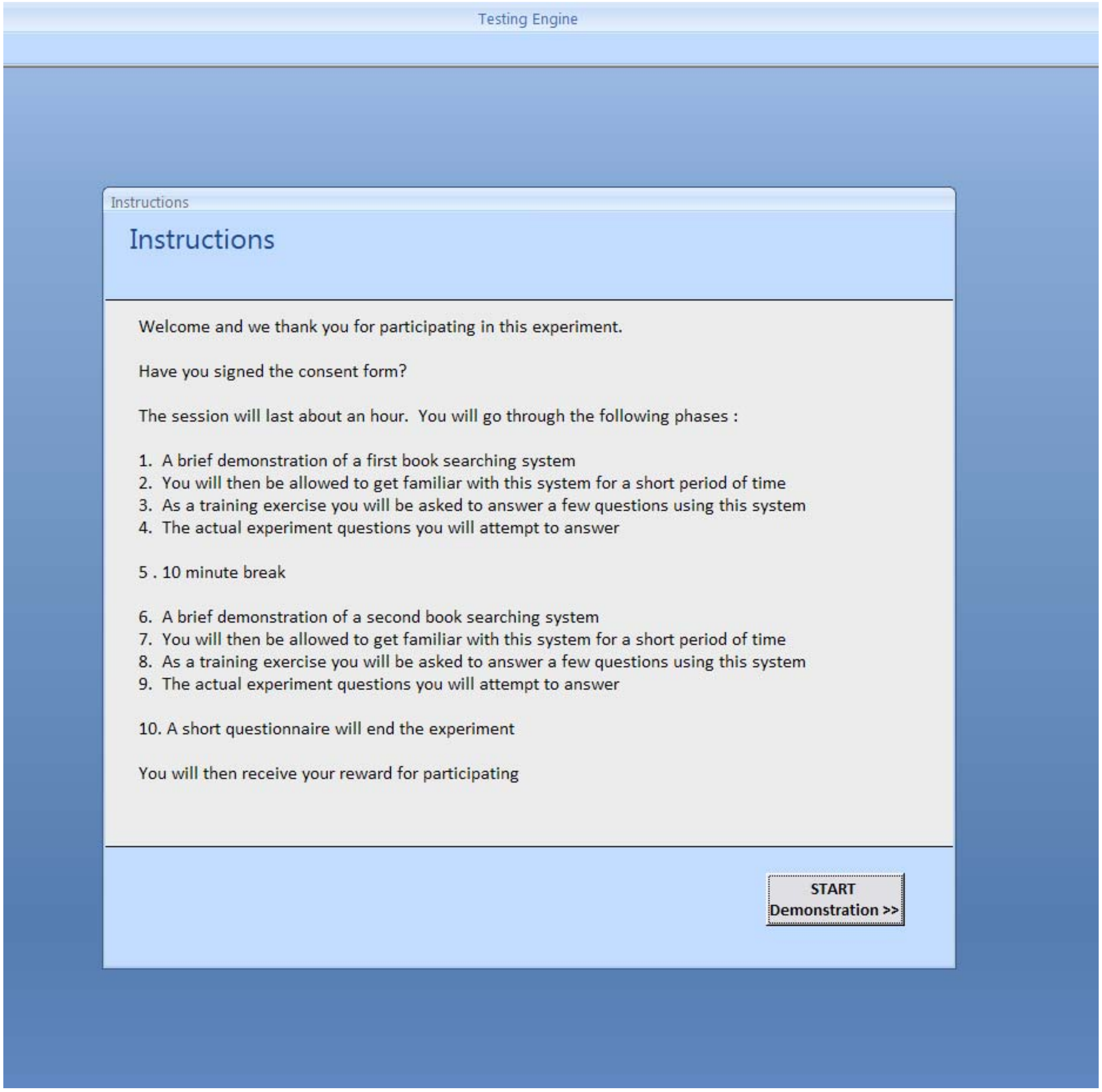


Figure 7.4: Test Engine Instruction Page

3. **Start of Training Slides**: User clicks on "START Demonstration" and flips through the PowerPoint training presentation guided by the researcher of the first system (i.e., Web baseline or SE-3D). Supported by the researcher, the participant replicates the slide demonstration on the currently tested system.
4. **Start of Training Test**: Once the user has viewed all the training slides he/she can view the slides again or choose to begin the training test questions.
5. **Training Test**: The user chooses to start the training test. The screen shown in Figure 7.5 is presented to the user.

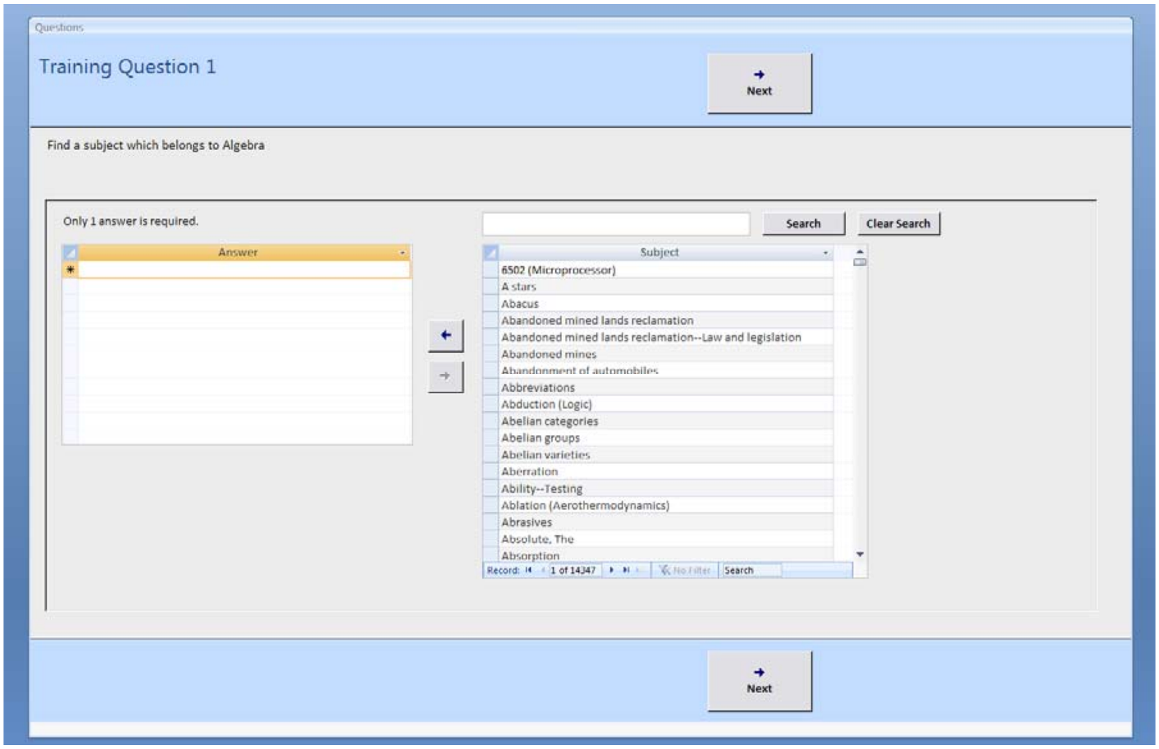


Figure 7.5. Testing Engine Question Entry for a Subject

Figure 7.5 shows the sequence of the current question at the top left and the question underneath. The user was meant to find a subject or book using the test system, locate it in the list of possible answers on the right, and transfer it to the answer list on the left using the arrows located between the lists. Participant had to enter the exact words starting from the beginning of a title (i.e., the left) for the search to function properly. This was meant to prevent participants from using the search box to find answers without using the tested interface.

Pressing the "Next" button brought the user to the next question if the current question had been answered. If the user had not entered any or enough answers, the system informed the user that the question was unfinished and asked if he/she still wanted to continue to the next question.

6. **Confirm Test Start**: Once the training questions are completed the user is asked to confirm he/she is ready to start the actual test.
7. **Test questions** are shown and answers recorded using the same interface as the training questions (see Figure 7.5). Completion time for each task is recorded from the moment the participant is shown the question to the moment he/she moves to the next question.

8. **End of Part 1- Break**: Once the first system questions are completed the user is forced to take a ten minute break.

9. **2nd System**: the user goes through steps 3 to 7 for the second system.

10. **Post-Test**: Once the 2nd system questions are complete the user is asked to complete the post-test questionnaire (see Figure 7.6).

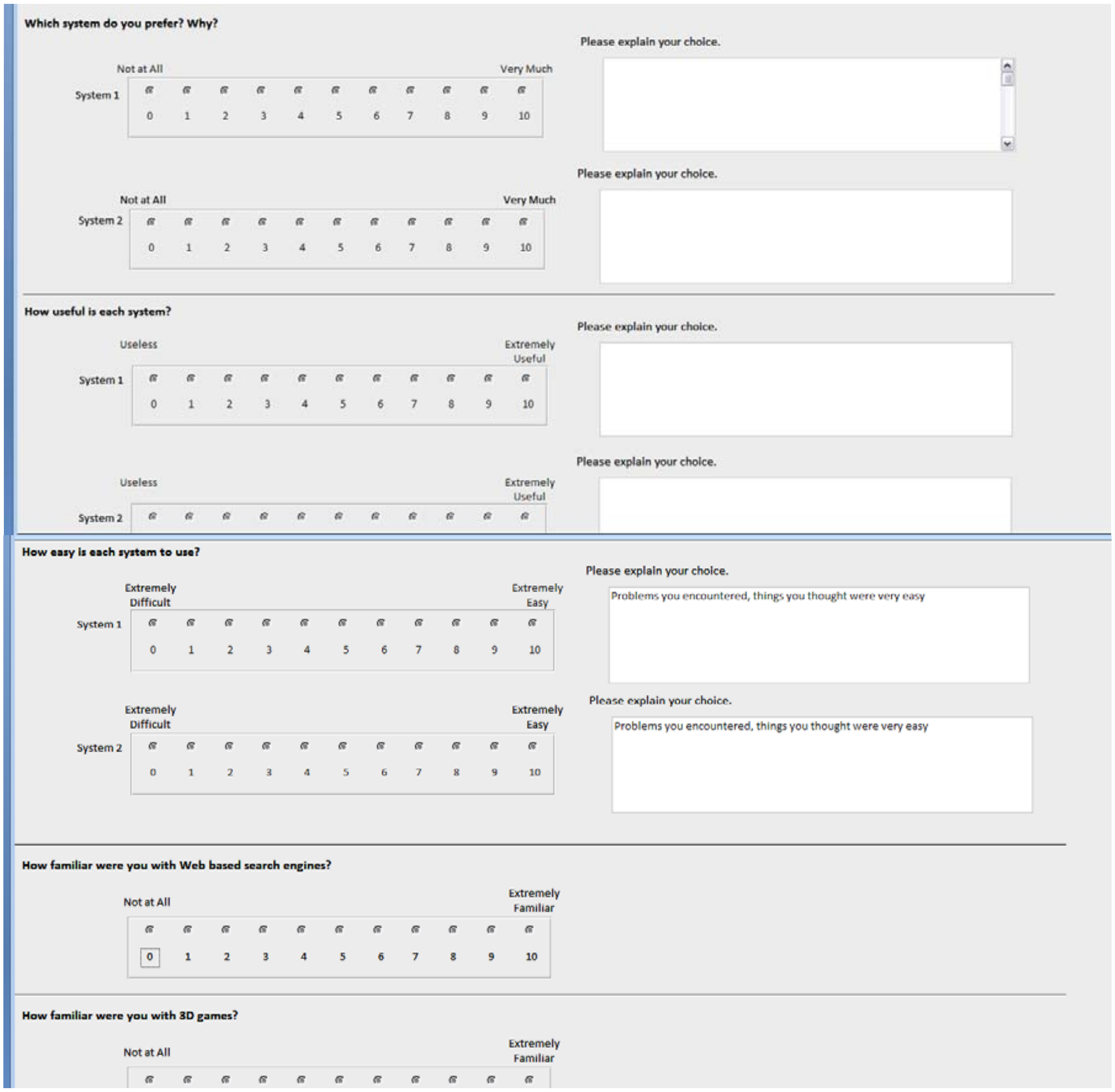


Figure 7.6. Test Engine Post-Test Questions 3 to 5

11. **End of Test**: The experimental session ends and the participant receives the monetary compensation.

7.5 Pilot-Tests

Pilot test objectives were to :

- practice and refine the experimental procedure to keep the session close to 60 minutes,
- gather user impressions of the procedure and the systems involved, and
- ensure stability of test engine, baseline Web system, SE-3D and their interactions.

The pilot tests were scheduled over a period of a month. Issues raised during a session were addressed before the next pilot session was scheduled. This allowed each new pilot participant to reveal unknown usability issues. By the time the third piloting session was completed the experimental design and systems had demonstrated their dependability.

7.6 Hypotheses

Visual IR interfaces are often "not immediately more useful than a traditional text list interface" (Cribbin & Chen, 2001, p. 207). This suggested non-directional hypotheses (i.e., no prediction as to which system will show better performance). IV testing research surveys showed "how infrequently information visualizations have been user tested and how poorly they fared when they were" (Morse, et al., 2002, p. 31).

There was likely to be a strong effect of the task type on measures of performance (see section 3.2.8.4 *Tasks*). There were indications that 3D Cone trees (Robertson, *et al.*, 1991) might "perform relatively better in tasks such as 'find the most densely populated directory' or 'find the deepest directory' (Cockburn & McKenzie, 2000, p. 434). This suggested a positive effect of SE-3D for hierarchy navigation tasks.

Hypotheses concerning simple and complex retrieval tasks were difficult based on the literature; however, IV seemed to be generally more suited to browsing tasks with unfamiliar domains (see section 3.2.3 *Tasks*). Smith *et al.* (2006) developed and tested a visual interface for user organized personal information collections (e.g., contents of My Documents folder in Microsoft Windows). They hypothesized that a traditional textual search box interface would be "difficult to beat for more textual, targeted searches" but they hoped their IV interface would be superior for "exploratory" tasks (p. 802). Their results supported the former but no effect was found for the latter. This suggested no effect of SE-3D for simple retrieval tasks but a positive effect for complex retrieval tasks.

Studies have reported users preferring an IV interface over a purely textual equivalent without any measurable performance advantage (see section 3.2.2 *Cognitive Affect*). SE-3D could be preferred and perceived as easier to use and more useful if correlations suggested by Davis (1989) hold true. These correlations were established in the context of a work setting where performance is paramount; therefore, they may not be found in an experimental setting where time pressures are not specifically applied. User preference for the IV interface was expected and correlations between post-test variables would be tested.

Based on these reports and an extensive review of IV evaluations (see review by Julien, et al., 2008), initial hypotheses are summarized and listed in Table 7.5.

Hypothesis	Null Hypothesis
H_1: Time will differ between systems for all task types	H_{1null}*: Time will not differ between systems for all task types
H_2*: Accuracy will be better for SE-3D for hierarchy navigation tasks	H_{2null}: Accuracy will not differ between systems for hierarchy navigation tasks
H_3: Accuracy will differ between systems for simple retrieval tasks	H_{3null}*: Accuracy will not differ between systems for simple retrieval tasks
H4*: Accuracy will be better for SE-3D for complex retrieval tasks	$H4_{null}$: Accuracy will not differ between systems for complex retrieval tasks
H_5*: SE-3D will be preferred as compared to the baseline	H_{5null}: Preference will not differ between systems
H_6*: Perceived ease of use and usefulness will be correlated with preference	H_{6null}: Perceived ease of use and usefulness and preference are not correlated

Table 7.5. Experiment hypotheses, asterisks (*) indicates anticipated results

7.7 Results

Table 7.6 is a simplification of Table 7.2 (i.e., the exact tasks tested) which provides an overview reminder of the tested tasks referred to in the following section.

Task	Question	Comment
1	Find a specific subject close to the top of the hierarchy	
2	Find the closest common parent to two subjects	The order of these questions was randomized.
3	Which one of the following 4 subjects contains more documents	
4	Find a specific subject close to the bottom of the hierarchy	
5	Which one of the following 4 subjects is more specific	
6	Return to the subject asked in number 1	
7	Simple Retrieval 1	The order of these two was randomized and measures were averaged.
	Simple Retrieval 2	
8	Complex Retrieval 1	The order of these two was randomized and measures were averaged.
	Complex Retrieval 2	

Table 7.6: Quick reminder of controlled experiment tested tasks

The discrepancy with the exact list of tasks shown in Table 7.2 is explained by the combination of tasks 7 and 8 averaged to a single simple retrieval result, and tasks 9 and 10 also averaged into a single complex retrieval result. In a perfect setting, measures would be averaged for all three task types (i.e., hierarchy navigation, simple retrieval and complex retrieval). This was not feasible since some of the hypotheses (see section 7.6) suggested there may be varying effects within the hierarchy navigation tasks. The experimental objective was to determine if the tested systems offered different performances for the tested task types as opposed to finding which specific task created the difference.

7.7.1 Time

Figure 7.7 presents the average time per task for each system.

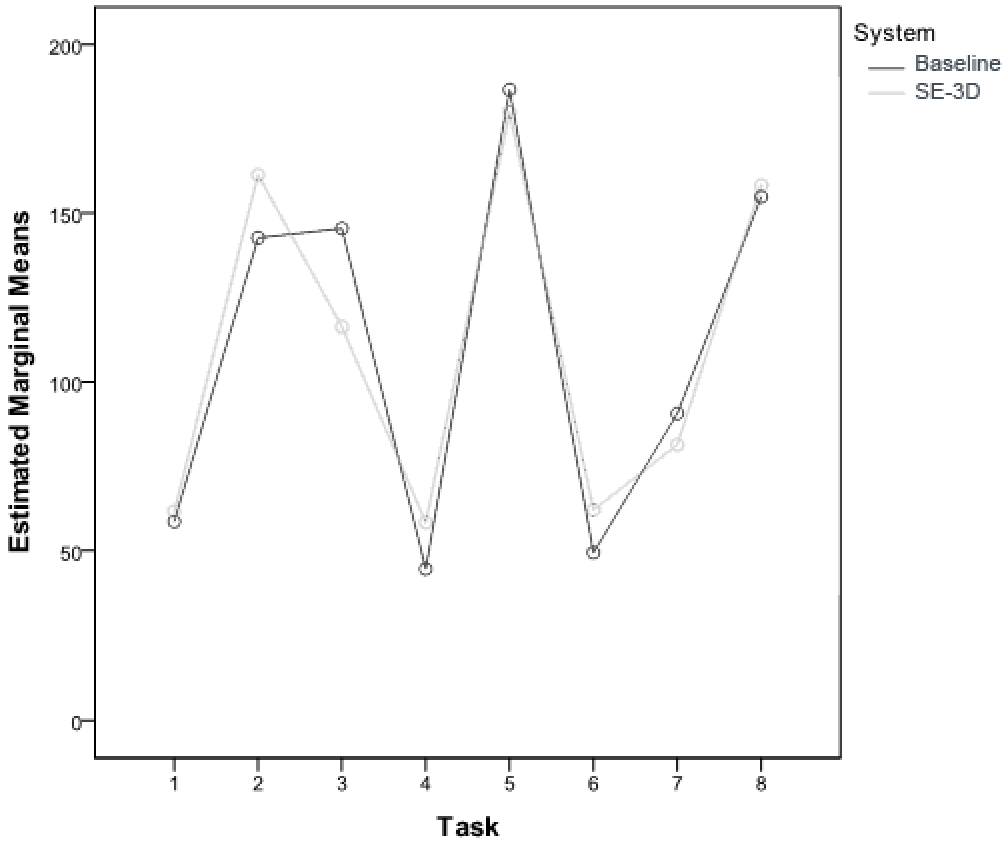


Figure 7.7: Line graph of time average time per task.

Figure 7.7 shows that average times for the individual tasks do not seem to be different between the two systems. The results were submitted to a 2 (System) x 8 (Task) x 2 (Question Set Order) x 2 (System Order) mixed ANOVA, with the first two factors as within subjects and the last two as between subjects variable.

Significant effects of task ($F(1,20) = 51.97$, $p = 0.000$) and interaction between System and System Order ($F(1,20) = 4.67$, $p = 0.043$) were observed. The effect size (partial eta squared = 0.189) of the interaction was weak (Cohen, 1988). The interaction effect suggests participants familiarized themselves with the testing engine and procedure. Their speed would increase slightly for the second system, no matter which one it was. This is in part attributable to participants becoming more efficient at using the testing engine. The random even distribution of system order amongst the participants controls the effects of this specific type of general learning. The lack of significant effects of task set order shows both task sets were equivalent.

These results do not allow the rejection of H_{1null} which means users of SE-3D did not suffer a time penalty as they familiarized themselves with the novel software. Inspection of actual training times showed no significant difference between the systems ($t(23) = 0.563$, $p = 0.579$, paired, 2-tail). This suggests differences between systems cannot be attributed to differences in training; instead, SE-3D was just as easy to learn and use as the traditional Web baseline.

7.7.2 Accuracy

Accuracy of hierarchy navigation tasks required no human judgment. This was not the case for simple and complex information retrieval tasks that relied on the subjective concept of relevance. Judgments of answer relevance were initially made by the researcher and later validated by a professional librarian. The following section describes this process and the results of the statistical analysis of accuracy measurements.

7.7.2.1 Accuracy Validation

For every question answered by each subject, the researcher evaluated the likely potential relevancy of each answer. Relevancy of books for a simple and complex retrieval tasks was determined using strictly the limited information provided by the title and series statement when available. The instructions given to participants and the validating librarian were to judge if a book was promising enough to request further details.

Table 7.7 presents numbers of unique books generated by all answers and their relevance assessments by the researcher and the validating librarian. The "Group Unique Books" column shows that the participant answers often included the same few books judged relevant by both evaluators.

Question Type	Question	Group Unique Books	Positive Relevance Judgments		
			Researcher	Librarian	Delta
Simple	You are looking for information on game design.	7	7	6	1
	You are looking for information on Visual Basic programming.	8	8	6	2
	You are looking for information on climate change.	18	18	18	0
	You are looking for information on internet security.	10	10	10	0
Complex	What kind of trees are used for construction of buildings?	27	23	18	5
	How would you go about fixing a broken computer?	16	14	8	6
	Why do concrete pipes eventually fall apart?	13	13	13	0
	How can you fix a leaky faucet?	11	11	5	6

Table 7.7: Relevance evaluations for retrieval tasks

Table 7.7 also shows that the relevance evaluations were most contentious for complex retrieval tasks. The difference between positive relevance judgments from researcher versus librarian was five or higher for three out of four tasks.

Table 7.8 shows that complex retrieval tasks were different from simple retrieval. The former shows a larger average number of unique answers (i.e., books) and standard deviation.

Task Type	Average Number of Books per Task	Standard Deviation
Simple	11	5
Complex	17	7

Table 7.8: Descriptive statistics of answered books by tasks type

For each task, each unique book answer may have been selected by multiple test participants. A power law distribution of answers was found for each question; in other words, a few answers were selected by many participants and most answers were selected by very few.

The frequency of book selection by participants was used to weight the relevance judgments by both researchers and validating librarian. This was followed by the analysis of differences between the two sets of judgments for each task. For example, Table 7.9 shows the weighted relevance judgments for one complex retrieval task (i.e., leaky faucet). It shows a large difference between absolute relevance judgments (11 vs. 5) but a much smaller difference when these are weighted by relative importance of the books based on their answer frequency (43 vs. 36).

How would you fix a leaky faucet?		**Relevance (1=Relevant; 0=Not Relevant)**		**Frequency Weighted Relevance**	
Unique Book Answers	**Frequency**	**Researcher**	**Librarian**	**Researcher**	**Librarian**
The plumbers handbook /by Joseph P. Almond, Sr.	15	1	1	15	15
Plumbing instant answers /R. Dodge Woodson.	14	1	1	14	14
Plumbing :installation and design /Thomas Philbin.	5	1	1	5	5
Uniform plumbing code illustrated training manual /International Association of Plumbing and Mechanical Officials. --	2	1	0	2	0
2003 international plumbing codes handbook /R. Dodge Woodson.	1	1	0	1	0
Advanced plumbing /Harry Slater & Lee Smith. --	1	1	0	1	0
Carroll Smith's nuts, bolts, fasteners, and plumbing handbook /Carroll Smith.	1	1	1	1	1
Options for leak and break detection and repair for drinking water systems /Lawrence A. Smith ... [et al.].	1	1	0	1	0
Plumbing dictionary /edited by A.S.S.E. Plumbing Nomenclature Committee ; J. Russell Boates, chairman. --	1	1	0	1	0
Practical plumbing design guide /James C. Church. --	1	1	1	1	1
The water supply of buildings, and rural communities, for engineers, architects, plumbers, and property owners.	1	1	0	1	0
Totals	43	11	5	43	36

Table 7.9: Example of weighted relevance judgments used for pair-wise t-tests

For each question, pair-wise t-tests were performed between the frequency weighted relevance judgments of the researcher versus the validating librarian. In general, both researcher and validating librarian agreed on the relevance of the top few most frequently answered books for each questions. This explains why only two of the eight tasks showed a significant difference between the weighted relevance judgments: the leaky faucet (t(23)= 2.23, p=0.01, paired, 2-tail) and the broken computer task (t(23)= 2.13 p=0.01, paired, 2-tail).

Figure 7.8 presents the distribution of answer frequencies for the leaky faucet task. It shows the top three most answered books for this question were judged relevant by both researcher and validating librarian and the total agreements included almost 84% of answers.

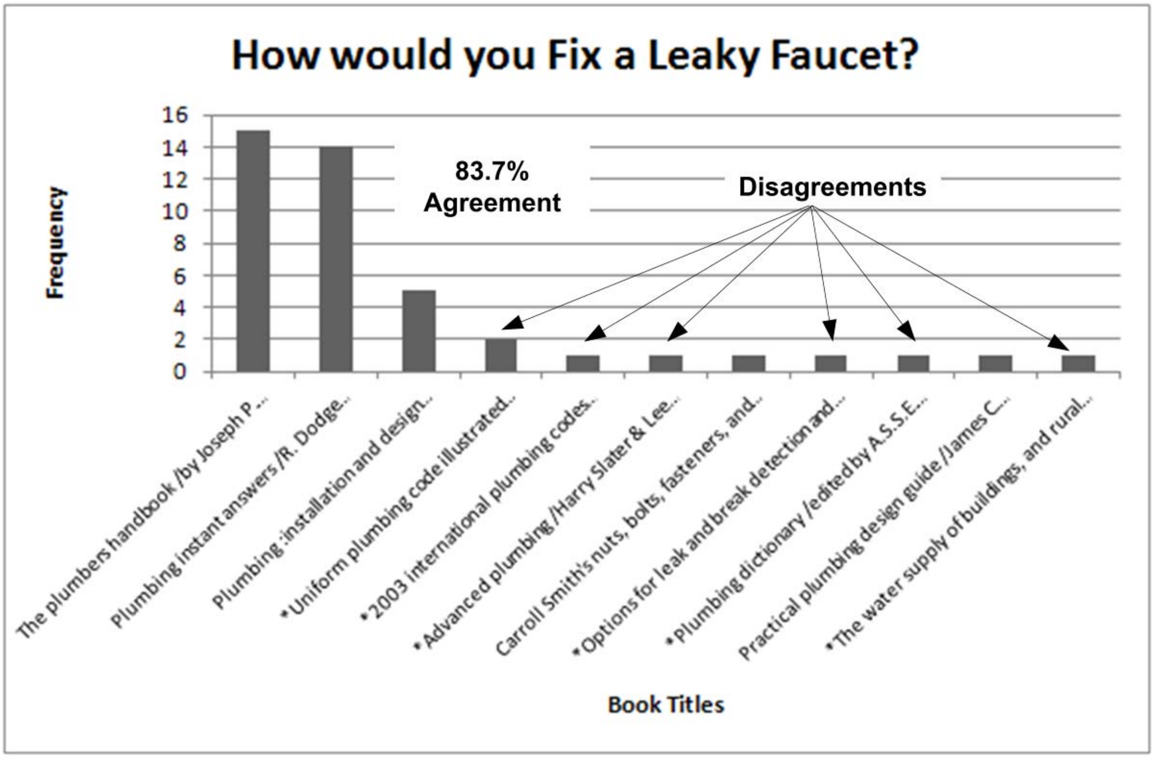


Figure 7.8: Leaky faucet frequency of answers and relevance judgment

A similar phenomenon was present for the broken computer task as shown in Figure 7.9.

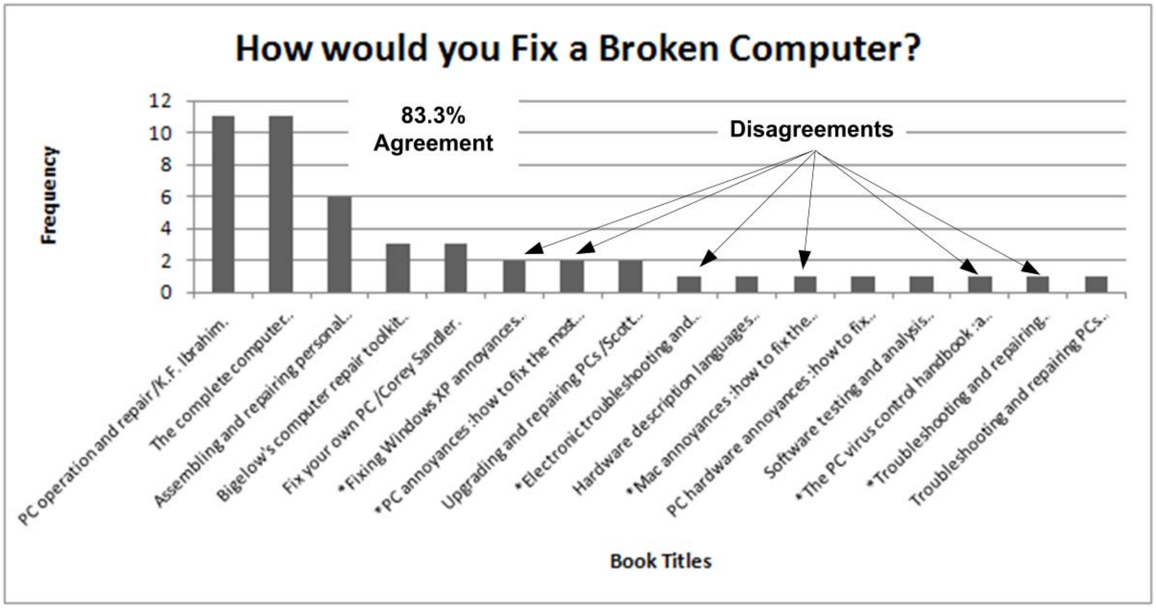


Figure 7.9: Broken computer frequency of answers and relevance judgment

The following conclusions can be drawn from the information presented in this section:

- Researcher vs. validating librarian relevance evaluations were not significantly different for 6 out of 8 tasks
- The two tasks which showed a significant difference still produced an agreement for over 83% of books answered. This is explained by the fact that, for all tasks, relevance evaluations were congruent for the few most frequently answered books.

Relevance is partly subjective (Chowdhury, 1999; Chu, 2003) depending on the user interpretation of the IR task. For this reason, the more generous relevance judgments made by the researcher were used for the statistical analysis presented in the following section.

7.7.2.2 Statistical Analysis

Task 7 (i.e., simple retrieval) accuracy was 100% on both systems which does not allow rejection of H_{3null}. This was the anticipated result (see section 7.6) since these tasks were directly answered using the traditional keyword search and textual result list which was available in both systems.

Task 3 (i.e., more documents) accuracy was also 100% on both systems. No specific hypothesis was made concerning this task, but H_2 stated participants would perform better using SE-3D for hierarchy navigation tasks in general. This accuracy ceiling suggests that some questions may have been too easy. Detailed question design is difficult; there is no standard set of IV evaluation tasks, and there is a wide variety of tested tasks in IV controlled experiments (see section 3.2.8.4–Tasks). This accuracy ceiling showed that participants performed equally with SE-3D and the more familiar textual baseline.

Since tasks 3 and 7 had no difference between the systems they were not considered for the subsequent analysis. Table 7.10 lists the reduced set of tasks used in this section.

Task	Question
1	Find a specific subject close to the top of the hierarchy
2	Find the closest common parent to two subjects
3	Find a specific subject close to the bottom of the hierarchy
4	Which one of the following 4 subjects is more specific
5	Return to the subject asked in number 1
6	Complex Retrieval

Table 7.10: Reduced list of tasks with only those with difference in accuracy

Figure 7.10 presents the average accuracy per task for each system. It shows a clear separation between the top line (SE-3D) and the bottom line (baseline).

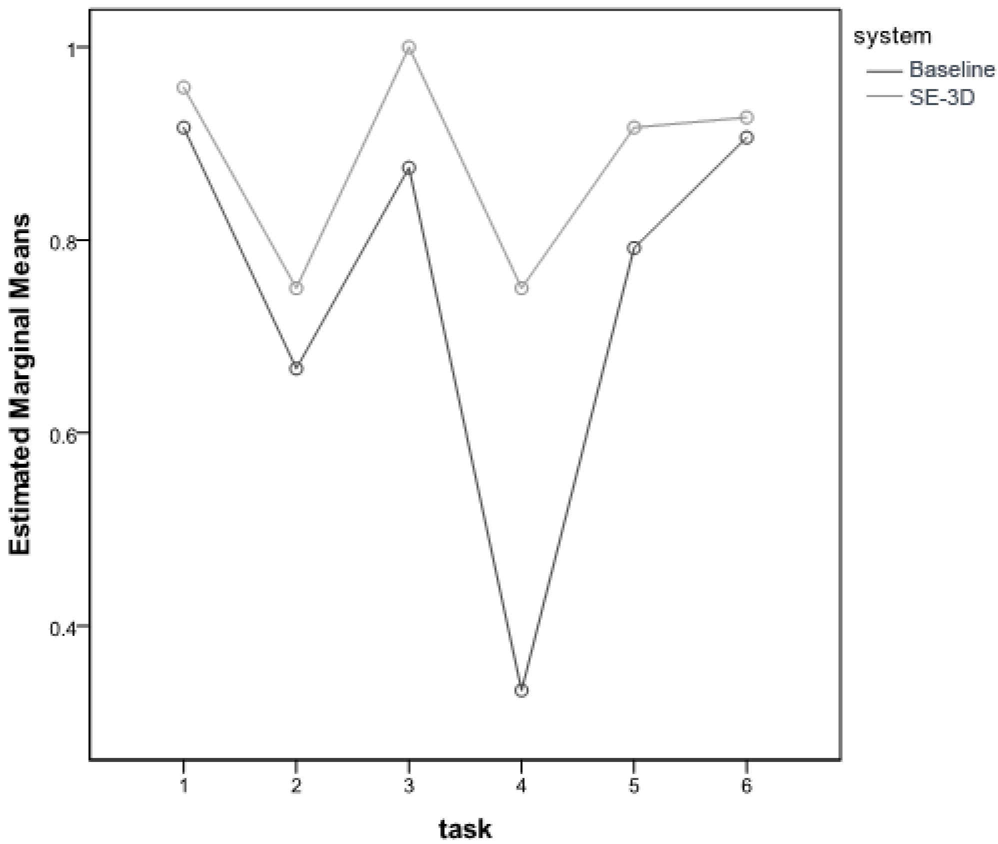


Figure 7.10: Graph of average accuracy for the 6 tasks with difference accuracies

The results were submitted to a 2 (System) x 6 (Task) x 2 (Question Set Order) x 2 (System Order) mixed ANOVA, with the first two factors as within-subjects and the last two as between-subjects variable. Sphericity was not respected for task and system*task interaction; the Greenhouse-Geisser correction was used to guard against type I errors (i.e., false positives). Significant effects of system ($F(1,20) = 13.425$, $p = 0.02$) and task ($F(1,20) = 12.032$, $p = 0.000$) were observed. According to Cohen (1988), the effect size of system (partial eta squared = 0.402) and task (partial eta squared = 0.376) can be considered medium.

A main effect of system without interaction with task suggests a uniform difference in accuracy across all tasks; however, this was explicitly tested by t-tests for each pair of system scores for each of the six tasks. A significant difference for task 4 (i.e., more specific) was observed ($t(23)=-3.500$, $p=0.002$, paired, 2-tail) with a small to medium size of effect (partial eta squared = 0.347). This result corresponds with the visual inspection of Figure 7.10 which shows a large difference for task 4 and relatively much smaller differences for all other tasks.

Table 7.11 provides mean accuracies and their distributions for both systems. For the six tasks listed in Table 7.10, Table 7.11 shows that the significant effect of system on accuracy translates to an average improvement for SE-3D of 13.6% (0.884–0.748).

System	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Baseline	.748	.040	.665	.831
SE-3D	.884	.024	.834	.933

Table 7.11: Descriptive statistics for the 6 tasks with different accuracies

7.7.2.3 Interpretation

Significant differences between system performances were not assured. There is a general lack of significant performance differences between the IV system and a text-only baseline (Chen & Czerwinski, 2000a; Julien, *et al.*, 2008). The results of this experiment are described below for each of the individual tasks tested.

7.7.2.3.1. Find Most Specific Subject

Task 4 asked users to evaluate the relative specificity of subjects. The baseline system offered no specific information for this type of task. Using the baseline Web system, a user had to successively navigate up the hierarchy until the top was reached, while counting the number of broader terms traversed. This was a tedious, error prone process and users were observed relying on the number of items assigned to a subject as an indication of its specificity. This tactic assumed that fewer items indicated a more specific subject. This might be intuitive but the baseline accuracy results show it was often false.

SE-3D offered explicit cues to the relative specificity of each subject. Searchers could observe that a subject situated higher in space was broader, and each subject map contained a label indicating the current abstraction level within the hierarchy. Either of these cues may explain the significant performance advantage provided by SE-3D for judging the relative specificity of subjects.

7.7.2.3.2. Find Nearest Common Parent

The lack of significant difference in task 2 (i.e., find nearest common parent) was surprising. Answering this type of question using the baseline system forced users to find each subject, then navigate up through their broader parents until they found a common one. Since the baseline only showed one hierarchy level at any one time, this forced users to start with one subject and remember or write down all the visited levels. This procedure was then repeated with the second subject until the user recognized a parent in common with the first subject. SE-3D allowed viewing of multiple hierarchy levels and broader terms were always visible all the way to the highest level of abstraction; therefore, an advantage was anticipated.

The specific questions designed for this task type may partly explain the lack of significant difference between the systems. These questions imposed many constraints on the choice of the two child subjects. Specifically:

- Both had to be unique non-replicated subjects. This left 1050 of the 1828 subjects shown in the visual structure.
- Of those, finding a valid pair meant finding two subjects which were not directly related by a succession of broader terms. This eliminated 965 pairs of subjects.
- Of the remaining pairs, both subjects had to have a common distant parent.

These constraints left few choices; as a result, both task set questions of this type asked for descendents of the broad subject of "Science". This broad subject was in fact a parent to the vast majority of this specialized collection. Observations made during testing showed that, when in doubt, some participants would guess "Science" since it was clearly a parent to most of the collection. They were not certain this was the lowest most specific answer required but at the very least they knew it was a common parent. Porting SE-3D to a different collection may allow mitigation of this issue. This will be the focus of future research.

7.7.2.3.3. *Specific Subject Searches*

Neither tasks 1 (i.e., find a subject situated near the top of the hierarchy) nor 5 (i.e., find a subject situated near the bottom of the hierarchy) showed a significant difference in accuracy between the systems. This was not surprising based on observations made during the tests which showed that for this task type, participants predominantly used the search box and result list available in both systems. Keyword searching is best for these kinds of known item searches (see section 3.1 - Information Retrieval). Returning to a previously visited subject (i.e., task 5) could arguably be a kind of known item search also well supported by keyword searching.

7.7.2.3.4. *Complex Retrieval*

Hypothesis H_{4null} cannot be rejected since the paired t-tests showed no significant difference for task 6 (i.e., complex retrieval). This finding was surprising since SE-3D aimed to facilitate browsing behavior often associated with broad subject search in unknown domains (see section 3.1.1–Browsing). This confirmed the known difficulties evaluating these types of browsing tasks; as described by Ellis *et al.* (2006)

> *"Visualisations are often at their best for more exploratory tasks, but these are precisely the tasks that are hardest to replicate in an experiment" (p. 3).*

7.7.3 Post-Test

The post-test questionnaire contained the following types of data:

Question	Likert Scales (1-10)	Qualitative Comments
1. Speed of each system	Baseline	Explain your choices
	SE-3D	
2. Usefulness of each system	Baseline	
	SE-3D	
3. Ease of use of each system	Baseline	
	SE-3D	
4. Preference for each system	Baseline	
	SE-3D	
5. Familiarity with search engines	Single Likert	
6. Familiarity with 3D games	Single Likert	

Table 7.12: Post-test data collected

This provided four pairs of data series (i.e., questions 1 to 4) which could be evaluated for correlations and significant differences. The following questions would be tested:

1. Are there differences across systems in terms of perceived speed, preference, usefulness, or ease of use?
2. Is there a correlation between perceived speed of the system and the actual completion times measured during the experiment?
3. Is there a relationship between familiarity with search engines or 3D gaming related to the accuracy for either system?

Results are described below.

7.7.3.1 Significant Differences between Systems

Figure 7.11 presents means (and standard errors) for questions 1 to 4. It shows that there was a 3.17 (8.21−5.04) difference in preference favoring SE-3D. It was also perceived as being more useful than the textual baseline by a difference of 1.79 (8.08−6.29).

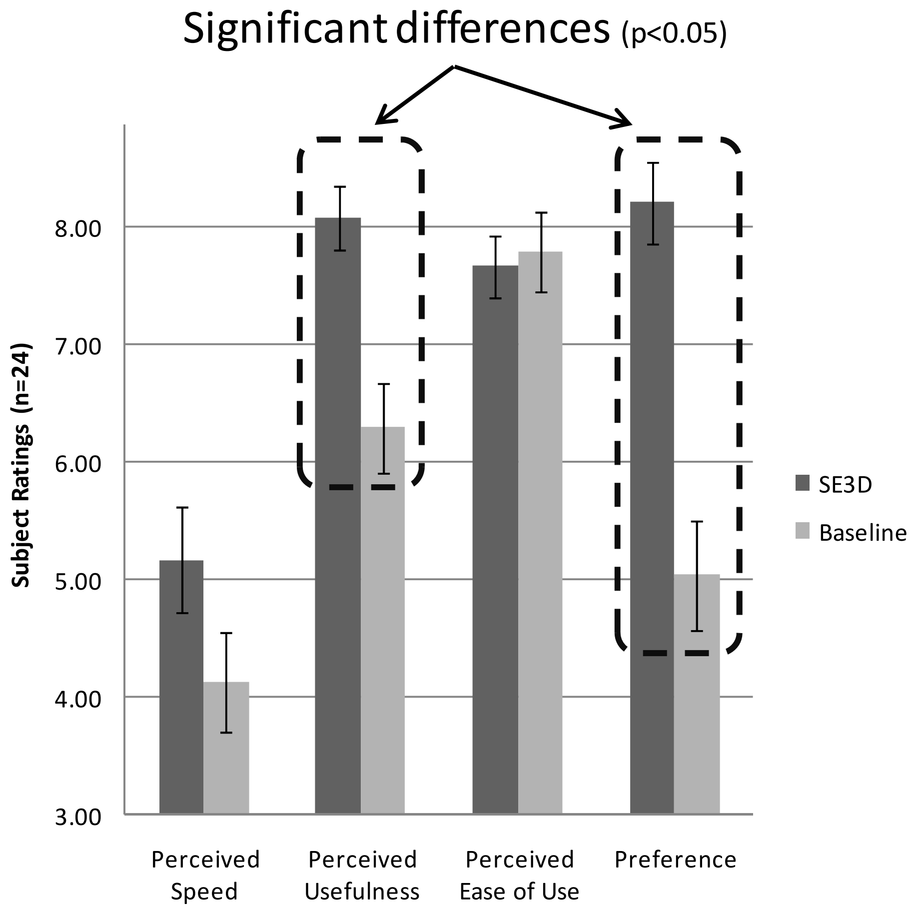


Figure 7.11: Graph of mean ratings for Post-Test questions 1 to 4

Paired sample T-tests between the two systems revealed that preference (t(23)= -4.789, p = 0.000, paired, two-tail) and usefulness (t(23)= -4.343, p = 0.000, paired, two-tail) differed significantly between systems. These results allow rejection of H_{5null} and support H_5 which states that participants would prefer SE-3D as compared to the baseline.

A correlation between actual usage and perceived usefulness was reported by Davis (1989). A correlation analysis for the baseline system showed a significant positive correlation (r= 0.696, p = 0.000 < 0.01) between preference and usefulness. The same analysis for the equivalent SE-3D variables also showed a significant positive correlation (r = 0.688, p = 0.000 < 0.01). A correlation analysis was performed between scores for ease of use and usefulness. This revealed no correlation between participant's ratings for these concepts. This was contrary to expectations stated by H_6 and did not allow rejection of H_{6null}.

These findings suggest that the known baseline web system and the novel SE-3D were perceived equally easy to use and responsive. This confirmed that SE-3D development and testing efforts had produced a usable tool with adequate speed of operation.

Participants preferred the system perceived as most useful. This would suggest SE-3D facilitated more kinds of tasks than the baseline web system. Participants may have appreciated that SE-3D included keyword searching in addition to visually enhanced subject hierarchy browsing features; in other words, SE-3D was a traditional keyword searching engine and more. These results suggest that a novel system may be preferred to a known tool if it offers equivalent ease of use and speed, and demonstrate how it supports more kinds of tasks.

7.7.3.2 Correlation between Perceived System Speed and Time Measures

For each system, the perceived system speed as reported by the participants versus measured average time per task were plotted. These are shown in Figure 7.12and Figure 7.13).

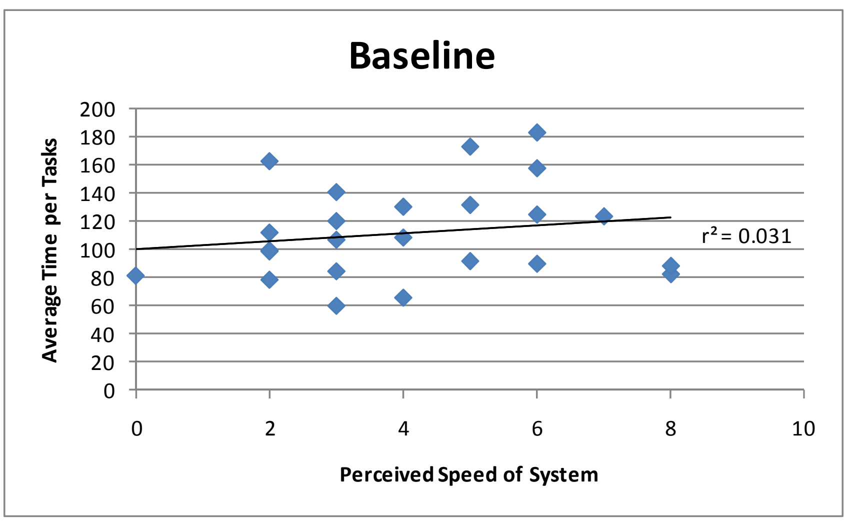


Figure 7.12: Scatter plot of baseline perceived speed vs. average times per task

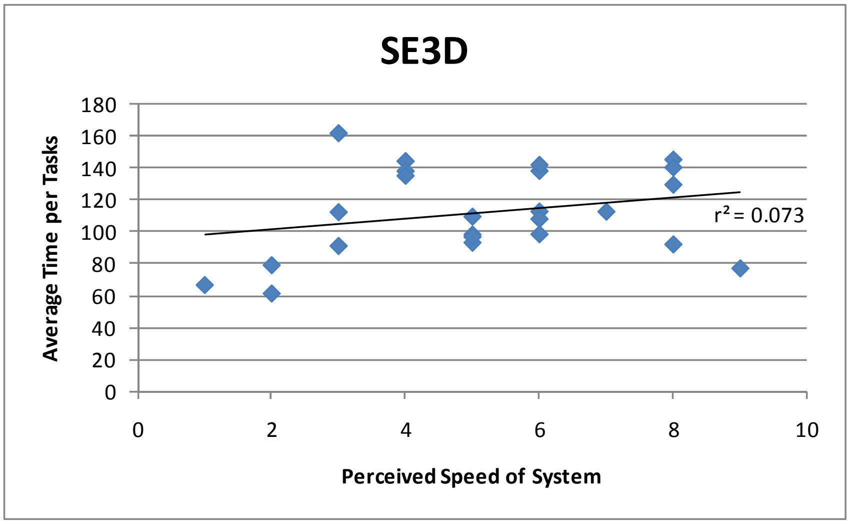


Figure 7.13: Scatter plot of SE-3D perceived speed vs. average time per task

Linear relations between these variables seemed unlikely. The coefficient of determination r^2 estimates how well the data might fit a predictive linear model (i.e., the trend line on the graphs). Values close to one suggest the data fits the model perfectly. r^2 values of 0.031 (baseline) and 0.073 (SE-3D) are very low and do not suggest the presence of a linear relation. Pearson's correlation coefficients showed little correlation between the variables (see Table 7.13).

System	Perceived Speed (Likert, 0-10)	Average Time per Task (Seconds)	r^2	Pearson's *r*
Baseline	4.13	111.8	0.031	0.176
SE-3D	5.16	111.9	0.073	0.270

Table 7.13: Comparison of perceived vs. measured speed

These numbers suggest there is no obvious linear relation between actual and perceived speeds of the systems.

7.7.3.3 Correlation between Familiarity and Accuracy

Familiarity with web based search engines was assessed using an 11-point Likert scale. Figure 7.14 presents the distribution of answers and shows participants felt they were generally very familiar with web based search engines.

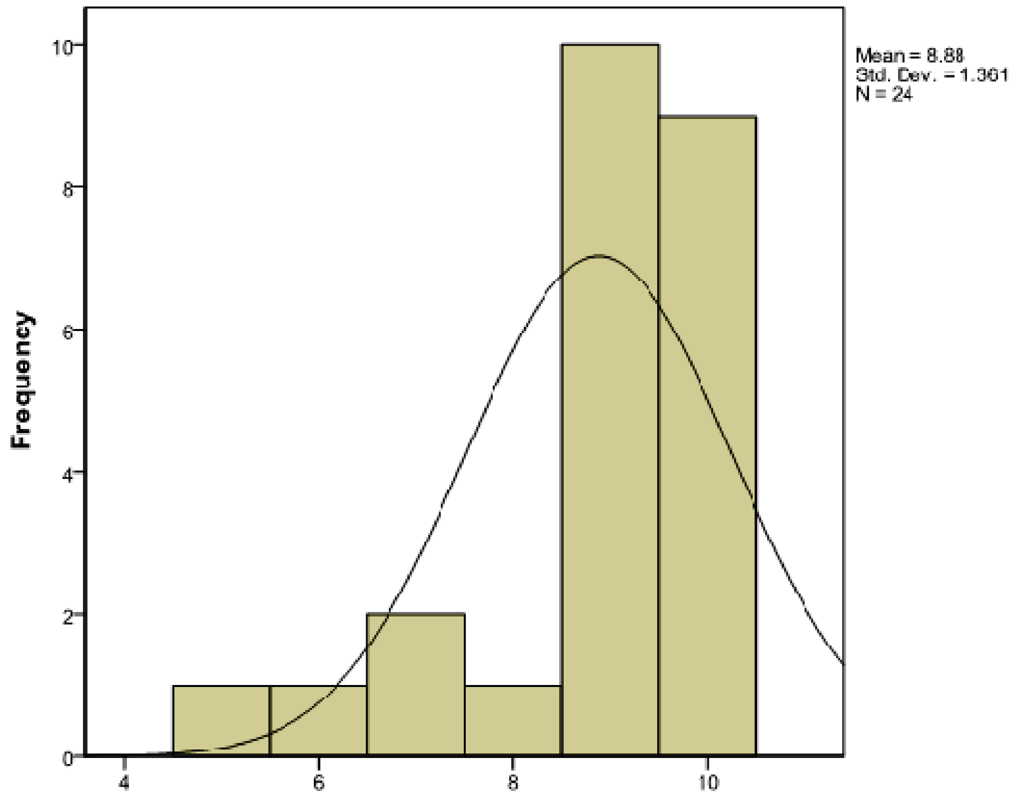


Figure 7.14: Distribution of Familiarity with Search Engines

There was little correlation between average baseline accuracy measures and search engine familiarity (Pearson's $r = 0.205$).

Familiarity with 3D games was also collected in a similar fashion and resulted in the distribution shown in Figure 7.15. It reveals three distinct groups of users with self-assessed low, medium and high familiarity with 3D games.

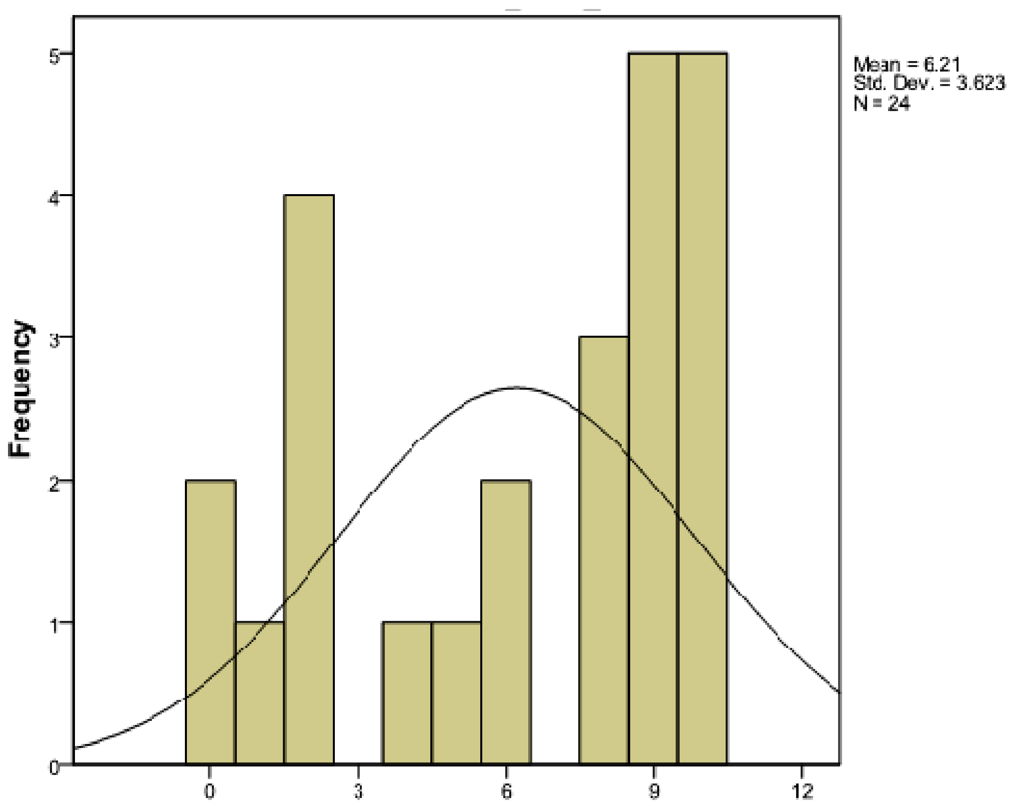


Figure 7.15: Distribution of Familiarity with 3D Games

There was little correlation between accuracy using SE-3D and familiarity with 3D game (Pearson's r = 0.114). Large, Beheshti, Tabatabaei & Nesset (2009) found a negative correlation between children's retrieval time and familiarity with video games. It would be interesting to create three groups of participants based on their self-assessed familiarity with 3D games (i.e., low, med, high familiarity), and perform an analysis of variance for accuracy or completion times. This will be the subject of future analysis.

These results show that familiarity with search engines did not translate into high accuracy with the baseline web search system, and self assessed familiarity (or the lack of it) with 3D games did not translate into better or worse accuracy with SE-3D.

7.7.3.4 Qualitative Analysis of Comments

Participants had the option to freely explain each of their Likert scale selections. Qualitative comments provided a rich set of unstructured prose containing information not captured by the Likert scales. A preliminary analysis of this data was performed by identifying themes and marking each comment with every theme it concerned. Themes were revised and refined in three successive passes. This provided a rough but representative picture of the prevalent themes related to each system. Additional analysis is planned based on Guastavino & Katz (2004) but is beyond the scope of the current research.

Table 7.14 shows the number of comments collected per question and system, and the number of themes extracted. It shows most participants were willing to explain their choices; however they generated a greater variety and number of themes when commenting about SE-3D.

	Speed	**Useful**	**Ease**	**Preference**	**Total Comments**	**Unique Themes**	**Total Themes**
Baseline	20	21	22	24	87	12	87
SE-3D	24	22	22	24	92	20	132

Table 7.14: Numbers of Qualitative Comments and Themes Collected

For example, Table 7.14 shows that 96% (92/96) of possible comments were filled for SE-3D. From these, 20 unique themes were identified, these were found multiple times for 132 total themes.

The occurrences of each theme were tabulated in order to identify the most prevalent themes per system. Those which made up at least 5% of the total themes per system were deemed significant for further discussion.

7.7.3.4.1. Baseline System Qualitative Analysis

Table 7.15 presents the significant themes found for the baseline system. It shows that the first three make up 60% of the total themes extracted from baseline comments.

Themes	Occurrences	Prevalence	Examples
1. Simple and Known	26	30%	• "a classic way of finding books in a library" • "It is familiar, I know what to do and what to expect"
2. Search is Slow	17	20%	• "Very long fetch times" • "delay between the times you pressed search and the time you obtained your results"
3. Relationships Difficult	9	10%	• "the links between different subjects were not clear" • "it was difficult to follow the path of subjects, from less specific to more specific, or vice-versa"
4. Fast for Known-item	7	8%	• "I preferred this system for looking for specific books" • "If I know what I want than this is better"
5. Lost	5	6%	• "really hard to see where your search was going- I was lost" • " The whole system seemed like an endless maze and I had no feeling for where I was in the database: every screen looked the same."
6. Crashes	4	5%	• "quick interactions and hasty corrections could crash the system" • "crashes if you double click search"
7. Keyword Search is Effective	4	5%	• "Fetched good results with just a keyword search" • "searching is good enough for books/subjects"
8. Multiple Parents	4	5%	• "It was confusing when each subject could have more than one more general subject" • "It was frustrating when a subject had more than one more general subject"
9. Subject Name vs. Subject Information	4	5%	• "I got confused whether I should click on the subject name or the subject information" • "The subject information link took a bit getting used to"

Table 7.15: Most common baseline concepts collected

Theme 1 "Simple and Known" concerned participants stating they recognized the baseline system as a familiar search box tool. This was the design goal behind the baseline system and this result shows it was met.

Theme 2 "Search is Slow" indicates participants noticed a delay between the time they initiated a search and the appearance of the results. Depending on the breadth of the search term(s) this operation could take up to three seconds. It is important to restate that the search engine powering the search box was exactly the same for the baseline and SE-3D; thus, both systems produced the same search result wait times.

Theme 3 "Relationships difficult" represents comments stating difficulties with finding and following relations between subjects that were critical for tasks 2 (i.e., common parent subject) and 5 (i.e., most specific subject) shown in Table 7.6. The baseline system offered no direct indication as to the current level of subject specificity and how many levels above or below users could expect to find. Text-based hierarchies are traditionally limited to showing the set of immediate children of a selected subject. The only way to determine the remaining depth of a subject branch was to systematically explore each successively more specific subject branch. This difficulty with subject relationships and single level hierarchy navigation may have contributed to a feeling of being lost identified in theme 5; users not remembering where they've come from and unsure of their current position or next paths to follow. Theme 8 "Multiple Parents" is also related with difficulties navigating subject relations; users had to choose between parents without knowing where each might lead.

Theme 6 "Crashes" refers to a programming error in the baseline interface which was detected after the testing started and kept for consistency. After initiating a search, this error would happen if the user pressed the 'Search' button a second time before the initial search was completed (i.e., before the results were shown). Double-clicking on the search button would also trigger this error, and since search times could take up to three seconds, some users impatiently clicked the search button. Participants were systematically warned they had to wait for the results to show before initiating a second search (i.e., wait for the hourglass to disappear); nonetheless, this error occurred for 6/24 participants and recorded times were adjusted.

7.7.3.4.2. SE-3D Qualitative Analysis

Table 7.16 presents the significant themes found for SE-3D. It shows that the first four themes made up more that 60% of the total themes extracted from SE-3D comments.

Themes	Occurrences	Prevalence	Examples
1. Good with Relationships	23	17%	• "it makes it more clear what subject the book falls under" • "the user can visualise the precise hierarchy"
2. Easy after Learning Curve	17	13%	• "Once you get a hold of how it works it's all pretty easy" • "Not hard but takes getting used to navigation keys and meaning of colours and visual symbols"
3. Post-Search Zoom-out Slow	15	11%	• "It took a while to see the visual representation of the library, when I wanted to move onto the next search" • "The transition between searches is a little bit slow"
4. Visual Good	14	11%	• " The whole visual aspect was the most useful part of the system. It allowed the user to have a more wholesome experience and understand better the hierarchy of the particular library." • " I loved how everything is visual. I could see where I was going"
5. Fun	8	6%	• "system is more fun to use" • "It is not boring" • " people will like to use the system because its different from the usual type-in-a-searchbox-press-enter-get-your-item system which is pretty mundane and all over the web"
6. Should be More responsive	8	6%	• "would be nice to be able to manually control how fast it goes" • "sometimes I wish I could manually control it, to stop zooming out."
7. Hidden Objects	7	5%	• "the titles may overlap and make it harder to understand but moving from side to side can fix that" • "Sometimes you needed to rotate to see all of the sub-categories, which was frustrating"
8. Best of Both Worlds	6	5%	• "It was very nice to see the list of search results and the library map at the same time." • "it catered to individual titles with the text box, while also catering to large topics with the 3D models."

Table 7.16: Significant SE-3D Concepts Collected

Theme 1 was the most prevalent theme and concerns how SE-3D clearly depicts relationships between subjects. This is followed by theme 2 which represents statements about the ease of use of SE-3D once an initial learning curve was completed.

Theme 3 "Post-Search Zoom-out Slow" was the direct result of a design decision to automatically zoom-out the user viewpoint after every search. The objective was to clearly show that the visible parts of the hierarchy had been modified by the keyword search filter. This animated movement quickly became anticipated and repetitive when performing successive searches. This is closely related with theme 6 "Should be more responsive" which concerned users stating they shouldn't have to wait for the animation to finish before triggering their next intended action.

Theme 4 "Visual Good" were positive statements concerning the highly visual nature of SE-3D. These positive reactions were often accompanied by descriptions of place and knowledge of explored areas. This might be associated with the theme of "Fun" which was often used in comparison with the "boring" and "usual" baseline.

Theme 7 "Hidden Objects" refers to overlapping labels in the 3D space. Thousands of overlapping subject labels will overwhelm users; as a result, SE-3D only showed the labels which were directly visible from the current user point-of-view (i.e., those that are not hidden). This forced users to navigate around the space in order to inspect all subject labels. Some users felt this was cumbersome or "frustrating" and wished they could simply inspect all the labels without having to move around the structure.

Theme 8 "Best of Both Worlds" were statements recognizing that SE-3D was a superset of the baseline search box. Six participants recognized that the search box and results list were identical in both interfaces. This made the interactive 3D model an additional visual feature as opposed to a replacement of the classic search box and result list. Some stated that visual browsing and keyword searching each had their own use within SE-3D.

7.8 Evaluation Conclusion

This chapter described a controlled comparative evaluation of a virtual reality information visualisation and retrieval tool. Participants performed a range of eight information retrieval tasks using the novel 3D IV for IR system versus an equivalent text-only baseline system. This was a repeated measures two-factor within subject design where 24 participants performed two sets of equivalent tasks on both systems, order of task set and system were counterbalanced.

7.8.1 Accuracy

Results showed that an IV application can provide an advantage for subject hierarchy navigation tasks. There are numerous hierarchy or tree navigation IV applications (see survey by Akrivi, *et al.*, 2007), but "few are evaluated" (Akrivi, *et al.*, 2006, p. 1). The few reports of evaluations (Akrivi, *et al.*, 2006; Plaisant, *et al.*, 2002; Risden, *et al.*, 2000) are difficult to compare since they use different baselines, tasks and experimental designs. In general, they do suggest an advantage of the visual system for tasks that require extraction of relationships between objects.

SE-3D produced a significant performance advantage for finding the most specific subject. This type of task required repeated evaluation of relations between subjects. This performance advantage may not be surprising since traditional OPACs represented by the baseline system provide little information about these relations. IV for IR may be most promising for supporting user navigating between related subjects; for example, in a rigorous study, Becks *et al.* (2002) showed that visually communicating the similarity between concepts in a text collection "significantly improved the effectiveness of task solutions" (p. 625).

7.8.2 Preference

There was a significant preference for the novel SE-3D as compared with the traditional text-only baseline system. This is consistent with IV evaluation reports (Becks, *et al.*, 2002; Chen, *et al.*, 1998; Chung, *et al.*, 2005; Rivadeneira & Bederson, 2003; Sutcliffe, *et al.*, 2000a; Yuan, Zhang & Trofimovsky, 2010). This preference may be due to an initial and potentially short-lived novelty effect. Longitudinal evaluations of IV usage could begin to answer this question—this seems to be unexplored territory.

Assuming preference for IV interfaces is strictly due to positive first impressions that fade over time, this may be a sufficient motivation for IV for IR development and use. As described in section 3.2.2.1 (Affects of Initial Impressions of a Design), there are indications positive first impressions of a system have a measurable effect on learning, user preference and performance. The computer's shift from a purely utilitarian business tool towards its widespread discretionary use suggests the emotive factor is critical to digital tool acceptance and usage. Löwgren (2006) reminds us of this historical technology usage shift and states:

> *It is clear that conditions for good use are no longer confined to efficient and error-free performance of tasks with set goals, but hinge on emotional and affective qualities of the use experience. (p. 383-384)*

There is a growing body of literature concerning the critical nature of emotions and aesthetics in computing (Fishwick, 2006). Emotions are considered a main cause of choice and action in HCI design (Brave & Nass, 2002; Cockton, 2002; Nielsen, 2003; Norman, 2002). Immediate and initial emotional responses precede the rationalized ones, and they may have a lasting effect (Lindgaard & Dudek, 2003). First impressions may be highly correlated with later evaluations of interactive systems (Fernandes, Lindgaard, Dillon & Wood, 2003; Tractinsky, Cokhavi & Kirschenbaum, 2004; Tractinsky, Katz & Ikar, 2000). To a large extent, emotions generated by a design "set the tone for the rest of the interaction" (Large, Beheshti, Tabatabaei & Nesset, 2009; Tractinsky & Zmiri, 2006, p. 407).

Positive first impressions and measured preference may partly explain why participants using SE-3D did not notice the time lag between search and result. This same query processing time was the second most prevalent qualitative comment concerning the text-only baseline system. What could explain the comparative patience of test participants when using the equivalent SE-3D search tool?

The lack of correlation between perceived speed and recorded times would suggest users' perception of time is modulated by affective reactions. This modulation in time perception is associated with moments of highest focus on the activity at hand (Csikszentmihalyi, 1990), and it has been found in technology usage scenarios (Chen, Wigand & Nilan, 1999; Hearst, Elliott, English, Sinha, Swearingen & Yee, 2002; Pilke, 2004). This is indicative of participant patience, error recovery, and preference associated with positive first impressions of a system (Ashby*, et al.*, 1999; Tractinsky, 1997).

This research adds to the few controlled comparative evaluations of functional IR system with IV interfaces (see review by Julien, et al., 2008), and builds on the body of knowledge concerning evaluations of IV for IR interfaces. The results show a significant advantage of the visual system (SE-3D) for effectively finding the most specific subject from a group of four. Participants significantly preferred SE-3D and thought it more useful. Future research plans in this area could include longitudinal experiments and task variations.

7.9 Limitations

The findings of this study are dependent on the tested tasks and systems. The conclusions drawn may not generalize to other tasks performed with different collections or other online search tools. The experimental design revealed performance differences between SE-3D and the chosen baseline system features. The experiment was performed with engineering students and extending the evaluation to participants from other disciplines may generate different results.

Chapter 8: CONCLUSION

Current information retrieval tools have limitations. Searchers have known **issues** of inadequate vocabulary, control of synonyms and homonyms, and lack of information concerning collection characteristics (e.g., relative coverage, matching results distribution). Browsing or explorative searching tasks are time-consuming because they are ill defined. This makes it difficult for searchers to formulate adequate search vocabulary. This produces long lists of low value information, and the success criteria is not apparent at the beginning of the task.

Information searching in large collections was described in terms of users **foraging for information** value (see section 3.3). Searchers extract valuable information from one source (e.g., a list of results, a group of items about a subject, etc.) while evaluating if other sources may offer a higher **information yield**. The **objective of the research** was to find and test ways to facilitate information value extraction by untrained searchers.

Representing information in a stable **3D virtual space** may facilitate information foraging by capitalizing on ubiquitous human experience with exploring and gathering in the physical world. The third dimension offered by virtual reality can potentially display more information as compared to 2D or 1D text-only result lists.

Design of **human-information interfaces** is guided by the demonstrated ease of learning and user preference for direct manipulation graphical user interfaces. Specifically, interactive visual representations of information collections stemming from the field of **information visualization**. IV techniques are said to be well suited to browsing and explorative searching tasks because they reveal patterns in the information.

Controlled subject vocabulary structures and their assignments to collections contain patterns which might be exploited by searchers. Specifically, this research has capitalized on the Bradford's Law distribution of subject assignments to dramatically reduce information structure noise, and explicitly represent the broad to narrow hierarchy of subjects offered by the **LCSH** structure. Controlled subject vocabulary offered an inherent hierarchical structure of broad to narrow terms which could intuitively be represented using known IV techniques for tree visualization. The **specific research question** asked if users of a novel 3D IV search tool would perform differently as compared with a text-only equivalent baseline.

The **research posited** that lack of explicit usage of CV was in part due to inadequate information retrieval tool designs. Adequacy is contextual and difficult to measure; the definition used by this research included time, accuracy, Likert scales, and qualitative comments. Using these measures, a novel IR tool was compared with a traditional text only keyword search interface acting as a baseline.

SE-3D is a kind of information retrieval environment created for this research. The design, development and testing of this 3D software is a **system study** in IV for IR. Chapter 5 described discoveries added to the IR domain, specifically research concerning the integration of information organization with information retrieval. It appears subject *hierarchy reconstruction* (Wang & Lee, 2007) of an LCSH hierarchy based on collection distribution has not been attempted before this research. Further research in this area could include studying the effect of different controlled vocabulary structures and the collections they organize.

This report has shown there are few fully functional 3D IR prototype reports, of these, less than a handful visually represent the information organization structure. Integrating keyword searching reduces this number down to one other report (Hearst 1997). Chapter 6 described SE-3D as an interactive 3D visual **navigation metaphor for information**, deeply integrated with **keyword searching** and ranked results lists. This integration takes the form of visual highlighting of the 3D visual subject structure based on keyword matches. Future research could explore mappings between collection characteristics (e.g., age, popularity, matches, etc.) and usable visual representations on the subject structure. The integration of multimodal communication is also of interest (e.g., audio or haptic).

Chapter 7 described the experimental design and procedure. Let us restate the specific question the controlled experiment hoped to address:

Are there differences between a 3D IV system and a text-only subject browser in terms of user performance and experience for undergraduate students performing IR tasks? Performance is measured by completion time and accuracy. Experience is measured by perceived speed, usefulness, ease of use and preference of the system.

The specific measured answers to this question are summarized in the following list of null hypotheses, the results found and their brief interpretation (see Table 8.1):

Null Hypothesis	Anticipated Result	Actual Result	Interpretation
H_{1null}: Time will not differ between systems for all task types	Not Rejected	Not Rejected	No effect of learning curve on the novel system
H_{2null}: Accuracy will not differ between systems for hierarchy navigation tasks	Rejected	Rejected	Subject hierarchy navigation tasks are a promising area for this IV application
H_{3null}: Accuracy will not differ between systems for simple retrieval tasks	Not Rejected	Not Rejected	Both system offered the same keyword searching used for these tasks
H_{4null}: Accuracy will not differ between systems for complex retrieval tasks	Rejected	Not Rejected	Complex retrieval tasks are difficult to test in a controlled setting
H_{5null}: Preference will not differ between systems	Rejected	Rejected	Users had positive impressions of the highly visual IV application and preferred it to the "boring" textual baseline
H_{6null}: Perceived ease of use and usefulness and preference are not correlated	Rejected	Not Rejected	Both system were rated as equally easy to use, usefulness differed. These concepts could not be correlated.

Table 8.1: Hypotheses tested by this research and results

SE-3D provided an advantage for a hierarchy navigation task. This suggests searchers may benefit from a visual representation of the subject structure for tasks which require exploration of subjects covered by a collection. For example, domain novices often search using very broad vocabulary resulting in long lists of irrelevant results. These searchers may acquire more specific previously unknown subject vocabulary by visually exploring the subject structure of the collection, and acquiring a mental model of the subjects covered by the collection.

SE-3D was preferred by participants and the preliminary analysis of qualitative comments suggests participants were more patient with it. It seems clear users prefer a pleasing visual search interface as opposed to a purely textual equivalent. These tools may be ready for the same historical shift from command based interfaces to GUIs. Beyond user preference, first impressions created by a product effect the quality of the ensuring interaction. Search tool designers are likely to prefer the positive impressions created by pleasing visual search tools as opposed to those created by purely utilitarian search box and ranked result list.

8.1 Contributions to Knowledge

Table 8.2 lists findings of interest generated by this research. They concern the domains of information visualization, information organization, usability testing, and the intersection between information retrieval and 3D interface design.

Research Area		Findings of Interest
Information Retrieval; Subject Hierarchy Reconstruction		• The vast majority of assigned subjects can be traced to an entry in the authority file, most by removing optional subdivisions. This produces a 70% reduction in subject hierarchy complexity. • LCSH contains conceptual traps or cycles which must be removed • Subject assignments follow a Bradford's Law distribution which is a form of power law: few subjects are assigned to many books, most subject are assigned to very few books. • Using subject assignments patterns, LCSH can be made into a clear hierarchy • As few as 22% of assigned subjects allowed access to over 92% of the collection. This is a dramatic reduction in complexity for a minimal loss of access.
Virtual Reality Information Visualisation Software Development		• Object positioning can be mapped to subject structure characteristics. • The vast majority of subjects have four or fewer children; the visual hierarchy is constrained and should be usable with a number of existing IV techniques. • Label occlusion management is best done dynamically as the user changes the viewpoint. • It is best to integrate the search indexes for subject terms and bibliographic information.
Evaluations of IV for IR Applications	**Quantitative**	• No significant difference in completion times between systems. • SE-3D produces significantly better accuracy for relative subject specificity tasks. • Users prefer SE-3D and find it more useful. • Preference and usefulness were correlated.
	Qualitative	• SE-3D creates positive first impressions. • Both systems could be more responsive. • Users were impatient with the baseline system search processing times. These same search processing times were present in SE-3D but they were not noticed by participants.

Table 8.2: Summary of Research Findings per Research Stage

The Semantic Web relies on controlled concept ontologies. Some of these structures are likely to come from the demonstrated practices offered by library and information science. As existing subject organization structures are ported to the Semantic Web (e.g., SKOS), concept ontology navigation tools will remain important. Their usable integration with keyword searching could increase information access to these ubiquitous and valuable information organization structures.

SE-3D is a novel IV for IR system, but is it a good design? Based on a list from White & McCain (1997), Buzydlowski *et al.* (2002) provide a list of criteria for evaluating a visual interface for information retrieval. The results shown in Table 8.3 suggest SE-3D is a promising design that can be improved and tested with larger collections.

Criteria	Answer	Supporting Evidence
Is the display an improvement over a simple list?	Yes	• Users prefer SE-3D • SE-3D offers higher performance for certain kinds of tasks
Does it provide new capabilities?	Yes	• SE-3D reveals the pattern of subject assignments and relations • SE-3D is novel integration of keyword searching and IV
Is it rapidly intelligible?	Yes	• Tests show users learn to use it quickly
Is it helpful in real time (or with an acceptable wait)?	Needs improvement	• SE-3D users require additional control over automated animated movements
Is it tied to an important collection?	Yes	• The 120,000+ collection is small by Web standards but is large by IV testing standards
Is it scalable upward to collections greater in size?	In theory	• Collection size does not instantly change the structure of information organization. The patterns they contain are relatively stable even if the collection grows drastically. • SE-3D scalability will be mostly constrained by text analysis and index searching capacity

Table 8.3: General evaluation of SE-3D

This research offers a contribution to the field of information visualization: Subject Explorer in 3D is a virtual reality information visualization application for information retrieval. SE-3D aimed to facilitate the visual exploration of information by offering searchers an interactive visual representation of the subject structure found in the Library of Congress Subject Headings (LCSH). The integration of visual exploration and ranked result lists is a step beyond text only interfaces towards direct manipulation information retrieval environments. SE-3D also demonstrated a novel technique to manage hundreds of overlapping textual labels in virtual reality, and a novel integration of explorative and specific keyword searching.

A second contribution is an automatic algorithm for the statistical analysis of an information collection. This revealed that subject assignments followed a power law; the top 1% most assigned subjects contained over 58% of the collection and 65% of non-empty subjects contained a single document. This meant that the vast majority of the collection could be represented by a small portion of the subjects it covered; specifically, a 78% reduction in subject structure complexity resulted in a loss of less than 8% access to the collection. This drastic simplification of the information structure may facilitate the exploration of the information for untrained searchers.

A third contribution is the testing of SE-3D by 24 undergraduate students during a repeated measures within-subject experiment. As compared with a text-only baseline, SE-3D produced an advantage in accuracy. Participants were more patient with SE-3D, they preferred it and perceived it as more useful.

These contributions suggest it is possible to extract additional value from organized collections by offering untrained users a reconstructed subject structure integrated with keyword searching. This research is significant for the development and testing of improved bridges between information organization and information retrieval, and interactive information visualization.

A future research direction may include the porting of SE-3D to other collections and their information organization schemes in order to improve bridges between information retrieval and information organization. SE-3D is still being refined and the addition of audio information representations are being developed within the context of another doctoral research project from the McGill School of Information Studies. People want pleasing interactive visual interfaces and search tools have yet to deliver this requirement. This research was a step in this direction which suggests it is possible to build better search tools.

References

Abello, J., van Ham, F., & Krishnan, N. (2006). ASK-GraphView : A Large Scale Graph Visualization System. *IEEE Transactions on Visualization and Computer Graphics, 12*(5), 1-8.

Ahlberg, C., & Shneiderman, B. (1994a). Alphaslider: a rapid and compact selector. *Paper presented at SIGCHI conference on Human factors in computing systems: celebrating interdependence. Boston, MA.* 365-371.

Ahlberg, C., & Shneiderman, B. (1994b). Visual information seeking: tight coupling of dynamic query filters with starfield displays. *Paper presented at SIGCHI conference on Human factors in computing systems: celebrating interdependence. Boston, MA.* 313-317.

Ahlberg, C., Williamson, C., & Shneiderman, B. (1992). Dynamic queries for information exploration: an implementation and evaluation. *Paper presented at Conf. on Human Factors in Computung Systems (CHI'92). Monterey, California.* 619-626.

Akrivi, K., Constantin, H., Georgios, L., Costas, V., & Eugenia, G. (2007). Ontology Visualiazation Methods - A Survey. *ACM Computing Surveys, 39*(4, Article 10), 1-43.

Akrivi, K., Elena, T., Constantin, H., Georgios, L., & Costas, V. (2006). A Comparative Study of Four Ontology Visualization Techniques in Protégé: Experiment Setup and Preliminary Results. *Paper presented at Information Visualization (IV). London.* 1-7.

Ali, K., Hartmann, K., & Strothotte, T. (2005). Label Layout for Interactive 3D Illustrations. *The Journal of WSCG, 13*. Retrieved Sept. 30th, 2010 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.68.9292&rep=rep1&type=pdf

Allan, J., Leuski, A., Swan, R., & Byrd, D. (2001). Evaluating combinations of ranked lists and visualizations of inter-document similarity. *Information Processing & Management, 37*(3), 435-458.

Allen, R. B. (1995). Two digital library interfaces that exploit hierarchical structure. *Paper presented at DAGS95 Conf.: Electronic Publishing and the Information Superhighway. Boston, MA.* Retrieved Oct. 2nd, 2010 from http://www.cis.drexel.edu/faculty/ballen/PAPERS/HOPAC/libr.pdf

Anderson, J. P., & Hofmann, M. A. (2006). A fully faceted syntax for Library of Congress subject headings. *Cataloging & Classification Quarterly, 43*(1), 7-37.

Andrews, K. (1995). Visualizing cyberspace: information visualization in the Harmony Internet browser. *Paper presented at Information Visualization Conf. Atlanta, GA.* 97-104.

Andrews, K. (2002). Visual Exploration of Large Hierarchies with Information Pyramids. *Paper presented at 6th Int. Conf. on Information Visualisation (IV'02). London, England.* 793-799.

Arnheim, R. (Ed.). (1972). *Visual Thinking*. Berkeley, CA.: University of California Press.

Ashby, F. G., Isen, A. M., & Turken, A. U. (1999). A neuropsychological theory of positive effect and its influence on cognition. *Psychological Review, 106*, 529-550.

Baddeley, A. D. (1998). *Human memory*. Boston: Allyn & Bacon.

Baeza-Yates, R., & Rivabeiro-Neto, B. (1999). *Modern Information Retrieval*. Boston, MA: Addison-Wesley Longman.

Barlow, T., & Neville, P. (2001). A comparison of 2D visualizations of hierarchies. *Paper presented at IEEE Symposium on Information Visualization. San Diego, CA.* 131-138.

Bates, M. J. (1986a). An exploratory paradigm for online information retrieval. *Paper presented at 6th IRFIS Conference: Intelligent Information Systems for the Information Society. London.* 91-99.

Bates, M. J. (1986b). Subject access for online catalogs: A design model. *Journal of the American Society for Information Science, 37*, 357-376.

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review, 13*(5), 407-424.

Bates, M. J. (2003). *Task force recommendation 2.3 Research and design review: Improving user access to library catalog and portal information. Final Report (V3).* Washington, DC: Library of Congress. Retrieved Sept. 29, 2010 from http://www.loc.gov/catdir/bibcontrol/2.3BatesReport6-03.doc.pdf

Battista, G. D., Eades, P., Tamassia, R., & Tollis, I. G. (1994). Algorithms for drawing graphs : an annotated bibliography. *Computational Geometry: Theory and Applications, 4*(5), 235-282.

Beagle, D. (1998). Visualization of metadata. *Information Technology and Libraries, 18*(4), 192-199.

Beaudoin, L., Parent, M.-A., & Vroomen, L. C. (1996). Cheops: a compact explorer for complex hierarchies. *Paper presented at IEEE Visualization Conf. San Francisco, California.* 87-92.

Becks, A., Seeling, C., & Minkenberg, R. (2002). Benefits of Document Maps for Text Access Knowledge Management: A Comparative Study. *Paper presented at ACM Symposium on Applied Computing. Madrid, Spain.* 621-626.

Bederson, B. B. (2003). Does animation help users build mental maps of spatial information? In B. B. Bederson & B. Shneiderman (Eds.), *The craft of information visualization: readings and reflections* (pp. 87-94). Amsterdam: Morgan Kaufmann Publishers.

Bederson, B. B., Hollan, J. D., Perlin, K., Meyer, J., Bacon, D., & Furnas, G. W. (1996). Pad++: A zoomable graphical sketchpad for exploring alternate interface physics. *Journal of Visual Languages and Computing, 7*, 3-31.

Bederson, B. B., Hollan, J. D., Stewart, J., Rogers, D., Vick, D., Ring, L., et al. (1998). A zooming Web browser. In C. Forsythe, J. Ratner & E. Grose (Eds.), *Human Factors in Web development* (pp. 255-266). New Jersey: Lawrence Erlbaum.

Bederson, B. B., & Shneiderman, B. (Eds.). (2003). *The craft of information visualization: readings and reflections*. Amsterdam: Morgan Kaufmann Publishers.

Belkin, N. J. (1980). Anomalous states of knowledge as the basis for information retrieval. *Canadian Journal of Information Science, 5*, 133-143.

Belkin, N. J., & Croft, W. B. (1987). Retrieval techniques. In M. E. Williams (Ed.), *Annual Review of Information Science and Technology* (Vol. 22, pp. 109-145). Amsterdam: Elsevier.

Belkin, N. J., Marchetti, P. G., & Cool, C. (1993). BRAQUE: design of an interface to support user interaction in information retrieval. *Information Processing and Management, Special issue on hypertext and information retrieval, 29*(3), 325-344.

Belkin, N. J., Oddy, R. N., & Brooks, H. M. (1982). ASK for information retrieval: Parts 1 and II. *Journal of Documentation, 38*(2,3).

Bender, M. A., Farach-Colton, M., Pemmasani, G., Skiena, S., & Sumazin, P. (2005). Lowest common ancestors in trees and directed acyclic graphs. *Journal of Algorithms, 57*(2), 75-94.

Berners-Lee, T., Hendler, J., & Lassila, O. (May 2001). The Semantic Web. *Scientific American, 284*(5), 28-37.

Bertin, J. (1967). *Semiologie graphique: les diagrammes, les réseaux, les cartes*. Paris: Gauthier-Villars.

Besant, L. (1982). Early survey findings: Users of Public Access Catalogs want Sophisticated Subject Access. *American Libraries, 13*(3), 160.

Bladh, T., Carr, D., & Scholl, J. (2004). Extending tree-maps to three dimensions: a comparitive study. *Paper presented at 6th Asia-Pacific Conf. on Computer-Human Interaction (APCHI 2004). Rotorua, New Sealand*. 50-59.

Blair, D. C., & Maron, M. E. (1985). An evaluation of retrieval effectiveness for a full-text document-retrieval system. *Communications of the ACM, 28*(3), 289-299.

Bleicher, S. (2005). *Contemporary color theory and use*. Clifton Park, NY: Thomson/Delmar Learning.

Borgman, C. (1996). Why are online catalogs still hard to use? *Journal of the American Society for Information Science, 47*(7), 493-503.

Borland, P., & Ingwersen, P. (1997). The development of a method for the evaluation of interactive information retrieval systems. *Journal of Documentation, 53*(3), 225-250.

Borner, K., & Chen, C. (2002). Visual interfaces to digital libraries. *Paper presented at 2nd ACM/IEEE-CS joint conf. on Digital libraries. Portland, Oregon, USA*. 425.

Borner, K., Chen, C., & Boyack, K. W. (2003). Visualizing Knowledge Domains *Annual Review of Information Science & Technology* (Vol. 37, pp. 179-255).

Bornstein, R. F. (1989). Exposure and Affect - Overview and Meta-Analysis of Research, 1968-1987. *Psychological Bulletin, 106*(2), 265-289.

Bosca, A., Bomino, D., & Pellegrino, P. (2005). OntoSphere: more than a 3D ontology visualization tool. *Paper presented at SWAP, the 2nd Italian Semantic Web Workshop. Trento, Italy*. Retrieved June 17th, 2009 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.62.8462&rep=rep1&type=pdf

Bosca, A., Bonino, D., Comerio, M., Grega, S., & Corno, F. (2007). A reusable 3D visualization component for the semantic web. *Paper presented at 12th Int. Conf. on 3D Web Technology. Perugia, Italy*. 89-96.

Bourgeois, F., Guiard, Y., & Lafon, M. B. (2001). Pan-zoom coordination in multi-scale pointing. *Paper presented at CHI '01 extended abstracts on Human factors in computing systems. Seattle, Washington.* 157-158.

Brady, R., Pixton, J., Baxter, G., Moran, P., Potter, C. S., Carragher, B., et al. (1995). Crumbs: A virtual Environment Tracking Tool for Biological Imaging. *Paper presented at IEEE Symposium on Frontiers in Biomedical Visualization. Atlanta, GA.* 18-25.

Brave, S., & Nass, C. (2002). Emotion in human-computer interaction. In J. Jacko & A. Sears (Eds.), *Handbok of human-computer interaction* (pp. 251-271). Mahwah, NJ: Lawrence Erlbaum Associates.

Brooke, J. (1996). SUS – A quick and dirty usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester & I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189-194). London: Taylor & Francis.

Brooks, F. P. J. (1995). *The mythical man-month (anniversary ed.)*: Addison-Wesley Longman Publishing Co., Inc.

Bruls, M., Huising, K., & van Wijk, J. J. (2000). Squarified treemaps. *Paper presented at joint Eurographics and IEEE TCVG Symp. on Visualization. Vienna.* 33-42.

Buzydlowski, J. W. (2003). *Comparison of self-organizing maps and pathfinder networks for the mapping of co-cited authors (Unpublished PhD Thesis).* Drexel University, Philadelphia.

Buzydlowski, J. W., White, H. D., & Lin, X. (2002). Term Co-occurrence Analysis as an Interface for Digital Libraries. *Lecture Notes in Computer Science, 2539*, 133-144.

Byrd, D. (1999). A scrollbar-based visualization for document navigation. *Paper presented at Fourth ACM Conf. on Digital Libraries. Berkeley, California.* 122-129.

Calhoun, K. (2006). *The changing nature of the catalog and its integration with other discovery tools: Prepared for the Library of Congress*. Retrieved Sept. 1st, 2010 from www.loc.gov/catdir/calhoun-report-final.pdf

Card, S., Pirolli, P., Van Der Wege, M., Morrison, J. B., Reeder, R. W., Schradley, P. K., et al. (2001). Information Scent as a Driver of Web Behavior Graphs: Results of a Prototype Analysis Method for Web Usability. *Paper presented at ACM CHI Conf. Seattle, Washington.* 498-505.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999a). Information visualization. In S. K. Card, J. D. Mackinlay & B. Shneiderman (Eds.), *Readings in information visualization: using vision to think* (1 ed., pp. 1-34). San Francisco, CA: Morgan-Kaufmann.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (Eds.). (1999b). *Readings in information visualization: using vision to think* (1 ed.). San Francisco, CA: Morgan-Kaufmann.

Card, S. K., & Robertson, G. (1996). The WebBook and the Web Forager: an information workspace for the World-Wide Web. *Paper presented at ACM CHI Conf. Vancouver, British Columbia, Canada.* 111-117.

Carlyle, A. (1989). Matching LCSH and user vocabulary in the library catalog. *Cataloging & Classification Quarterly, 10*(1-2), 37-63.

Carriere, J., & Kazman, R. (1995). Research report: Interacting with huge hierarchies: beyond cone trees. *Paper presented at IEEE Symposium on Information Visualization. Atlanta, Georgia*. 74-78.
Carroll, J. M., & Thomas, J. C. (1982). Metaphor and the Cognitive Representation of Computing Systems. *Systems, Man and Cybernetics, IEEE Transactions on, 12*(2), 107-116.
Carter, D. E. (2006). *The new big book of color in design*. New York: Collins Design.
Celentano, A., & Pittarello, F. (2004). Observing ad adapting user behavior in navigational interfaces. *Paper presented at ACM AVI Conf. Gallipoli, Italy*. 275-282.
Chalmers, M. (1993). Using a landscape metaphor to represent a corpus of documents. *Paper presented at European Conf. on Spatial Information Theory. Elba, Italy*. 377-390.
Chalmers, M., & Chitson, P. (1992). BEAD: Exploration in information visualization. *Paper presented at ACM SIGIR Conf. Copenhagen, Denmark*. 330-337.
Chan, L. (1989). Library of Congress class numbers in online catalog searching. *RQ. 28, 530, 536*.
Chan, L. (2000). Exploiting LCSH, LCC, and DDC To Retrieve Networked Resources: Issues and Challenges. *Paper presented at Bicentennial Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web. Washington, DC*. 159-178.
Chan, L., & Hodges, T. (2000). Entering the millennium: a new century for LCSH. *Cataloging & Classification Quarterly, 29*(1), 225-234.
Chan, L. M., & Hodges, T. (2007). *Cataloging and classification : an introduction* (3rd ed.). Lanham, Md.: Scarecrow Press.
Chang, N. L., & Said, A. (2003). *Automatically designed 3-D environments for intuitive browsing and discovery* (HP Labs Technical Report HPL-2003-92). Retrieved Sept. 30th, 2010 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.58.6238&rep=rep1&type=pdf
Chang, S.-J. L. (2005). Chang's Browsing. In K. E. Fisher, S. Erdelez & L. E. F. McKenchnie (Eds.), *Theories of Information Behavior*. Medford, NJ: Information Today, Inc. for ASIST.
Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology, 9*(2), 129-136.
Chen, C. (2000). Individual differences in a spatial-semantic virtual environment. *Journal of the American Society for Information Science, 51*(6), 529-542.
Chen, C. (2005). Top 10 Unsolved Information Visualization Problems. *IEEE Computer Graphics and Applications, 25*(4), 12-16.
Chen, C. (Ed.). (1999). *Information Visualisation and Virtual Environments*. London: Springer Verlag.
Chen, C. (Ed.). (2004). *Information visualization: beyond the horizon* (2nd ed.). London: Springer-Verlag.
Chen, C., & Czerwinski, M. (2000a). Empirical studies of information visualization: a meta-analysis. *International Journal of Human Computer Studies, 53*(5), 851-866.

Chen, C., & Czerwinski, M. (2000b). Empirical studies of information visualization: an introduction. *International Journal of Human Computer Studies, 53*(5), 631-635.

Chen, H., & Dhar, V. (1990). User Misconceptions of Information Retriaval Systems. *International Journal of Man-Machine Studies, 32*, 673-692.

Chen, H., Houston, A. L., Sewell, R. R., & Schatz, B. R. (1998). Internet browsing and searching: user evaluations of category map and concept space techniques. *Journal of the American Society of Information Science (and Technology), 49*(7), 582-603.

Chen, H., Wigand, R. T., & Nilan, M. (1999). Flow experiences on the Web. *Computers in Human Behavior, 15*, 454-608.

Chen, H., Yim, T., Fye, D., & Schatz, B. (1995). Automatic thesaurus generation for an electronic community system. *Journal of the American Society for Information Science, 46*(3), 175-193.

Chi, E. H., Hong, L., Heiser, J., Card, S. K., & Gumbrecht, M. (2007). ScentIndex and ScentHighlights: productive reading techniques for conceptually reorganizing subject indexes and highlighting passages. *Information Visualization, 6*, 32-47.

Chi, E. H., Rosien, A., Suppattanasiri, G., Williams, A., Royer, C., & Chow, C. (2003). The Bloodhound Project: Automating discovery of Web usability issues using the InfoScent simulator. *CHI Letters, 5*, 505-512.

Chowdhury, G. G. (1999). *Introduction to modern information retrieval*. London: Library Association.

Chu, H. (2003). *Information representation and retrieval in the digital age*. Medford, N.J.: Published for the American Society for Information Science and Technology by Information Today.

Chung, W., Chen, H., & Nunamaker Jr., J. F. (2005). A Visual Framework for Knowledge Discovery on the Web: An Empirical Study of Business Intelligence Exploration. *Journal of Management Information Systems, 21*(4), 57-84.

Cleveland, W. S., & McGill, R. (1984). Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Assoc., 79*(387), 531-554.

Cmolík, L., & Bittner, J. (2010). Layout-aware optimization for interactive labeling of 3D models. *Computers & Graphics, 34*(4), 378-387.

Cochrane, P. A. (1986). *Improving LCSH for use in online catalogs: exercises for self-help with a selection of background readings*. Englewood, CO, USA: Libraries Unlimited, Inc. .

Cochrane, P. A. (2000). Improving LCSH for Use in Online Catalogs Revisited-What Progress Has Been Made? What Issues Still Remain? *Cataloging & Classification Quarterly, 29*(1), 73-89.

Cockburn, A., & McKenzie, B. (2000). An evaluation of cone trees. *Paper presented at People and Computers XV (the 2000 British Computer Society Conference on Human-Computer Interaction). University of Sunderland.* 425-436.

Cockton, G. (2002). From doing to being: bringing emotion into interaction. *Interacting with Computers, 14*(2), 89-92.

Cohen, J. (Ed.). (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Colour Source Book (2006). London: Flame Tree.

Conway, M. (1993). Characteristics of Subject Headings in the Library of Congress BOOKSM Database. *Library Resources and Technical Services, 37*(1), 47-58.

Costabile, M. F., Malerba, D., Hemmje, M., & Paradiso, A. (1998). Building metaphors for supporting user interaction with multimedia databases. *Paper presented at 4th IFIP Conf. on Visual Database Systems*. 47-66.

Cribbin, T., & Chen, C. (2001). Visual-Spatial Exploration of Thematic Spaces: A Comparative Study of Three Visualisation Models. *Paper presented at SPIE Conf. Electronic Imaging 2001: Visual Data Exploration and Analysis VIII. San Jose, CA*. 199-209.

Cribbin, T., & Westerman, S. J. (1999). Spatial data management systems: mapping semantic distance. *Paper presented at INTERACT '99 Conf. on Human-Computer Interaction. Edinburgh, Scotland.*

Csikszentmihalyi, M. (Ed.). (1990). *Flow. The psychology of optimal experience*. New York: Harper and Row.

Cubaud, P., Stokowski, P., & Topol, A. (2002, 2002). Binding browsing and reading activities in a 3D digital library. *Paper presented at the Proc. of JCDL, Portland, OR*. 281-282.

Cutting, D. R., Karger, D. R., Pedersen, J. O., & Tukey, J. W. (1992). Scatter/Gather: a cluster-based approach to browsing large document collections. *Paper presented at ACM Conf. on Research and Development in Information Retrieval. Copenhagen, Denmark*. 318-329.

Czumaj, A., Kowaluk, M., & Lingas, A. (2007). Faster algorithms for finding lowest common ancestors in directed acyclic graphs. *Theoretical Computer Science, 380*(1-2), 37-46.

Darken, R. P., & Sibert, J. L. (1996). Wayfinding strategies and behaviors in large virtual worlds. *Paper presented at CHI '96 Conf. Vancouver, Canada*. 142-149.

Davidson, G. S., Hendrickson, B., Johnson, D. K., Meyers, C. E., & Wylie, B. N. (1998). Knowledge mining with VxInsight: Discovery through interaction. *Journal of Intelligent Information Systems, 11*(3), 259-285.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly, 13*(3), 319-340.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science, 35*(8), 982-1003.

de Groot, A. D. (1966). Perception and memory versus thought: some old ideas and recent findings. In B. Kleinmuntz (Ed.), *Problem Solving* (pp. 19-50). New York: Wiley.

Di Battista, G. (1999). *Graph drawing : algorithms for the visualization of graphs*. Upper Saddle River, N.J.: Prentice Hall.

Dillon, A., McKnight, C., & Richardson, J. (1990). Navigation in hypertext: a critical review of the concept. *Paper presented at INTERACT '90 Conf. Cambridge, UK.* Retrieved Oct. 2nd, 2010 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.90.2750&rep=rep1&type=pdf

Ding, Y., Chowdhury, G. G., Foo, S., & Qian, W. (2000). Bibliometric Information Retrieval System (BIRS): A Web Search Interface Utilizing Bibliometric

Research Results. *Journal of the American Society for Information Science 51*(13), 1190-1204.

Douglas, S., & Moran, T. P. (1983). Learning text editor semantics by analogy. *Paper presented at CHI Conf. Boston, MA.* 207-211.

Drabenstott, K., Simcox, S., & Fenton, E. (1999). End-user understanding of subject headings in library catalogs. *Library Resources & Technical Services, 43*(3), 140-160.

Drabenstott, K. M., & Weller, M. S. (1996a). The exact-display approach for online catalog subject searching. *Information Processing and Management, 32*(6), 719-745.

Drabenstott, K. M., & Weller, M. S. (1996b). Failure Analysis of Subject Searches in a Test of a New Design for Subject Access to Online Catalogs. *Journal of the American Society for Information Science, 47*(4), 519-537.

Driessen, S., Jacobs, J., & Huijsen, W.-O. (2006). Combining query and visual search for knowledge mapping. *Paper presented at InfoVis Conf. London, England.* 216-224.

Dumais, S., Cutrell, E., & Chen, H. (2001). Optimizing search by showing results in context. *Paper presented at Conf. on Human Factors in Computing Systems. Seattle, WA.* 227-284.

Dushay, N. (2004). Visualizing Bibliographic Metadata - A Virtual (Book) Spine Viewer. *D-Lib Magazine, 10*(10). Retrieved Oct. 2nd, 2010 from http://www.cs.cornell.edu/naomi/ECDL2004/dushayECDL2004.pdf

Eades, P., Lai, W., Misue, K., & Sugiyama, K. (1991). Preserving the mental map of a diagram. *Paper presented at the Int. Conf. on Computational Graphics and Visualization Techniques (COMPUGRAPHICS '91).* 34-43.

Edmondson, S., Christensen, J., Marks, J., & Shieber, S. (1996). A General Cartographic Labelling Algorithm. *Cartographica: The International Journal for Geographic Information and Geovisualization, 33*(4), 13-24.

Edwards, D., & Hardman, L. (1988). 'Lost in hyperspace':cognitive mapping and navigation in a hypertext environment. In R. McAleese (Ed.), *Hypertext I: Theory into Practice* (pp. 105-125). Wiltshire, UK: Cromwell Press for Intellect Books.

Ellis, G., & Dix, A. (2006). An explorative analysis of user evaluation studies in information visualisation. *Paper presented at the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization. Venice, Italy.* 1-7.

Emond, B., & West, R. L. (2003). Cyberpsychology: A Human-Interaction Perspective Based on Cognitive Modeling. *CyberPsychology & Behavior, 6*(5), 527-536.

Fabrikant, I., & Skupin, A. (2003). Spatialization Methods: A Cartographic Research Agenda for Non-Geographic Information Visualization. *Cartography and Geographic Information Science, 30*(2), 99-119.

Fabrikant, S. (2001). Evaluating the usability of the scale metaphor for querying semantic spaces. *Lecture Notes in Computer Science, 2205*, 156-173.

Fairchild, K. M., Serra, L., Hem, N., Beng Hai, L., & Tin Leong, A. (1993). Dynamic FishEye Information Visualizations. In R. A. Earnshaw, M. A. Gigante & H. Jones (Eds.), *Virtual Reality Systems* (pp. 161-177): Academic Press.

Farber, E. (1984). Catalog Dependency. *Library Journal, 109*(3), 325-328.

Fekete, J.-D., & Plaisant, C. (1999). Excentric labeling: dynamic neighborhood labeling for data visualization. *Paper presented at the SIGCHI conference on Human factors in computing systems: the CHI is the limit. Pittsburgh, Pennsylvania, United States.* 512-519.

Fekete, J.-D., & Plaisant, C. (2003). Interactive information visualization of a million items. In B. B. Bederson & B. Shneiderman (Eds.), *The craft of information visualization: readings and reflections* (pp. 279-286). Amsterdam: Morgan Kaufmann Publishers.

Fernandes, G., Lindgaard, G., Dillon, R., & Wood, J. (2003). Judging the appeal of web sites. *Paper presented at the 4th World Congress on the Management of Electronic Commerce. McMaster University, Hamilton, Ontario, Canada.*

Fischer, K. (2005). Critical Views of LCSH, 1990–2001: The Third Bibliographic Essay. *Cataloging & Classification Quarterly, 41*(1), 63-109.

Fishwick, P. A. (2006). *Aesthetic computing*. Cambridge, MA: The MIT Press.

Fluit, C., Sabou, M., & van Harmelen, F. (2003). Ontology-based Information Visualization. In V. Geroimenko & C. Chen (Eds.), *Visualizing the semantic Web : XML-based Internet and information visualization* (pp. 36-48). London: Springer.

Foote, H., & Thomas, J. (2005). Dynamic visualization of graphs with extended labels. *Paper presented at IEEE Symposium on Information Visualization, INFOVIS 2005. Minneapolis, MN.* 73 - 80.

Fowler, R. H., Folwer, W. A. L., & Wilson, B. A. (1991). Integrating query, thesaurus, and documents through a common visual representation. *Paper presented at Annual Int. ACM/SIGIR Conf. Chicago, Illinois.* 142-151.

Fox, E. A., Das Neves, F., Yu, X., Shen, R., Kim, S., & Fan, W. (2006). Exploring the Computing Literature with Visualization and Stepping Stones & Pathways. *Communications of the ACM, 49*(4), 53-58.

Fox, E. A., Urs, S. R., & Cronin, B. (2002). Digital libraries. *Annual Review of Information Science and Technology, 36*, 503-589.

Frank, E., & Paynter, G. W. (2004). Predicting Library of Congress classifications from Library of Congress subject headings. *Journal of the American Society for Information Science and Technology, 55*(3), 214-227.

Freitas, C. S., Luzzaerdi, P., Cava, R., Winckler, M., Pimenta, M., & Nedel, L. (2002). Evaluating usability of information visualization. *Paper presented at Symposium on Human Factors in computer systems IHC. Fortaleza, Ceara.* Retrieved Oct. 2nd, 2010 from http://www.irit.fr/recherches/IHCS/winckler/projects/spider/publications/IHC2002-FreitasEtAl.pdf

Frost, C., & Dede, B. (1988). Subject heading compatibility between LCSH and catalog files of a large research library: A suggested model for analysis. *Information Technology and Libraries, 7*(3), 288-299.

Furnas, G. W., & Rauch, S. J. (1998). Considerations for information environments and the NaviQue workspace. *Paper presented at the third ACM conference on Digital libraries. Pittsburgh, Pennsylvania, United States.* 79-88.

Gansner, E., & North, S. (1998). Improved Force-Directed Layouts *Graph Drawing* (Vol. 1547, pp. 364-373). Heidelberg: Springer Berlin.

Gattis, M., & Holyoak, K. J. (1996). Mapping conceptual to spatial relations in visual reasoning. *Journal of Experimental Psychology-Learning Memory and Cognition, 22*(1), 231-239.

Gentner, D., & Nielsen, J. (1996). The anti-Mac interface. *Communications of the Acm, 39*(8), 70-82.

Gibson, J. J. (Ed.). (1986). *The ecological approach to visual perception*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Graham, R. L., & Hell, P. (1985). On the History of the Minimum Spanning Tree Problem. *IEEE Annals of the History of Computing, 7*(1), 43-57.

Grun, C., Gerken, J., Jetter, H.-C., Konig, W., & Reitener, H. (2005). MedioVis- A user-centered library metadata browser. In A. Rauber, S. Christodoulakis & A. M. Tjoa (Eds.), *Research and advanced technology for digital libraries 2005: the 9th European Conference (ECDL 2005)* (pp. 174-185). Berlin: Springer-Verlag.

Guastavino, C., & Katz, B. F. G. (2004). Perceptual evaluation of multi-dimensional spatial audio reproduction. *The Journal of the Acoustical Society of America, 116*(2), 1105-1115.

Guiard, Y., & Beaudouin-Lafon, M. (2004). Target acquisition in multiscale electronic worlds. *International Journal of Human-Computer Studies, 61*(6), 875-905.

Gutwin, C. (2002). Improving focus targeting in interactive fisheye views. *Paper presented at the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves. Minneapolis, Minnesota, USA.* 267-274.

Halasz, F., & Moran, T. P. (1982). Analogy considered harmful. *Paper presented at the 1982 conference on Human factors in computing systems. Gaithersburg, Maryland, United States.* 383-386.

Harper, B. D., & Norman, K. L. (1993). Improving User Satisfaction: The Questionnaire for User Interaction Satisfaction Version 5.5. *Paper presented at the 1st Annual Mid-Atlantic Human Factors Conf. Virginia Beach, VA.* 224-228.

Harper, C. A. (2006). Encoding Library of Congress Subject Headings in SKOS: Authority Control for the Semantic Web. *Paper presented at the Int. Conf. on Dublin Core and Metadata Aplications. Manzanillo.* Retrieved June 4th, 2009 from https://scholarsbank.uoregon.edu/xmlui/handle/1794/3268

Hartmann, K., Ali, K., & Strothotte, T. (2004). Floating Labels: Applying Dynamic Potential Fields for Label Layout. In A. Butz, A. Krüger & P. Olivier (Eds.), *Smart Graphics* (Vol. 3031, pp. 101-113). Berlin: Springer.

Healey, C. G., Booth, K. S., & Enns, J. T. (1996). High-speed visual estimation using preattentive processing. *ACM Transactions on Computer-Human Interaction, 3*(2), 107-135.

Hearst, M. A. (1999). User interfaces and visualization. In R. Baeza-Yates & B. Ribeiro-Neto (Eds.), *Modern Information Retrieval* (pp. 257-323). New York: ACM Press.

Hearst, M. A. (2009). *Search User Interfaces*. Cambridge, MA: Cambridge University Press.

Hearst, M. A., Elliott, A., English, J., Sinha, R., Swearingen, K., & Yee, K.-P. (2002). Finding the flow in Web site search. *Communications of the ACM, 45*(9), 42-49.

Hearst, M. A., & Karadi, C. (1997). Cat-a-Cone: an interactive interface for specifying searches and viewing retrieval results using a large category hierarchy. *Paper presented at Int. Conf. on Research and Development in Information Retrieval (SIGIR). Philadelphia, PA.* 246-255.

Hedman, A., Carr, D. A., & Nassla, H. (2004). Browsing thumbnails: a comparison of three techniques. *Paper presented at Int. Conf. Information Technology Interfaces (ITI). Cavtat, Croatia.* 353-360.

Hemmje, M., Kunkel, C., & Willet, A. (1994). LyberWorld - A visualization user interface supporting fulltext retrieval. *Paper presented at SIGIR. Dublin, Ireland.* 249-259.

Henderson, J., & Card, S. K. (1986). Rooms: the use of multiple virtual workspaces to reduce space contention in window-based graphical user interfaces. *ACM Transactions on Graphics, 5*(3), 211-241.

Henzler, R. G. (1978). Free or controlled vocabularies: Some statistical user-orient ed evaluations of biomedical information systems. *International Classification, 5*, 21-26.

Heo, M. (2000). *A usability study on web visualization techniques and user mental models.* Unpublished Faculty of Information Science Doctoral Thesis, University of Pittsburgh.

Herman, I., Melancon, G., & Marshall, M. S. (2000). Graph visualization and navigation in information visualization: A survey. *Visualization and Computer Graphics, IEEE Transactions on, 6*(1), 24-43.

Hetzler, B., Harris, W. M., Havre, S., & Whitney, P. (1998). Visualizing the full spectrum of document relationships. *Paper presented at the 5th International ISKO Conf.: Structures and Relations in the Knowledge Organization. Lille, France.* Retrieved Sept. 30th, 2010 from http://infoviz.pnl.gov/pdf/isko.pdf

Hetzler, E., & Turner, A. (2004). Analysis experiences using information visualization. *IEEE Computer Graphics and Applications, 24*(5), 22-26.

Hoare, C., & Sorensen, H. (2005). Information Foraging with a Proximity-Based Browsing Tool. *Artificial Intelligence Review, 24*, 233-252.

Hoffman, H. (1999). Non-Standard Chronological Subdivisions In United States History. *TECHNICALITIES, 19*(1), 5.

Holzschlag, M. E. (2001). *Color for websites*. Crans-Près-Céligny; Hove: RotoVision.

Hornbaek, K., & Frokjaer, E. (2003). Evaluating User Interfaces with Metaphors of human Thinking. In N. Carbonell & C. Stephanidis (Eds.), *User Interfaces for All* (pp. 486-507). Berlin: Springer-Verlag.

Iglesias, E., & Stringer Hye, S. (2008). Topic maps and the ILS: an undelivered promise. *Library Hi Tech, 26*(1), 12-18.

Indurkhya, B. (Ed.). (1992). *Metaphor and cognition*. Dordrecht, Netherland: Kluwer Academic.

Iwayama, M., & Tokunaga, T. (1995). Cluster-Based Text Categorization: A Comparison of Category Search Strategies. *Paper presented at SIGIR-95, 18th ACM International Conference on Research and Development in Information Retrieval. Seattle, WA.* 273-281.

Jackendoff, R. (Ed.). (1983). *Semantics and Cognition*. London: MIT Press.

Janosky, B., Smith, P. J., & Hildreth, C. R. (1986). Online library catalog systems: An analysis of user errors. *International Journal of Man-Machine Studies, 25*(5), 573-592.

Jarvenpaa, S. L. (1989). The Effect of Task Demands and Graphical Format on Information-Processing Strategies. *Management Science, 35*(3), 285-303.

Jasen, B. J., Spink, A., Bateman, J., & Saracevic, T. (1998). Real Life Information Retrieval: A Study of User Queries on the Web. *ACM SIGIR Forum, 32*(1), 5-17.

Jih, H. J., & Reeves, T. C. (1992). Mental models: A research focus for interactive learning systems. *Educational Technology Research and Development, 40*, 1042-1629.

Johnson, B., & Shneiderman, B. (1991). Tree-maps: A space filling approach to the visualization of hierarchical information structures. *Paper presented at IEEE Conf. on Visualization. San Diego, CA.* 248-291.

Jul, S., & Furnas, G. W. (1997). Navigation in electronic worlds: A CHI 97 workshop. *SIGCHI Bulletin, 29*(4), 44-49.

Julien, C.-A., & Bouthillier, F. (2008). Le catalogue réinventé. *Documentation et bibliothèques, 54*(3), 229-239.

Julien, C.-A., & Cole, C. (2009). Capitalizing on Controlled Subject Vocabulary by Providing a Map of Main Subject Headings : an Exploratory Design Study. *Canadian Journal of Information and Library Sciences, 33*(1/2), 67-83.

Julien, C.-A., Leide, J. E., & Bouthillier, F. (2008). Controlled Comparative User Evaluations of Information Visualization Interfaces for Text Retrieval: Review and Meta-Analysis. *Journal of the American Society for Information Science and Technology, 59*(6), 1012-1024.

Juran, J. M., & Godfrey, A. B. (1999). *Juran's quality handbook* (5th ed.). New York: McGraw Hill.

Katz, C. A., & Byrne, M. D. (2003). Effects of Scent and Breath on Use of Site Specific Search on E-Commerce Web Sites. *ACM Transactions on Computer–Human Interaction, 10*(3), 198-220.

Keim, D. A. (2002). Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics, 8*(1), 1-8.

Keller, P. R., & Keller, M. M. (Eds.). (1993). *Visual clues: Practical Data Visualization*. Los Alamitos, CA: IEEE Computer Society Press.

Keller, T., & Tergan, S.-O. (2005). Visualizing knowledge and information: an introduction. In S.-O. Tergan & T. Keller (Eds.), *Knowledge and Information Visualization* (pp. 1-23). Berlin: Springer-Verlag.

Kieras, D. E., & Bovair, S. (1984). The role of a mental model in learning to operate a device. *Cognitive Science, 8*(3), 255-273.

Kim, H., & Hirtle, S. C. (1995). Spatial metaphors and the disorientation in hypertext browsing. *Behaviour & Information Technology, 14*, 239-250.

Kleiberg, E., van de Wetering, H., & van Wijk, J. J. (2001). Botanical Visualization of Huge Hierarchies. *Paper presented at IEEE Symp. on Information Visualization (InfoVis '01). San Diego, CA.* 87-94.

Kobsa, A. (2001). An empirical comparison of three commercial information visualization systems. *Paper presented at IEEE Symposium on Information Visualization. San Diego, CA.* 123-130.

Kobsa, A. (2004). User Experiments with Tree Visualization Systems. *Paper presented at IEEE Symposium on Information Visualization. Austin, TX.* 9-16.
Kohonen, T. (Ed.). (1997). *Self-Organizing Maps*. Berlin: Springer-Verlag.
Korfhage, R. (1991). To see or not to see - Is that the query? *Communications of the ACM, 34*, 134-141.
Korn, F., & Shneiderman, B. (1995). *Navigating terminology hierarchies to access a digital library of medical images* (Technical Report HCIL-TR-94-03). College Park, MD: University of Maryland
Kosara, R., & Hauser, H. (2003). An interaction view on information visualization. *Paper presented at Eurographics 2003 State-of-the Art Reports. Grenada, Spain.* 123-137.
Koshman, S. (2006). Visualization-based information retrieval on the Web. *Library & Information Science Research, 28*, 192-207.
Krohn, U. (1995). Visualization of navigational retrieval in virtual information spaces. *Paper presented at the Workshop on New Paradigms in Information Visualization and Manipulation. Baltimore, MD.* 26-32.
Krowne, A., & Halbert, M. (2005). An initial evaluation of automated organization for digital library browsing. *Paper presented at JCDL '05: Proceedings of the 5th ACM/IEEE-CS joint conference on Digital libraries. Denver, CO, USA.* 246-255.
Kumar, H. P., Plaisant, C., & Shneiderman, C. (1997). Browsing hierarchical data with multilevel dynamic queries and pruning. *International Journal of Human-Computer Studies, 46*, 103-124.
Lakoff, G., & Johnson, M. (1980). *Metaphors we live by* Chicago: University of Chicago Press.
Lamping, J., Rao, R., & Pirolli, P. (1995). A focus+context technique based on hyperbolic geometry for visualizing large hierarchies. *Paper presented at Human Factors in Computing Systems (SIGCHI) Conf. Denver, Colorado.* 401-408.
Lancaster, F. (Ed.). (1986). *Vocabulary control for information retrieval* (2nd ed.). Arlington, VA: Information Resources.
Large, A., Beheshti, J., Clement, I., Tabatabaei, N., & Tam, M. (2009). Visualizing a hierarchical taxonomy in a children's web portal: User Evaluations of a prototype *Canadian Journal of Information and Library Science 33*(2/4), 255-282.
Large, A., Beheshti, J., & Cole, C. (2002). Information Architecture for the Web: The IA Matrix Approach to Designing Children's Portals. *Journal of American Society for Informational Science and Technology, 53*(10), 831-838.
Large, A., Beheshti, J., Tabatabaei, N., & Nesset, V. (2009). Developing a visual taxonomy: Children's views on aesthetics. *Journal of the American Society for Information Science and Technology, 60*(9), 1808-1822.
Larivière, V., Archambault, E., Gingras, Y., & Vignola-Gagné, É. (2006). The place of serials in referencing practices: Comparing natural sciences and engineering with social sciences and humanities: Research Articles. *Journal of the American Society for Information Science and Technology, 57*(8), 997-1004.
Larson, K., & Czerwinski, M. (1998). Web page design: Implications of memory, structure and scent for information retrieval. *Paper presented at the SIGCHI conference on Human factors in computing systems. Los Angeles, CA.* 25-32.

Larson, R. R. (1991). The decline of subject searching: Long-term trends and patterns of index use in an online catalog. *Journal of the American Society for Information Science, 42*(3), 197-215.

Larson, R. R. (1992). Experiments in automatic Library of Congress Classification. *Journal of the American Society for Information Science, 43*(2), 130-148.

Laurel, B. (Ed.). (1993). *Computers as theatre* (1st ed.). Reading, MA: Addison-Wesley Publishing Company.

Le Grand, B., & Soto, M. (2006). Topic Maps, RDF Graphs, and Ontologies Visualization. In V. Geroimenko & C. Chen (Eds.), *Visualizing the Semantic Web* (pp. 59-79). London: Springer.

Lee, B., Parr, C., Plaisant, C., Bederson, B. B., Veskler, V. D., Gray, W. D., et al. (2006). TreePlus: Interactive exploration of networks with enhanced tree layouts. *IEEE Transactions On Visualization And Computer Graphics, 12*(6), 1414-1426.

Leung, Y. K., & Apperley, M. D. (1994). A review and taxonomy of distortion-oriented presentation techniques. *ACM Transactions on Computer-Human Interaction (TOCHI), 1*(2), 126-160.

Leuski, A., & Allen, J. (2000). Improving interactive retrieval by combining ranked lists and clustering. *Paper presented at RIAO'2000. College de France.* 665-681.

Levy, E., Zacks, J., Tversky, B., & Sciano, D. (1996). Gratuitous graphics? Putting preferences in perspectives. *Paper presented at CHI Conf. Vancouver, Canada.* 42-49.

Li, J., Plaisant, C., & Shneiderman, B. (1998). Data object and label placement for information abundant visualizations. *Paper presented at the 1998 workshop on New paradigms in information visualization and manipulation. Washington, D.C.* 41-48.

Li, Y., & Belkin, N. J. (2008). A faceted approach to conceptualizing tasks in information seeking. *Information Processing & Management, 44*(6), 1822-1837.http://www.sciencedirect.com/science/article/B6VC8-4TFH2HS-1/2/b6b0cf3b993c5ad39d189261294cbcc1

Lidwell, W., Holden, K., & Butler, J. (2003). *Universal principles of design*. Gloucester, Mass.: Rockport.

Liebig, T., & Noppens, O. (2004). OntoTrack: Combining Browsing and Editing with Reasoning and Explaining for OWL Lite Ontologies. *Paper presented at Semantic Web ISWC 2004 Conf. Hiroshima, Japan.* 244-258.

Lin, X. (1997). Map displays for information retrieval. *Journal of the American Society for Information Science, 48*, 40-54.

Lindgaard, G., & Dudek, C. (2003). What is this evasive beast we call user satisfaction? *Interacting with Computers, 15*(3), 429-452.

Liu, T., Yang, Y., Wan, H., Zeng, H., Chen, Z., & Ma, W. (2005). Support vector machines classification with a very large-scale taxonomy. *ACM SIGKDD Explorations Newsletter, 7*(1), 36-43.

Loftin, R. B., & Kenney, P. J. (1995). Training the Hubble Space Telescope Flight Team. *IEEE Computer Graphics and Applications, 15*(5), 31-37.

Logie, R. H. (Ed.). (1995). *Visuo-spatial working memory*. Hove: Lawrence Erlbaum Associates.

Lohse, G., Rueter, H., Biolsi, K., & Walker, N. (1990). Classifying visual knowledge representations: A foundation for visualization research. *Paper presented at First IEEE Conf. on Visualization. San Francisco, CA.* 131-138.
Lohse, G. L., & Walker, N. (1993). Classifying graphical knowledge. In A. Kent (Ed.), *Encyclopedia of library and information science* (Vol. 53, pp. 38-82).
Löwgren, J. (2006). Articulating the use qualities of digital designs. In P. A. Fishwick (Ed.), *Aesthetic computing* (pp. 383–403). Cambridge, MA: The MIT Press.
Ludy, L., HAMMELL, K., & GOLDBERG, K. (1985). OSU libraries' use of Library of Congress subject authorities file. *Information Technology and Libraries, 4*(2), 155-160.
Lynch, K. (Ed.). (1960). *The image of the city*: MIT Press.
Maass, S., & Döllner, J. (2006). Dynamic Annotation of Interactive Environments using Object-Integrated Billboards. *Paper presented at 14th Int. Conf. in Central Europe on Computer Graphics, Visualization and Computer Vision. Plzen, Czech Republic.* 327-334.
Mack, R. L., Lewis, C. H., & Carroll, J. M. (1983). Learning to use office systems: problems and prospects. *ACM Transactions on Office Information Systems, 1*, 254-271.
Manber, U., Smith, M., & Gopal, B. (1997). Webglimpse: Combining browsing and searching. *Paper presented at Usenix Technical Conference. Anaheim, CA.* 15.
Mann, T. M., & Reiterer, H. (2000). Evaluation of different visualizations of Web search results. *Paper presented at 11th Int. Workshop on Database and Expert Systems Applications Greenwich, London, U.K.* 586-590.
Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. New York: Cambridge University Press.
Marchionini, G. (1987). An invitation to browse: designing full text system for novice users. *Canadian Journal of Information Science, 12*(3), 69-79.
Marchionini, G. (Ed.). (1995). *Information Seeking in Electronic Environments*. Cambridge, UK: Cambridge University Press.
Marchionini, G., & Shneiderman, B. (1988). Finding facts vs. browsing knowledge in hypertext systems. *IEEE Computer, 21*(1), 70-80.
Mariani, J. A., & Lougher, R. (1992). Triplespace: an experiment in a 3D graphical interface to a binary relational database. *Interacting with Computers, 4*, 147-162.
Markey, K. (1984). *Subject Searching in Library Catalogs: Before and after the introduction of online Catalogs*. Dublin, OH: Online Computer Library Center OCLC.
Markey, K. (1985). Subject-Searching experiences and needs of online catalog users : Implications for library classification. *Library Resources and Technical Services, 29*, 34-51.
Markey, K. (1989). Alphabetical searching in an online catalog. *Journal of Academic Librarianship, 4*(6), 353-360.
Markey, K. (2007). The Online Library Catalog: Paradise Lost and Paradise Regained? *D-Lib, 13*(1/2). Retrieved Oct. 2nd, 2010 from http://www.biblioteche.unibo.it/sba/progetti/formazione-utenti/bonora3.doc
Markey, K., Atherton, P., & Newton, C. (1982). An analysis of controlled vocabulary and free text search statements in online searches. *Online Review, 4*, 225-236.

Matthews, J. R., Lawrence, G. S., & Ferguson, D. K. (Eds.). (1983). *Using Online Catalogs. A nationwide survey*. New York: Neal-Schuman.

McAleese, R. (1989). Navigation and browsing in hypertext. In R. McAleese (Ed.), *Hypertext: Theory into Practice* (pp. 5-38). Oxford: Intellect.

McCain, K. W. (1990). Mapping Authors in Intellectual Space: A Technical Overview. *Journal of the American Society for Information Science, 41*, 433-443.

McClure, C. (1976). Subject and Added Entries as Access to Information. *Journal of academic librarianship, 2*(1), 9-14.

McDonald, S., & Stevenson, R. J. (1996). Disorientation in hypertext: The effects of three text structures on navigation performance. *Applied Ergonomics, 27*(1), 61-68.

McDonald, S., & Stevenson, R. J. (1998). Navigation in hyperspace: An evaluation of navigational tools and subject matter expertise on browsing and information retrieval in hypertext. *Interacting with Computers, 10*, 129-142.

McKim, R. H. (Ed.). (1980). *Experiences in Visual Thinking* (2nd ed.). Monterey, CA: Brooks/Cole.

McKnight, C., Dillon, A., & Richardson, J. A. (1990). A comparison of linear and hyper text formats in information retrieval. In R. McAleese & G. C. (Eds.), *Hypertext: the state of art* (pp. 10-19). Oxford, UK: Intellect Books.

Meadow, C. T., Boyce, B. R., & Kraft, D. H. (2007). *Text information retrieval systems* (3rd ed.). Amsterdam ; London: Academic.

Mercun, T., & Zumer, M. (2008). New generation of catalogues for the new generation of users: A comparison of six library catalogues. *Program: electronic library and information systems, 42*(3), 243-261.

Merton, R. K. (Ed.). (1957). *Social Theory and Social Structure*. New York: The Free Press.

Mitchell, S. (1999). Interface design considerations in libraries. *Science & Technology Libraries, 17*(3/4), 131-181.

Modjeska, D., & Waterworth, J. (2000). Effects of desktop 3D world design on user navigation and search performance. *Paper presented at IEEE Symposium on Information Visualization. Salt Lake City, Utah.* 215-220.

Monk, A., Walsh, P., & Dix, A. (1988). A comparison of hypertext, scrolling, and folding as mechanisms for program browsing. In D. Jones & R. Winder (Eds.), *Peopl and Computers IV* (pp. 421-436). Cambridge: Cambridge University Press.

Morse, E., Lewis, M., & Olsen, K. (2002). Testing visual information retrieval methodologies case study: comparative analysis of textual, icon, graphical and 'spring' displays. *Journal of the American Society for Information Science and Technology, 53*(1), 28-40.

Morse, E., Lewis, M., & Olsen, K. A. (2000). Evaluating visualizations: using a taxonomic guide. *International Journal of Human-Computer Studies, 53*(5), 637-662.

Motro, A. (1986). BAROQUE: A browser for relational databases. *ACM Transactions on Office Information Systems, 4*, 164-181.

Moya-Anegón, F., Vargas-Quesada, B., Herrero-Solana, V., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., & Munoz-Fernández, F. J. (2004). A new technique for

building maps of large scientific domains based on the cocitation of classes and categories. *Scientometrics, 61*(1), 129-145.
Mullet, K., Fry, C., & Schiano, D. (1997). On your marks, get set, browse. *Paper presented at Human Factors in Computing Systems, CHI '97 (Extended Abstracts). Atlanta, GA.* 113-114.
Mullet, K., & Sano, D. (Eds.). (1995). *Designing Visual Interfaces: Communication Oriented Techniques*. Englewood Cliffs, NJ: Prentice Hall.
Munzner, T. (1997). H3: laying out large directed graphs in 3D hyperbolic space. *Paper presented at IEEE Symp. on Information Visualization. Phoenix, AZ.* 2-10.
Nahl, D. (2007). Social-Biological information technology: an integrated conceptual framework. *Journal of the American Society for Information Science and Technology, 58*(13), 2021-2046.
Nation, D. A., Plaisant, C., Marchionini, G., & Komladi, A. (1997). Visualizing websites using a hierarchical table of contents browser: WebTOC. *Paper presented at 3rd Conf. on Human Factors in and the Web.* Retrieved Oct. 1st, 2010 from http://hcil.cs.umd.edu/trs/97-10/97-10.html
Navarro-Prieto, R., Scaife, M., & Rogers, Y. (1999). Cognitive Strategies in Web Searching. *Paper presented at the 5th Conf. on Human Factors & the Web.* Retrieved June 8th. 2009 from http://zing.ncsl.nist.gov/hfweb/proceedings/navarro-prieto/
Newby, G. B. (2002). Empirical study of a 3D visualization for information retrieval tasks. *Journal of Intelligent Information Systems, 18*, 31-53.
Nguyen, Q. V., Huang, M. L., & Hawryszkiewycz, I. (2004). A new visualization approach for supporting knowledge management and collaboration in E-Learning. *Paper presented at Int. Conf. on Information Visualization. London, UK.* 693-700.
Nguyen, T. N., & Zhang, J. (2006). A Novel Visualization Model for Web Search Results. *IEEE Transactions of Visualization and Computer Graphics, 12*(5), 981-988.
Nielsen, J. (2003). User empowerment and the fun factor. In M. A. Blythe, K. Overbeeke, A. F. Monk & P. C. Wright (Eds.), *Funology: From Usability to Enjoyment* (pp. 103-105). Dordrecht, Netherlands: Kluwer Academic Publishers.
Nielsen, J. (Ed.). (1993). *Usability Engineering*. Boston, MA: Academic Press.
Norman, D., A. (Ed.). (2004). *Emotional design : why we love (or hate) everyday things*. New York, NY: Basic Books.
Norman, D. A. (1998). *The invisible computer : why good products can fail, the personal computer is so complex, and information appliances are the solution*. Cambridge, Mass.: MIT Press.
Norman, D. A. (2002). Emotion & Design, Attractive things work better. *Interactions, IX.4*, 36-42.
Norman, D. A. (Ed.). (1988). *The psychology of everyday things*. New York, NY: Basic Books.
Noy, N. F., & McGuiness, D. L. (March 2001). *Ontology Development 101: A Guide to Creating Your First Ontology. Technical Report KSL-01-05*: Stanford Knowledge Systems Laboratory
O'Neill, E., & Aluri, R. (1981). Library of Congress subject heading patterns in OCLC monographic records. *Library Resources and Technical Services, 25*, 63-80.

Oakhill, J., & Johnson-Laird, P. (1984). Representation of spatial descriptions in working memory. *Current Psychology, 3*(1), 52-62.
Olsen, K. A., Korfhage, R. R., Sochats, K. M., Spring, M. B., & Williams, J. G. (1993). Visualization of a document collection: The VIBE system. *Information Processing & Management, 29*(1), 69-81.
Olson, H. A. (2004). The ubiquitous hierarchy: an army to overcome the threat of a mob. *Library Trends, 52*(3), 604-616.
Olston, C., & Chi, E. (2003). ScentTrails: Integrating Browsing and Searching the Web. *ACM Transactions on Computer–Human Interaction, 10*(4), 177-197.
Overbeeke, K., Djajadiningrat, T., Hummels, C., Wensveen, S., & Frens, J. (2003). Let's make things engaging. In M. A. Blythe, K. Overbeeke, A. F. Monk & P. C. Wright (Eds.), *Funology: From Usability to Enjoyment* (pp. 7-18). Dordrecht, Netherlands: Kluwer Academic Publishers.
Papadakis, I., Stefanidakis, M., & Tzali, A. (2008). Visualizing OPAC subject headings. *Library Hi Tech, 26*(1), 19-23.
Parsia, B., Wang, T., & Goldbeck, J. (2005). Visualizing Web ontologies with CropCircles. *Paper presented at the 4th Int. Semantic Web Conf. Galway, Ireland.* Retrieved Oct. 2nd, 2010 from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.77.5408&rep=rep1&type=pdf
Pedroza, C. (2005). The Encyclopedia of Educational Technology. Retrieved 21 October 2005, from http://coe.sdsu.edu/eet/articles/visualperc1/start.htm
Pejtersen, A. M. (1988). Search strategies and database design for information retrieval in libraries. In L. P. Goodstein, H. B. Andersen & S. E. Olsen (Eds.), *Tasks, Errors and Mental Models, Hampshire* (pp. 171-192). England: Taylor & Francis.
Perez, C., & de Antonio, A. (2004). 3D Visualization of text collections: an experimental study to assess the usefulness of 3D*. Paper presented at IEEE Symposium on Information Visualization. Austin, TX.* 317-323.
Peterson, S. D., Axholt, M., Cooper, M., & Ellis, S. (2009a). Evaluation of Alternative Label Placement Techniques in Dynamic Virtual Environments *Lecture Notes in Computer Science* (Vol. 5531, pp. 43-55). Berlin: Springer.
Peterson, S. D., Axholt, M., Cooper, M., & Ellis, S. R. (2009b). Visual clutter management in augmented reality: Effects of three label separation methods on spatial judgments*. Paper presented at the IEEE Symposium on 3D User Interfaces. Lafayette, Louisiana, USA.* 111-118.
Petrie, H. G., & Oshlag, R. S. (1993). Metaphor and learning. In A. Ortony (Ed.), *Metaphor and thought* (pp. 579-609). Cambridge, England: Cambridge University Press.
Pietriga, E., Appert, C., & Beaudouin-Lafon, M. (2007). Pointing and beyond: an operationalization and preliminary evaluation of multi-scale searching*. Paper presented at the SIGCHI conference on Human factors in computing systems. San Jose, California, USA.* 1215-1224.
Piggott, M. (1988). *A topography of cataloguing: showing the most important landmarks, communications and perilous places*. London: Library Assn Pub Ltd.
Pilke, E. M. (2004). Flow Experiences in information technology use. *International Journal of Human-computer Studies, 61*(3), 347-357.

Pirolli, P. (1998). Exploring browser design trade-offs using a dynamical model of optimal information foraging. *Paper presented at CHI. Los Angeles, CA.* 33-40.
Pirolli, P., Card, S., & Van Der Wege, M. (2003). The Effects of Information Scent on Visual Search in the Hyperbolic Tree Browser. *ACM Transactions on Computer-Human Interaction, 10*(1), 20-53.
Pirolli, P., & Card, S. K. (1999). Information foraging. *Psychological Review, 106*(4), 643-675.
Pirolli, P., Card, S. K., & Van Der Wege, M. M. (2000). The effect of information scent on searching information visualizations of large tree structures. *Paper presented at Int. Conf. on Advanced Visual Interfaces. Palermo, IT.* 161-172.
Plaisant, C. (2004). The challenge of information visualization evaluation. *Paper presented at Proceedings of of AVI 2004. Gallipoli, Italy.* 109-116.
Plaisant, C., Grosjean, J., & Bederson, B. B. (2002). Spacetrees: Supporting Exploration in Large Node Link Tree, Design Evolution and Empirical Evaluation. *Paper presented at IEEE Symposium on Information Visualization. Boston, MA.* 57-64.
Plaisant, C., Grosjean, J., & Bederson, B. B. (2003). SpaceTree: supporting exploration in large node link tree, design evolution and empirical evaluation. In B. B. Bederson & B. Shneiderman (Eds.), *The craft of information visualization: readings and reflections* (pp. 287-294). Amsterdam: Morgan Kaufmann Publishers.
Porter, M. F. (1980). An algorithm for suffic stripping. *Program: electronic library and information systems, 14*(3), 130-137.
Raeithel, A., & Velichkovshy, B. M. (1996). Joint attention and co-construction: new ways to foster user-designer collaboration. In B. A. Nardi (Ed.), *Context and Consciousness: Activity theory and Human-computer Interaction* (pp. 199-233). Cambridge, MA: MIT Press.
Raubal, M., & Worboys, M. (1999). A Formal Model of the Process of Wayfinding in Built Environments *Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science* (Vol. 1661, pp. 748). Berlin: Springer.
Reese, T. (2000). MarcEdit (Version 5.0). Corvallis, OR: Oregon State University.
Reitman, W. R. (1964). Heuristic decision procedures, open constraints,and the structure of ill-defined problems. In M. W. Shelly & G. L. Bryan (Eds.), *Human judgements and optimality* (pp. 282-315). New York: Wiley.
Reitman, W. R. (Ed.). (1965). *Cognition and thought: An information-processing approach*. New York: Wiley.
Reitz, J. M. (Ed.). (2004). *Dictionary for Library and Information Science*. Westport, CT: Libraries Unlimited.
Resnikoff, H. L. (Ed.). (1989). *The illusion of reality*. New York: Springer-Verlag.
Rhyne, T.-M. C., Tory, M., Munzner, T., Ward, M., Johnson, C., & Laidlaw, D. H. P. (2003). Panel: Information and Scientific Visualization: Separate but Equal or Happy Together at Last. *Paper presented at IEEE Visualization.* 115. Retrieved Oct. 2nd, 2010 from http://portal.acm.org/citation.cfm?id=1081559
Risden, K., Czerwinski, M. P., Munzner, T., & Cook, D. B. (2000). An initial examination of ease of use for 2D and 3D information visualizations of web content. *International Journal of Human Computer Studies, 53*(5), 695-714.

Ritter, F., Sonnet, H., Hartmann, K., & Strothotte, T. (2003). Illustrative shadows: integrating 3D and 2D information displays. *Paper presented at the 8th international conference on Intelligent user interfaces. Miami, Florida, USA.* 166-173.

Rivadeneira, W., & Bederson, B. B. (2003). *A Study of Search Result Clustering Interfaces: Comparing Textual and Zoomable User Interfaces* (Report no HCIL-2003-36, Univ. of Maryland). College Park, Maryland

Robertson, G. G., Card, S. K., & Mackinlay, J. D. (1993). Information visualization using 3-D interactive animation. *Communications of the ACM, 36*(4), 57-71.

Robertson, G. G., Mackinlay, J. D., & Card, S. K. (1991). Cone trees: animated 3D visualizations of hierarchical information. *Paper presented at SIGCHI Conf. on Human Factors in Computing. New Orleans, Louisiana.* 189-194.

Rogers, Y. (1989). Icons at the interface: their usefulness. *Interacting with Computers, 1*(1), 105-117.

Rosenthal, R. (Ed.). (1991). *Meta-Analytic Procedures for Social Research* (Rev. ed.). Newbury Park, CA: Sage Publications.

Roussinov, D., & Chen, H. (2001). Information navigation on the web by clustering and summarizing query results. *Information Processing and Management, 37*(6), 789-816.

Salton, G. (Ed.). (1989). *Automatic text processing: The transformation, analysis, and retrieval of information by computer*. Reading, MA: Addison-Wesley.

Sanchez, J. A., Twidale, M. B., Nichols, D. M., & Silva, N. N. (2005). Experiences with starfield visualizations for analysis of library collections. *Paper presented at Visualization and Data Analysis 2005. San Jose, CA, USA.* 215-225.

Sandstrom, P. E. (2001). Scholarly communication as a socioecological system. *Scientometrics, 51*(3), 573-605.

Santos, B. S., Zamfir, F., Ferreira, C., Mealha, O., & Nunes, J. (2004). Visual application for the analysis of Web-bases information systems usage: a preliminary usability evaluation. *Paper presented at IEEE Int. Conf. on Information Visualization. London, UK.* 812-818.

Schank, R., & Abelson, R. (Eds.). (1976). *Scripts, Plans, Goals and Understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schvaneveldt, R. W., Dearholt, D. W., & Durso, F. T. (1988). Graph theoretic foundations of pathfinder networks. *Computers & Mathematics with Applications, 15*(4), 337-345.

Sebrechts, M. M. (2005). Visualizing Information in Virtual Space: Prospects and Pitfalls. In S.-O. Tergan & T. Keller (Eds.), *Knowledge and Information Visualization* (pp. 136-166). Berlin: Springer-Verlag.

Sebrechts, M. M., Vasilakis, J., Miller, M. S., Cugini, J. V., & Laskowski, S. J. (1999). Visualization of search results: a comparative evaluation of text, 2D, and 3D interfaces. *Paper presented at Annual Int. ACM SIGIR Conf. Berkeley, CA.* 3-10.

Seeling, C., & Becks, A. (2004). Analysing associations of textual and relational data with a multiple views system. *Paper presented at Int. Conf. on Coordinated and Multiple Views in Exploratory Visualization. London, UK.* 61-70.

Sherman, W. R., & Craig, A. B. (Eds.). (2003). *Understanding virtual reality: interface, application and design*. San Francisco, CA: Morgan Kaufman.

Shiaw, H.-y., Jacob, R. J. K., & Crane, G. R. (2004). The 3D vase museum, a new approach to context in a digital library. *Paper presented at Proc. of the Joint Conference on Digital Libraries. Tucson, AZ.* 125-134.
Shiri, A. A., Revie, C., & Chowdhury, G. (2002). Thesaurus-enhanced search interfaces. *Journal of Information Science, 28*, 111-122.
Shneiderman, B. (1992). Tree visualization with treemaps: a 2-d space-filling approach. *ACM Transactions on Graphics, 11*(1), 92-99.
Shneiderman, B. (1996). The eyes have it: a task by data type taxonomy for information visualizations. *Paper presented at IEEE Sym. on Visual Languages. Boulder, CO.* 336-343.
Shneiderman, B. (2003). The eyes have it: a task by data type taxonomy for information visualizations. In B. B. Bederson & B. Shneiderman (Eds.), *The craft of information visualization: readings and reflections* (pp. 364-371). Amsterdam: Morgan Kaufmann Publishers.
Shneiderman, B. (Ed.). (1998). *Designing the user interface*. Reading, MA: Addison-Wesley.
Shneiderman, B., Feldman, D., Roseand, A., & Grau, X. F. (2000). Visualizing digital library search results with categorical and hierarchical axes. *Paper presented at 5th Int. Conf. on Digital Libraries (DL'00). San Antonio, TX.* 57-65.
Shneiderman, B., & Plaisant, C. (Eds.). (2005). *Designing the user interface: strategies for effective human-computer interaction* (4th ed.). Reading, MA: Addison-Wesley.
Shubert, S. (1992). Critical Views of LCSH-Ten Years Later. *Cataloging & Classification Quarterly, 15*(2), 37-97.
Silverstein, C., Marais, H., Henzinger, M., & Moricz, M. (1999). Analysis of a very large web search engine query log. *SIGIR Forum, 33*(1), 6-12.
Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics, 69*, 99-118.
Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). Cambridge, Mass.: MIT Press.
Simpson, A. (1989). Navigation in hypertext: design issues. *Paper presented at Paper presented at OnLine '89 Conference. London, UK.* 241-255.
Skupin, A. (2000). From metaphor to method: cartographic perspectives on information visualization. *Paper presented at InfoVis. Salt Lake City, Utah.* 91-97.
Small, H. (2000). Charting pathways through science: exploring Garfield's vision of a unified index to science. In B. Cronin & H. B. Atkins (Eds.), *The Web of Knowledge: A Festschrift in Honor of Eugene Garfield* (pp. 449–473). Medford, N. J.: Information Today.
Smallman, H. S., St John, M., Oonk, H. M., & Cowen, M. B. (2001). Information availability in 2D and 3D displays. *Ieee Computer Graphics and Applications, 21*(5), 51-57.
Smith, G., Czerwinski, M., Meyers, B., Robbins, D., Robertson, G., & Tan, D. S. (2006). FacetMap: A Scalable Search and Browse Visualization. *IEEE Transactions on Visualization and Computer Graphics, 12*(5), 797-804.

Smith, M., & Cochrane, P. (1999). Creating Better Subject Access with Multiple Vocabularies: Upgrading the Subject Heading List for the Alzheimer's Association. *Library Resources and Technical Services, 43*, 53-58.

Smith, P. A. (1986). Towards a practical measure of hypertext usability. *Interacting with Computers, 4*, 365-381.

Song, H., Curran, E. P., & Sterritt, R. (2004). Multiple foci visualisation of large hierarchies with FlexTree. *Information Visualization, 3*(1), 19-35.

Spence, R. (Ed.). (2001). *Information Visualization*: ACM Press.

Spivey, M. J., Tyler, M. J., Eberhard, K. M., & Tanenhaus, M. K. (2001). Linguistically mediated visual search. *Psychological Science, 12*(4), 282-286.

Stappers, P. J., Gaver, W., & Overbeeke, K. (2003). Beyond the limits of real-time realism: Moving from stimulation correspondence to information correspondence. In L. J. Hettinger & M. Haas (Eds.), *Virtual and Adaptive Environments: Applications, Implications, and Human Performance Issues* (pp. 91-110). Mahwah, NJ: Lawrence Erlbaum Associates.

Stein, T., & Décoret, X. (2008). Dynamic label placement for improved interactive exploration. *Paper presented at the 6th international symposium on Non-photorealistic animation and rendering. Annecy, France.* 15-21.

Stephens, G., & Handzic, M. (2004). Knowledge Discovery Through Visualising Using Virtual Reality. *Paper presented at Proceedings of the 37th Hawaii International Conference on system Sciences. Big Island, Hawaii.* 1-10.

Strausfeld, L. (1995). Financial viewpoints: using point-of-view to enable understanding of information. *Paper presented at CHI '95. Denver, CO.* 208-209.

Summers, E., Isaac, A., Redding, C., & Krech, D. (2008). LCSH, SKOS and Linked Data. *Paper presented at Proc. Int'l Conf. on Dublin Core and Metadata Applications. Berlin.* Retrieved June 4th, 2009 from http://dcpapers.dublincore.org/ojs/pubs/article/viewFile/916/912

Sutcliffe, A. (1996). 3D or not 3D: is it nobler in the mind? *Paper presented at HCI Conf. People and Computers XI. London, UK.* 79-94.

Sutcliffe, A. (Ed.). (2003). *Multimedia and Virtual Reality: designing multisensory User Interfaces*. Mahwa, NJ: Lawrence Erlbaum Associates.

Sutcliffe, A. G., Ennis, M., & Hu, J. (2000a). Evaluating the effectiveness of visual user interfaces for information retrieval. *International Journal of Human-Computer Studies, 53*(5), 741-763.

Sutcliffe, A. G., Ryan, M., Springett, M. V., & Doubleday, A. (2000b). Model mismatch analysis: towards a deeper explanation of users' usability problems. *Behaviour and Information Technology, 19*, 43-55.

Svenonius, E. (1986). Unanswered questions in the design of controlled vocabularies. *Journal of the American Society for Information Science*, *37*, 331-340.

Swan II, J. E., Gabbard, J. L., Hix, D., Schulman, R. S., & Kim, K. P. (2003). A comparative study of user performance in a map-based virtual environment. *Paper presented at IEEE Symp. on Virtual Reality. Chicago, Illinois.* 259-266.

Swan, R. C., & Allen, J. (1998). Aspect windows, 3-D visualizations, and indirect comparisons of information retrieval systems. *Paper presented at Annual Int. ACM SIGIR Conf. Melbourne, Australia.* 171-181.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*(3), 185-233.

Synnestvedt, M., & Chen, C. (2005). Design and Evaluation of the Tightly Coupled Perceptual-Cognitive Tasks in Knowledge Domain Visualization. *Paper presented at 11th Int. Conf. on Human-Computer Interaction. Las Vegas, Nevada.* Retrieved Oct. 2nd, 2010 from http://naglexbutbus.59.to/handle/1860/1937

Tang, M.-C. (2007). Browsing and searching in a faceted information space: a neturalistic study of PubMed Users' interaction wiht a display tool. *Journal of the American Society for Information Science and Technology, 58*(13), 1998-2006.

Taylor, A. G., & Miller, D. P. (2006). *Introduction to cataloging and classification* (10th ed.). Westport, Conn.: Libraries Unlimited.

Tesler, J., & Strasnick, S. (1992). FSN: The 3D File System Navigator. Mountain View, CA: Silicon Graphics Inc.

Teyseyre, A. R., & Campo, M. R. (2009). An Overview of 3D Software Visualization. *Visualization and Computer Graphics, IEEE Transactions on, 15*(1), 87-105.

Thompson, R. H., & Croft, W. B. (1989). Support for browsing in an intelligent text retrieval system. *International Journal of Man-Machine Studies, 30*, 639-668.

Toms, E. G. (2002). Information Interaction: Providing a Framework for Information Architecture. *Journal of the American Society for Information Science and Technology, 53*(10), 855-862.

Toms, E. G., Dufour, C., & Hesemeier, S. (2004). Measuring the User's experience with Digital Libraries. *Paper presented at Joint Conf. on Digital Libraries '04. Tuscon, AZ.* 51-52.

Tonta, Y. (1992). LCSH and PRECIS in library and information science: a comparative study. *Occasional Papers (University of Illinois at Urbana-Champaign,Graduate School of Library & Information Science), 194*, 1-68.

Tory, M., & Moller, T. (2004). Rethinking Visualization: A High-Level Taxonomy. *Paper presented at IEEE Symp. on Information Visualization.* 151-158.

Tractinsky, N. (1997). Aesthetics and Apparent Usability: Empirically Assessing Cultural and Methodological Issues. *Paper presented at Conf. on Human factors in computing systems (CHI '97). Atlanta, Georgia.* 115-122.

Tractinsky, N., Cokhavi, A., & Kirschenbaum, M. (2004). Using ratings and response latencies to evaluate the consistency of immediate aesthetic perceptions of web pages. *Paper presented at the 3rd Annual Workshop on HCI Reserach in MIS. Washington, DC.* 40-44.

Tractinsky, N., Katz, A., & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers, 13*(2), 127-145.

Tractinsky, N., & Zmiri, D. (2006). Exploring attributes of skins as potential antecedents of emotion in HCI. In P. A. Fishwick (Ed.), *Aesthetic computing* (pp. 405-422). Cambridge, MA: The MIT Press.

Trafton, J. G., Kirschenbaum, S. S., Tsui, T. L., Miyamoto, R. T., Ballas, J. A., & Raymond, P. D. (2000). Turning pictures into numbers: extracting and generating information from complex visualizations. *Int. Journal of Human-Computer Studies, 53*, 827-850.

Trafton, J. G., Marshall, S. P., Mintz, F., & Trickett, S. B. (2002). Extracting explicit and implicit information from complex visualizations *Diagrammatic Representation and Inference* (Vol. 2317, pp. 25-48). Berlin: Springer.

Triesman, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing, 31*, 156-177.

Trißl, S., & Leser, U. (2005). Querying Ontologies in Relational Database Systems. *Paper presented at Data Integration in the Life Sciences.* 63-79.

Tufte, E. R. (Ed.). (1983). *The Visual Display of Quantitative Information*. Chesshire, CT: Graphics Press.

Tufte, E. R. (Ed.). (1990). *Envisioning Information*. Cherchire, Conn.: Graphics Press.

Turetken, O., & Sharda, R. (2005). Clustering-Based Visual Interfaces for Presentation of Web Search Results: An Empirical Investigation. *Information Systems Frontiers, 7*(3), 273-297.

Turo, D. (2003). Hierarchical Visualization with Treemaps: Making sense of Pro Basketball data. In B. B. Bederson & B. Shneiderman (Eds.), *The craft of information visualization: readings and reflections* (pp. 237-238). Amsterdam: Morgan Kaufmann Publishers.

Tversky, B., Zacks, J., Lee, P., & Heiser, J. (2000). Lines, Blobs, Crosses and Arrows: Diagrammatic Communication with Schematic Figures *Theory and Application of Diagrams* (pp. 221-230). Berlin: Springer.

United States (1979). *Report of the President's Commission on the Accident at Three Mile Island : the need for change, the legacy of TMI*. Washington, D.C. : President’s Commission on the Accident at Three Mile Island.

Valduriez, P., & Boral, H. (1986). Evaluation of recursive queries using join indices. *Paper presented at 1st Int. Conf. on Expert Database Systems. Redwood, CA.* 271-293.

van Ham, F., & van Wijk, J. J. (2003). Beamtrees: compact visualization of large hierarchies. *Information Visualization, 2*, 31-39.

Van Pulis, N., & Ludy, L. E. (1988). Subject searching in an online catalog with authority control. *College & Research Libraries, 49*(6), 523-533.

Veerasamy, A., & Heikes, R. (1997). Effectiveness of a graphical display of retrieval results. *Paper presented at Annual Int. ACM SIGIR Conf. Philadelphia, PA, USA.* 236-245.

Veith, R. H. (Ed.). (1988). *Visual information systems: The power of graphics and video*. Boston: G. K. Hall & Co.

Veling, A. (1997). The Aqua Browser: visualization of large information spaces in context. *Journal of AGSI, 6*(3), 136-142.

Voorhees, E. M., & Harman, D. K. (Eds.). (2005). *TREC: Experiment and Evaluation in Information Retrieval*. Cambridge, MA: MIT Press.

Wang, J., & Lee, M. (2007). Reconstructing ddc for interactive classification. *Paper presented at Conf. on information and knowledge management. Lisbon, Portugal.* 137-146.

Wang, T., & Parsia, B. (2006). CropCircles: Topology Sensitive Visualization of OWL Class Hierarchies. *Paper presented at the Int. Semantic Web Conf. (ISWC), Athens, GA.* 695-708.

Wang, W., Wang, H., Dai, G., & Wang, H. (2006). Visualization of large hierarchical data by circle packing. *Paper presented at the SIGCHI conference on Human Factors in computing systems. Montreal, Canada.* 517-520.
Ware, C. (2008). *Visual thinking for design* Burlington, Mass: Morgan Kaufmann.
Ware, C., Purchase, H., Colpoys, L., & McGill, M. (2002). Cognitive measurements of graph aesthetics. *Information Visualization, 1*(2), 103-110.
Waterworth, J. A. (1996). A pattern of islands: exploring public information space in a private vehicle. In P. Brusilovsky, P. Kommers & N. Streitz (Eds.), *Multimedia, hypermedia and virtual reality: models, systems, and applications* (pp. 266-279). Berlin, Germany: Springer.
Wertheimer, M. (1925). *Drei abhandlungen zur gestalttheorie*. Erlangen: Philosophische akademie.
Westerman, S. J., Collins, J., & Cribbin, T. (2005). Browsing a document collection represented in two- and three-dimensional virtual information space. *International Journal of Human-Computer Studies, 62*(6), 713-736.
Westerman, S. J., & Cribbin, T. (2000). Mapping semantic information in virtual space: dimensions, variance and individual differences. *International Journal of Human-Computer Studies, 53*, 765-787.
White, H. D. (1990). Author Co-Citation Analysis: Overview and Defense. In C. L. Borgman (Ed.), *Scholarly Communication and Bibliometrics* (pp. 84-106). Newbury Park, CA: Sage Publications.
White, H. D. (2003). Pathfinder networks and author cocitation analysis: A remapping of paradigmatic information scientists. *Journal of the American Society for Information Science and Technology, 54*(5), 423-434.
White, H. D., Lin, X., Buzydlowski, J. W., & Chen, C. (2004). User-controlled mapping of significant literatures. *the National Academy of Sciences of the United States of America, 101*(Suppl 1), 5297-5302.
White, H. D., & McCain, K. W. (1997). Visualization of Literatures *Annual Review of Information Science and Technology (ARIST)* (Vol. 32, pp. 99-168).
White, H. D., & McCain, K. W. (1998). Visualizing a Discipline: An Author Co-Citation Analysis of Information Science. *Journal of the American Society for Information Science, 49*, 327-355.
Wiberley, S. E., & Daugherty, R. A. (1988). Users' persistence in scanning lists of references. *College & Research Libraries, 49*(2), 149-156.
Wilk, D., Rotenberg, S., Schacham, S., Hoffman, G., & Liebman, S. (2001). Problems in the use of Library of Congress Subject Headings as the basis for Hebrew Subject Headings in the Bar-Ilan University library. *International cataloguing and bibliographic control, 30*(3), 43-46.
Winckler, M. A., Palanque, P., & Freitas, C. M. D. S. (2004). Tasks and Scenario-based evaluation of information visualization techniques. *Paper presented at TAMODIA '04. Prague, Czeck Replublic.* 165-172.
Wise, J. A. (1998). The ecological approach to text visualization. *Journal of the American Society for Information Science, 50*, 1224-1233.
Wise, J. A., Thomas, J. J., Pennock, K., Lantrip, D., Pottier, M., Schur, A., et al. (1995). Visualizing the non-visual: Spatial analysis and interaction with information from

text documents. *Paper presented at IEEE Information Visualization. Atlanta, Georgia.* 51-58.

Wiss, U., Carr, D., & Jonsson, H. (1998). Evaluating three-dimensional information visualization designs: a case study of three designs. *Paper presented at IEEE Conf. on Information Visualization London, UK.* 137-144.

Yang, Y., Zhang, J., & Kisiel, B. (2003). A scalability analysis of classifiers in text categorization. *Paper presented at Annual Int. ACM SIGIR conf. on Research and development in informaion retrieval. Toronto, Canada.* 96-103.

Yee, K.-P., Swearingen, K., Li, K., & Hearst, M. (2003). Faceted metadata for image search and browsing. *Paper presented at Proc. of CHI 2003 conference. Ft. Lauderdale, Florida, USA.* 401-408.

Yi, K., & Chan, L. M. (2010). Revisiting the syntactical and structural analysis of Library of Congress Subject Headings for the digital environment. *Journal of the American Society for Information Science and Technology, 61*(4), 677-687.

Yuan, X., Zhang, X., & Trofimovsky, A. (2010). Testing visualization on the use of information systems. *Paper presented at Proceeding of the third symposium on Information interaction in context. New Brunswick, New Jersey, USA.* 365-370.

Zacks, J., & Tversky, B. (1999). Bars and lines: A study of graphic communication. *Memory & Cognition, 27*(6), 1073-1079.

Zaphiris, P., Gill, K., Ma, T. H.-Y., Wilson, S., & Petrie, H. (2004). Exploring the use of information visualization for digital libraries. *The New Review of Information Networking, 10*(1), 51-69.

Zhang, J., & Marchionini, G. (2005). Evaluation and evolution of a browse and search interface: Relation Browser++. *Paper presented at National Conf. on Digital Government Research. Atlanta, GA.* 179-188.

Zhu, B., & Chen, H. C. (2005). Information Visualization. *Annual Review of Information Science and Technology, 39*, 139-177.

Ziemkiewicz, C., & Kosara, R. (2008). The Shaping of Information by Visual Metaphors. *Visualization and Computer Graphics, IEEE Transactions on, 14*(6), 1269-1276.

ANNEXES

Annex 1. Subject Recruitment Add

Participate in research at McGill, help out a fellow student for about an hour and receive 20$.

It's not difficult, we're looking for people to try and test new kinds of software interfaces.

If you are a first year student enrolled in McGill undergraduate engineering program between 18 and 21 years old then you are eligible to receive a pre-selection questionnaire.

(A valid McGill student ID is necessary)

Interested? charles.julien@mail.mcgill.ca

The research is supervised by Prof. Catherine Guastavino from the School of Information Studies and will take place at the Multimedia Interaction Lab (Room B120 and B121, Education Building, 3700 McTavish Street).

Annex 2. Pre-selection questionnaire

Gender (choose one): Male ________Female______Choose Not to Answer____________

Have you ever been diagnosed or do you believe you have some level of color blindness (difficulty differentiating between red and blue or red and green for example)? Yes No

What is your current age? ___ years old

List the languages you speak, starting with your most fluent or best:_________________________

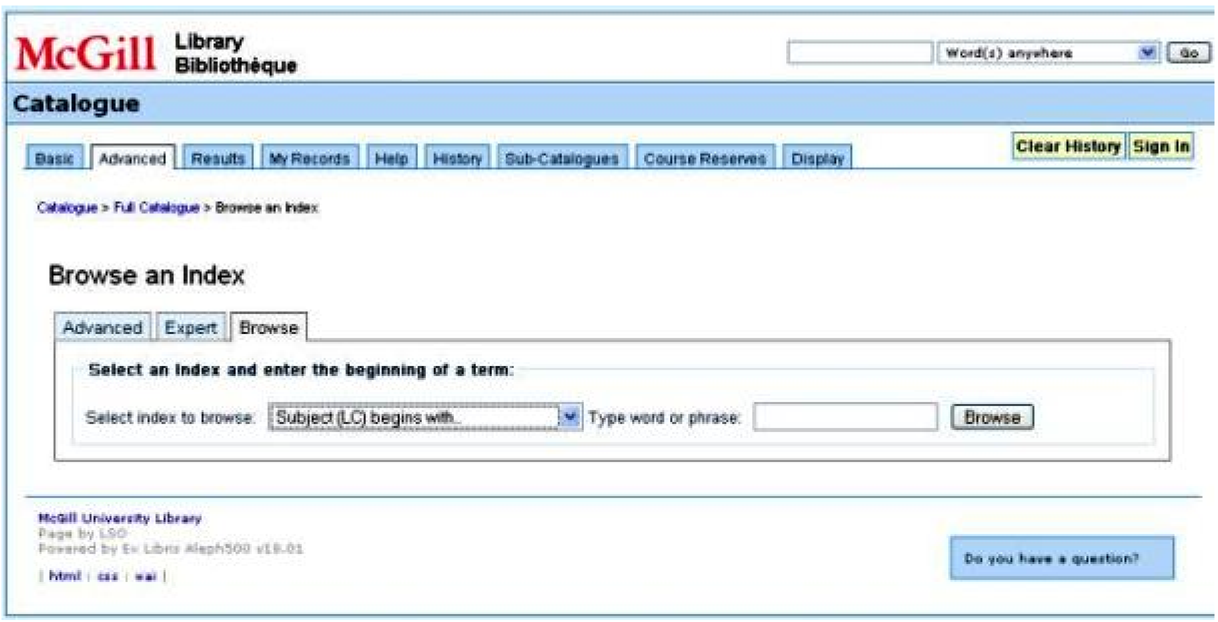
The answers you provide to the following four (4) questions will dictate if you are selected for this study (and receive $20 when completed). There are no wrong or right answers and we are looking for your current personal understanding.

Question 1. Please estimate how often you have used the McGill Library Online Catalogue MUSE (choose one):

Never; Less than 5 times; Less than 10 times; Less than 20 times; More than 20 Times

Question 2. In MUSE, how often have you used the subject index browsing feature shown below?

Never; Less than 5 times; Less than 10 times; Less than 20 times; More than 20 Times



Annex Figure 1. MUSE Subject Browse Feature

Question 3.

- How familiar are you with the Library of Congress Subject Headings or LCSH (Choose one):

Not at all, don't know / I have heard of it but I'm not sure I understand what it is / I understand what it is but I never use it / I sometimes use it when searching / I often search using LCSH

Annex 3. Informed Consent Form

Dear McGill University Student:

I am conducting a study about computer tools for searching. Your participation in the study will provide you with opportunities to learn about new searching tools and how we make sure they are well made for people to use.

If you agree to participate, you will be asked to meet with a researcher for one session that will last for about 60 minutes. The session takes place at a prearranged time at our lab in the Education building. The activities consist of using search interfaces and answering questions with them. Notes will be taken during the session. You will receive a 20$ compensation once you have completed the activities.

If you are willing to participate in this out-of-class session, please sign and return the consent form at the bottom of this letter. If you have any further questions about the study, you may contact me directly email.

Charles-Antoine Julien
Ph.D. Candidate, SIS
charles.julien@mail.mcgill.ca

Prof. Catherine Guastavino (Supervisor)
catherine.guastavino@mcgill.ca

Your signature below serves to signify that you agree to participate in this study.

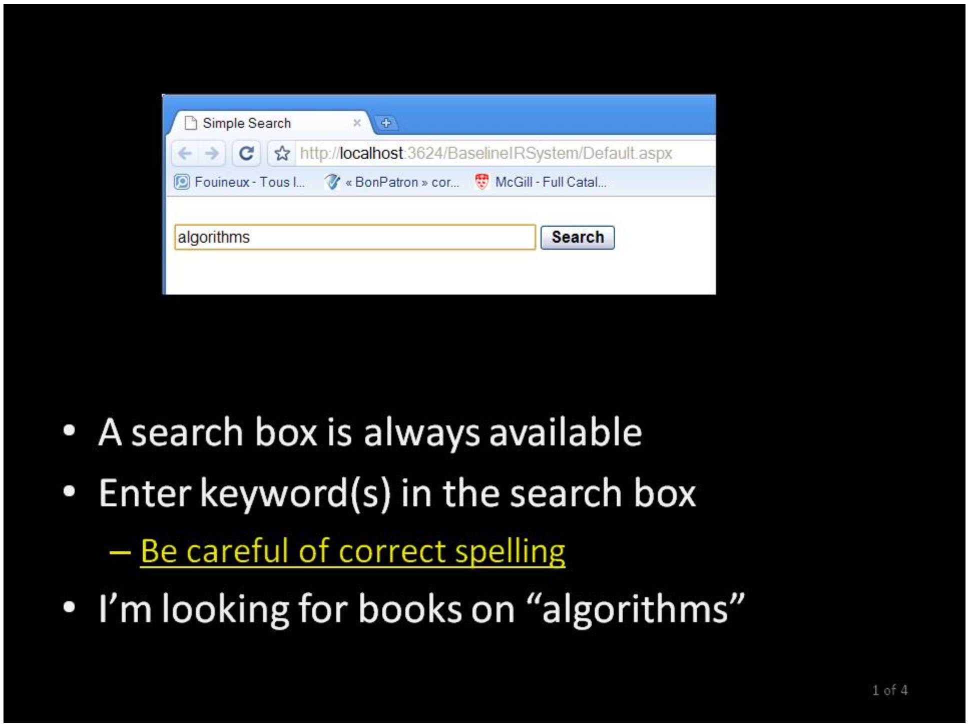
Your participation is entirely voluntary and you can choose to decline to answer any question or even to withdraw at any point from the project. Anything you say will only be attributed to you with your permission; otherwise the information will be reported in such a way as to make direct association with yourself impossible. My pledge to confidentiality also means that no other person or organization will have access to the interview materials and that they will be coded and stored in such as way as to make it impossible to identify them directly with any individual (e.g. they will be organized by number rather than by name)

I have read the above information and I agree to participate in this study

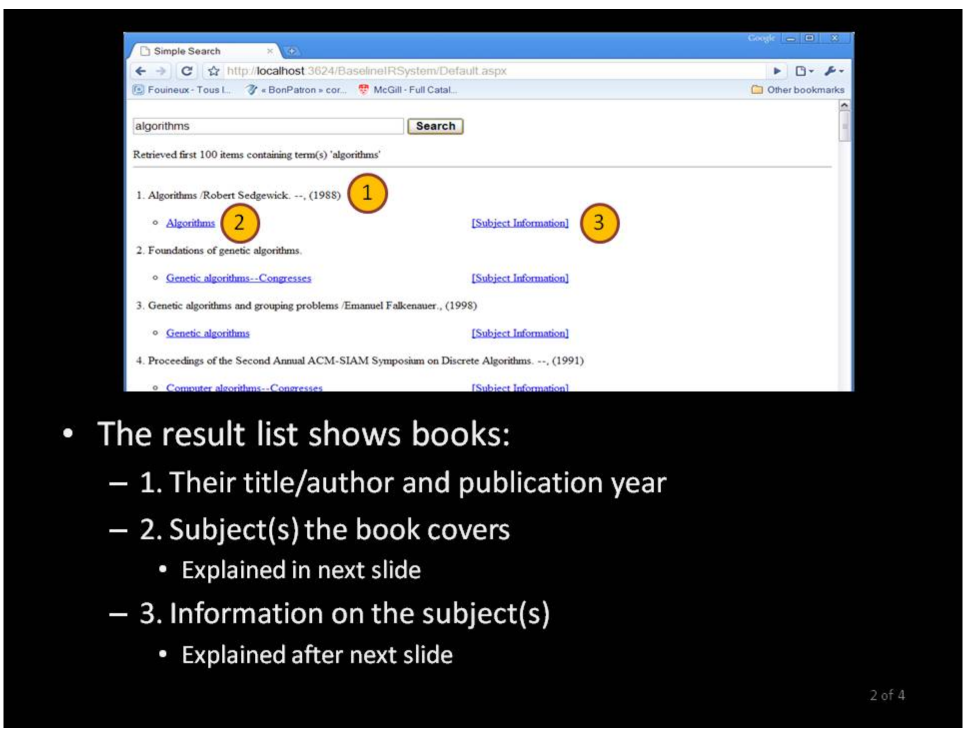
Signature: ___________________ Researcher's signature: _____________________

Name: _____________________________ Date: ______________________________

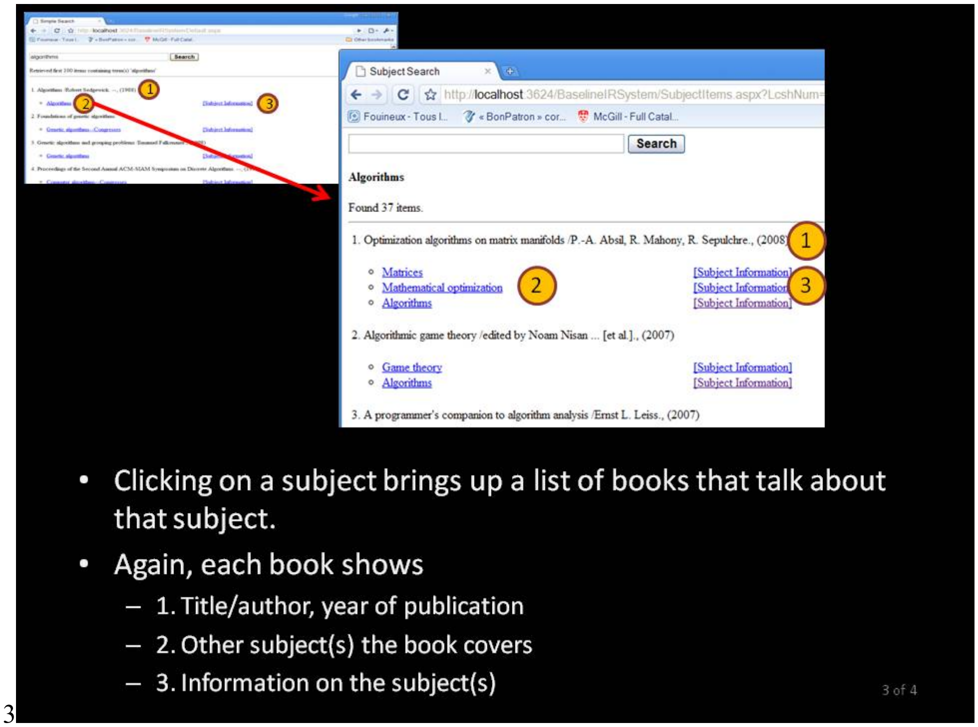
Annex 4. Baseline Web System Training Slides



Annex Figure 2. Baseline Web System Training Slide 1

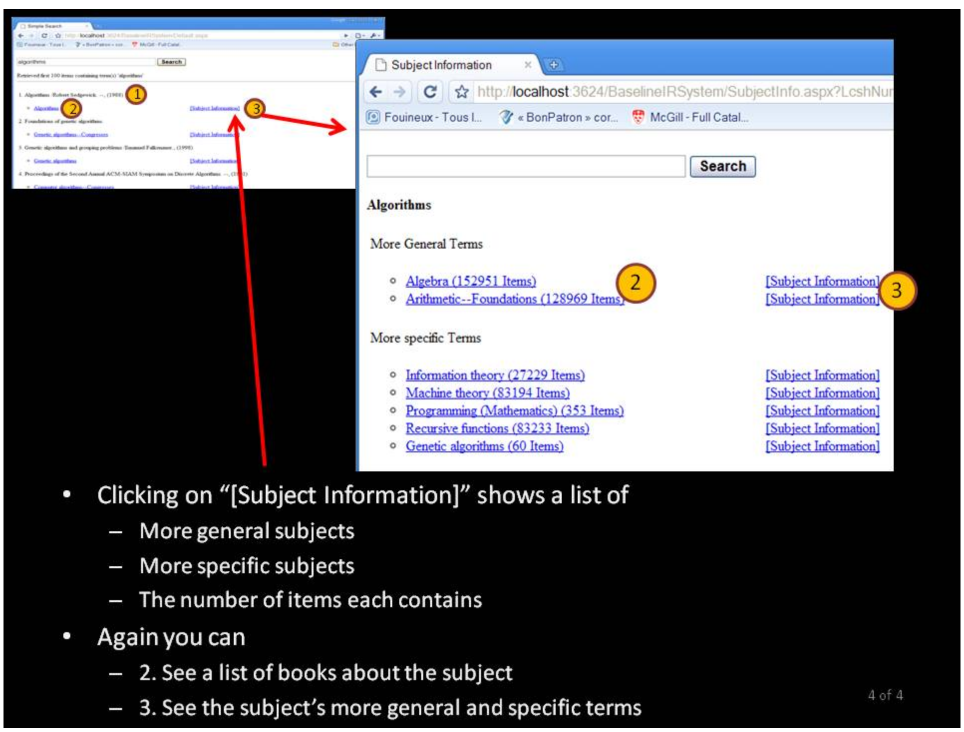


Annex Figure 3. Baseline Web System Training Slide 2



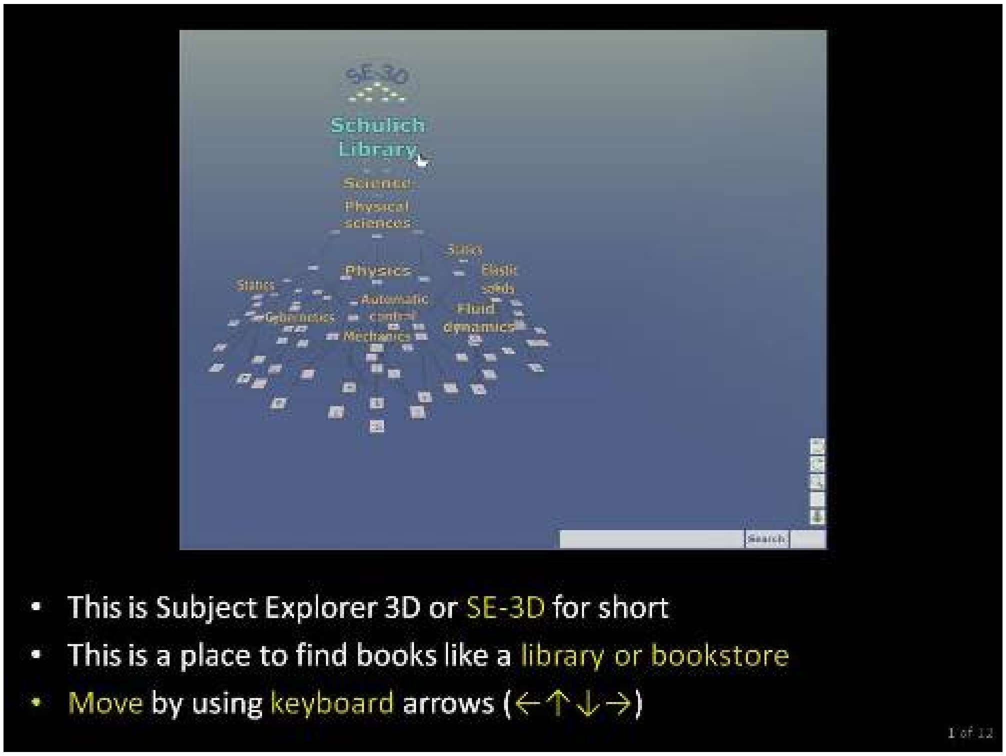
3

Annex Figure 4. Baseline Web System Training Slide 3

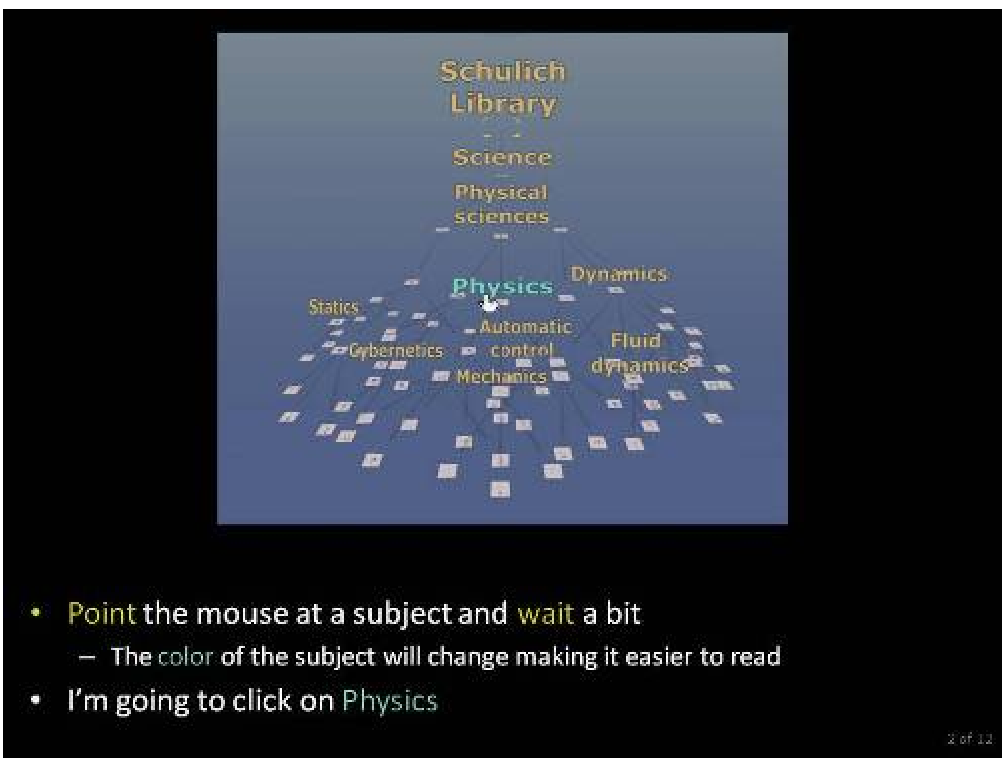


Annex Figure 5. Baseline Web System Training Slide 4

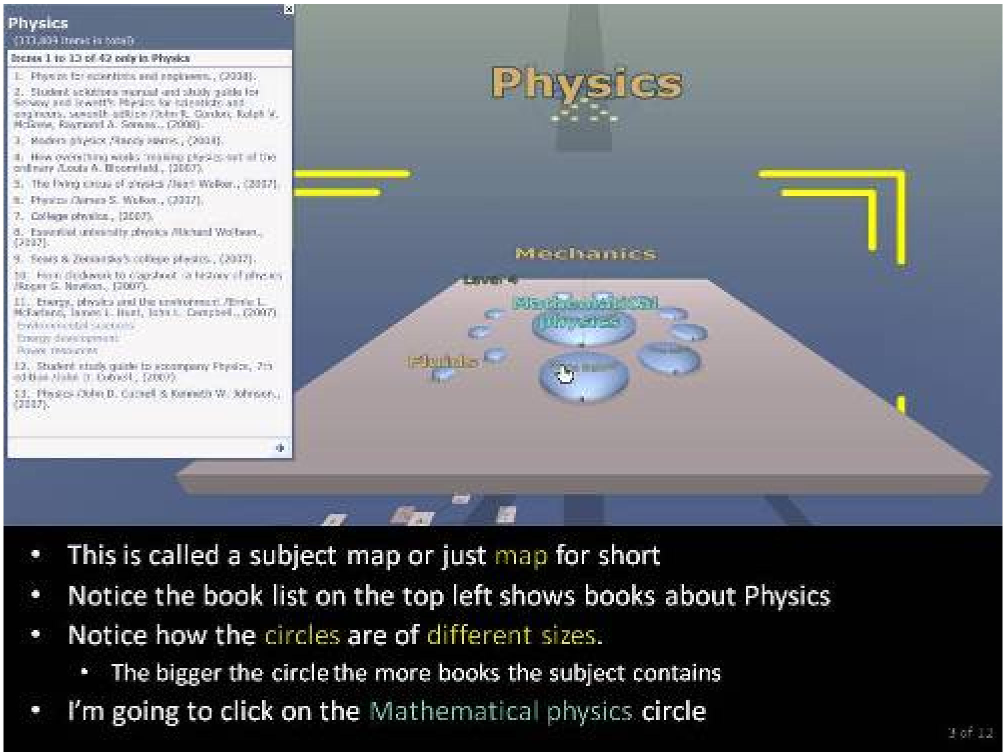
Annex 5. SE-3D Training Slides



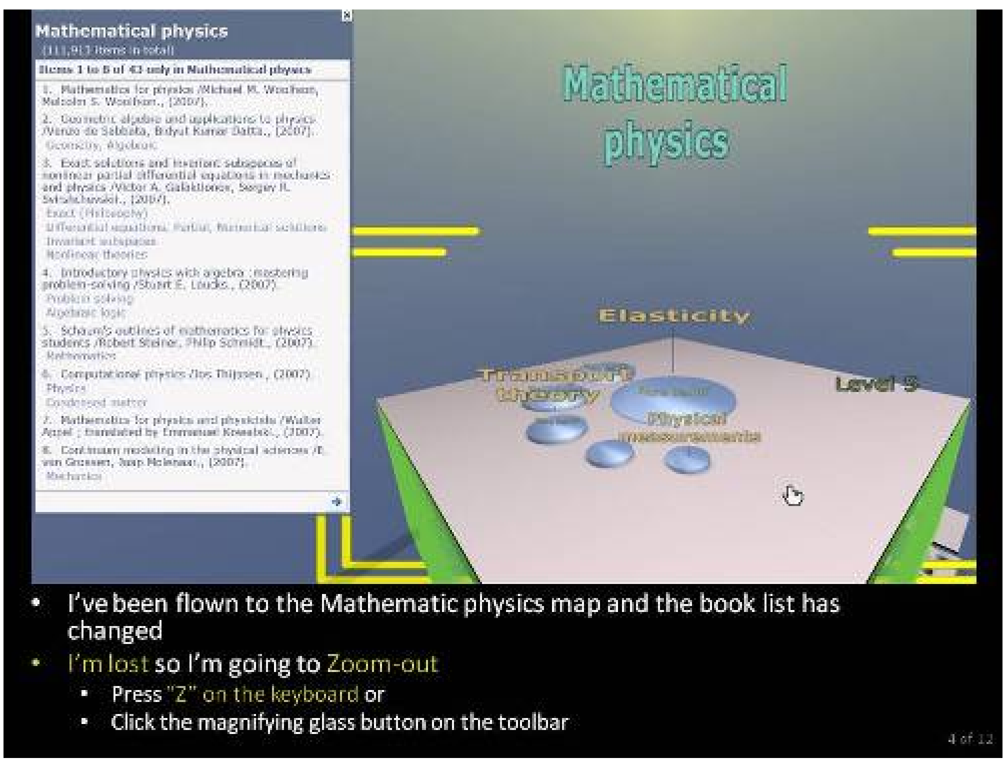
Annex Figure 6. SE-3D Training Slide 1



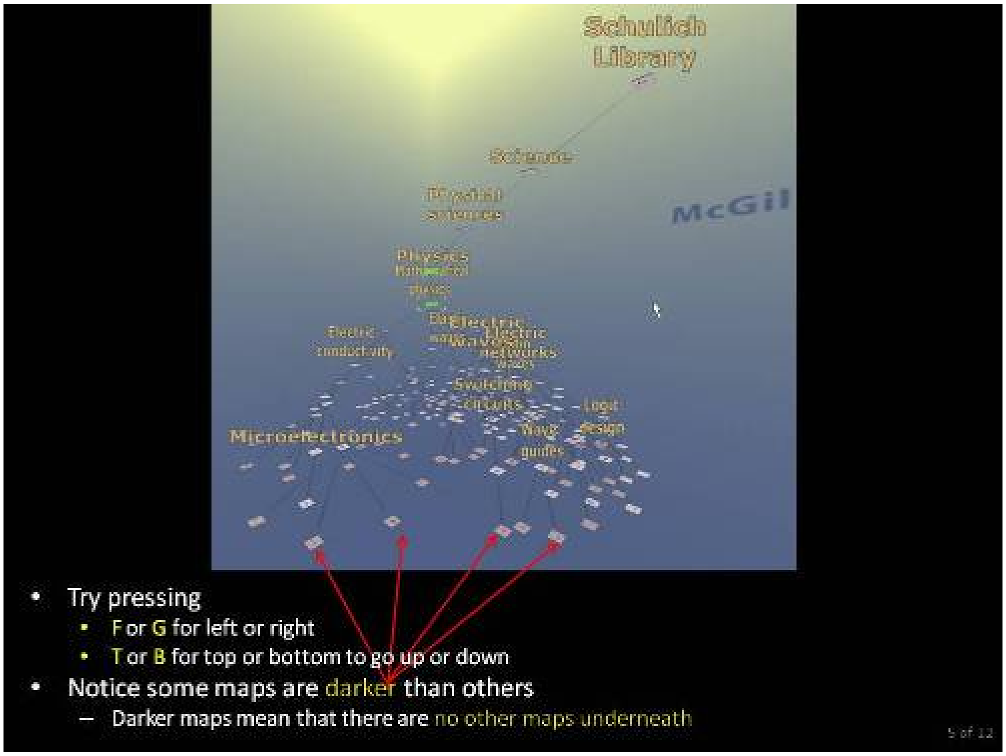
Annex Figure 7. SE-3D Training Slide 2



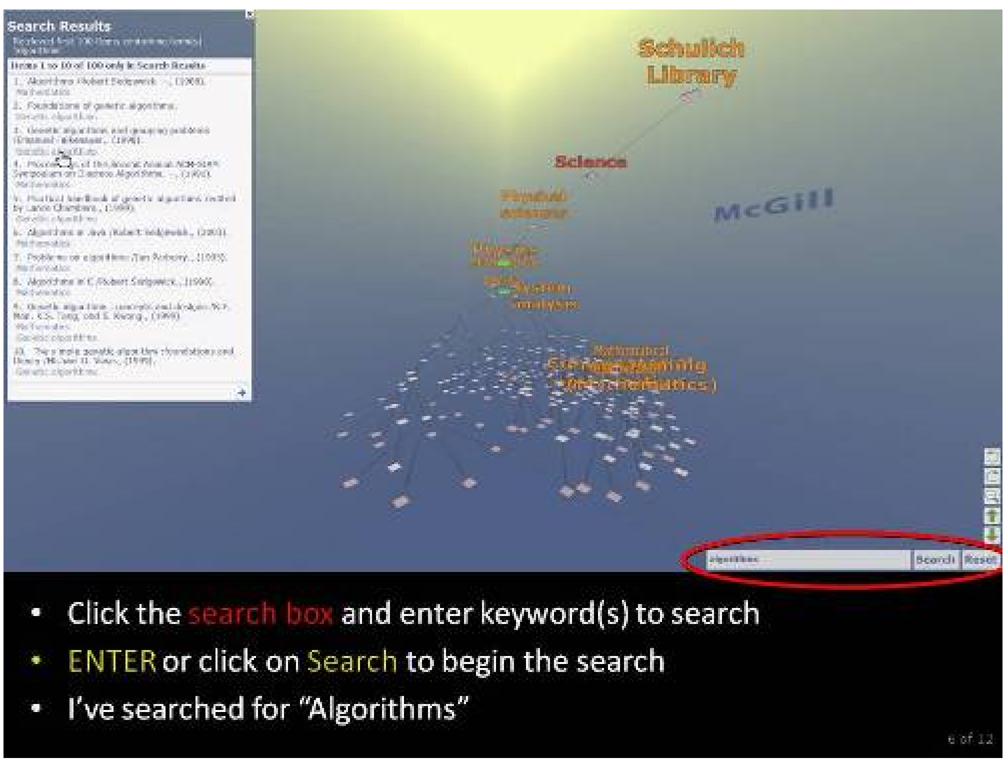
Annex Figure 8. SE-3D Training Slide 3



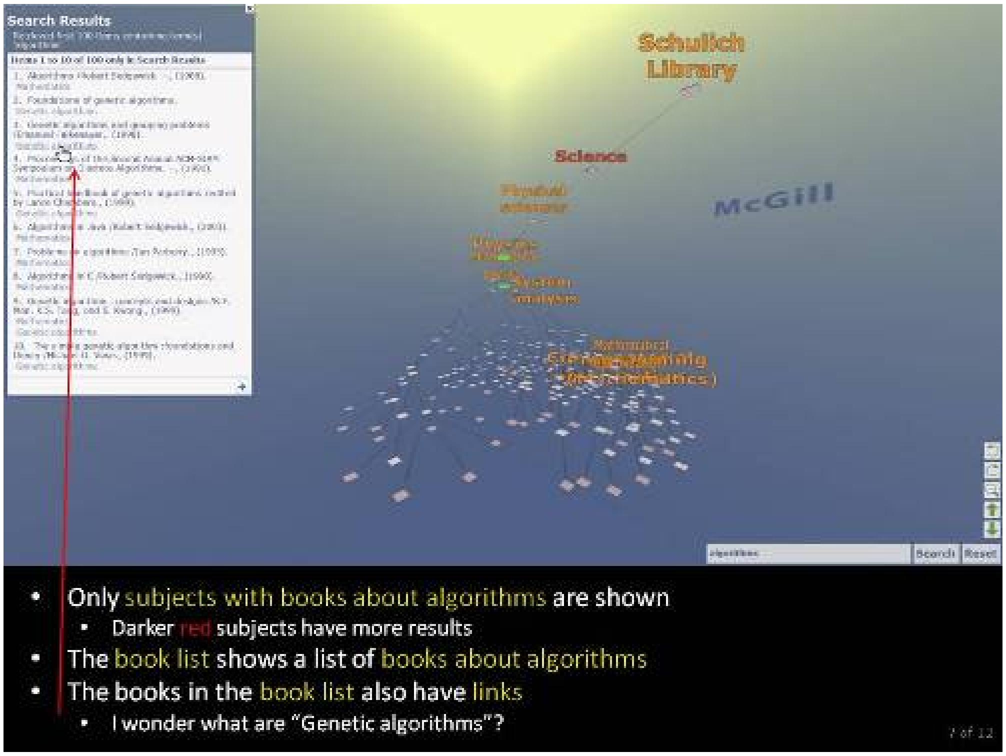
Annex Figure 9. SE-3D Training Slide 4



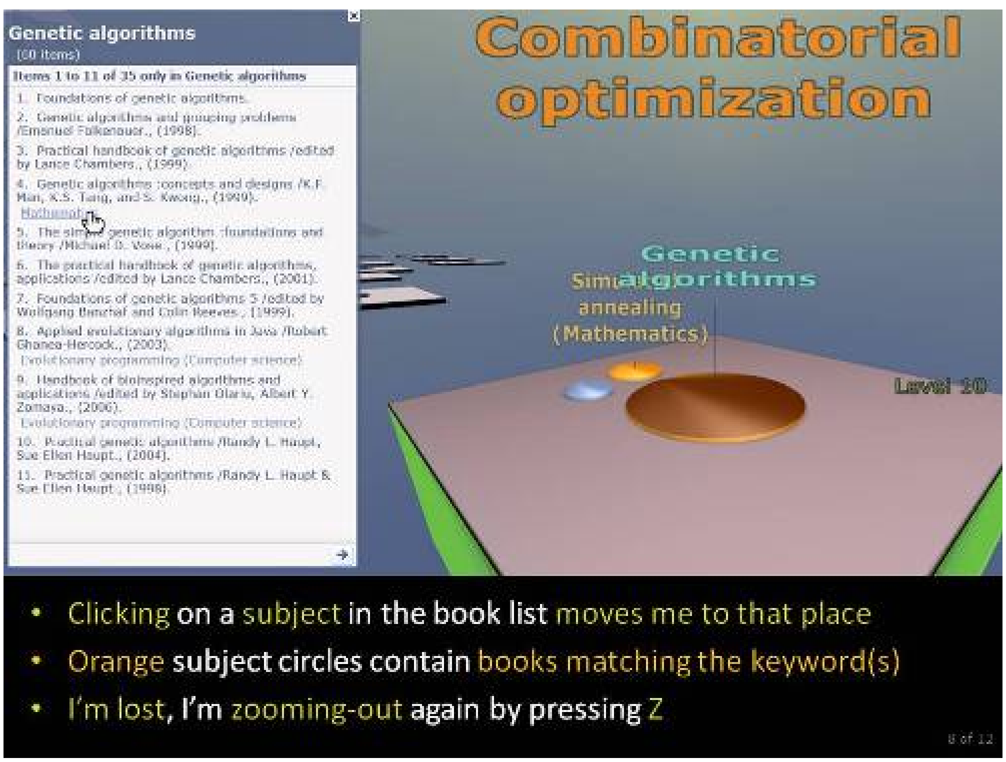
Annex Figure 10. SE-3D Training Slide 5



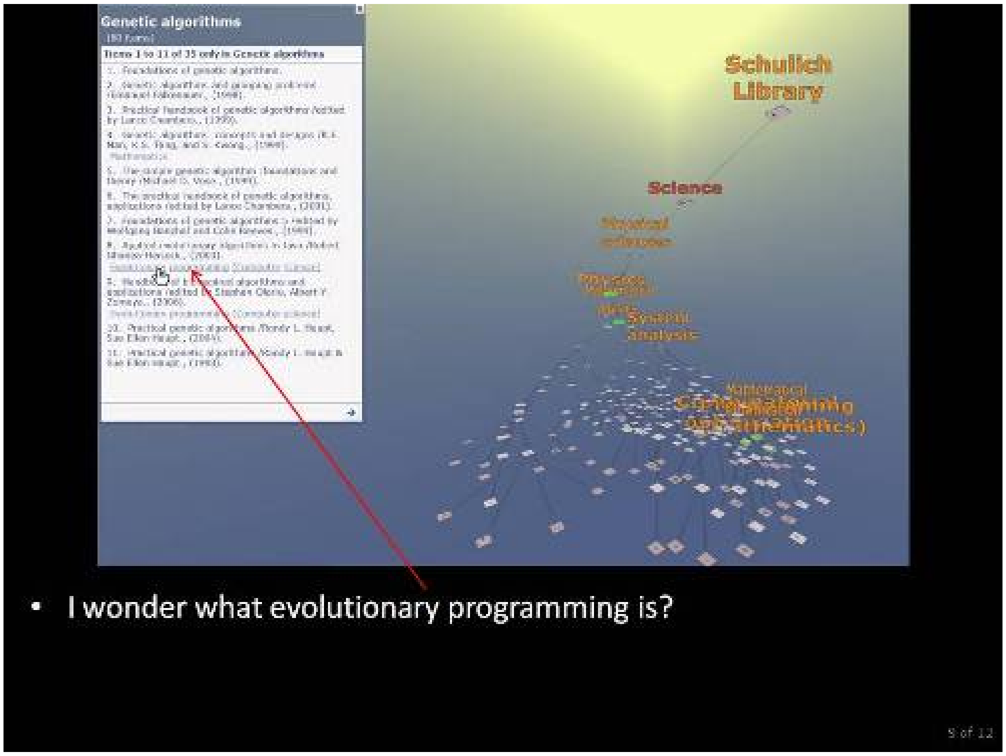
Annex Figure 11. SE-3D Training Slide 6



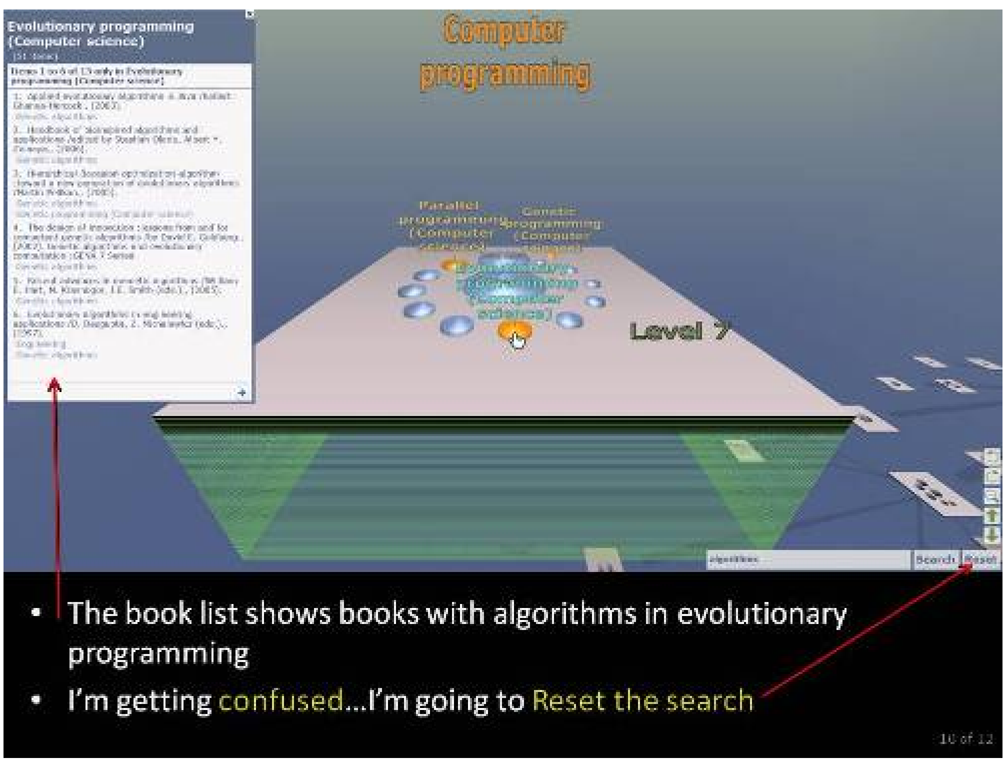
Annex Figure 12. SE-3D Training Slide 7



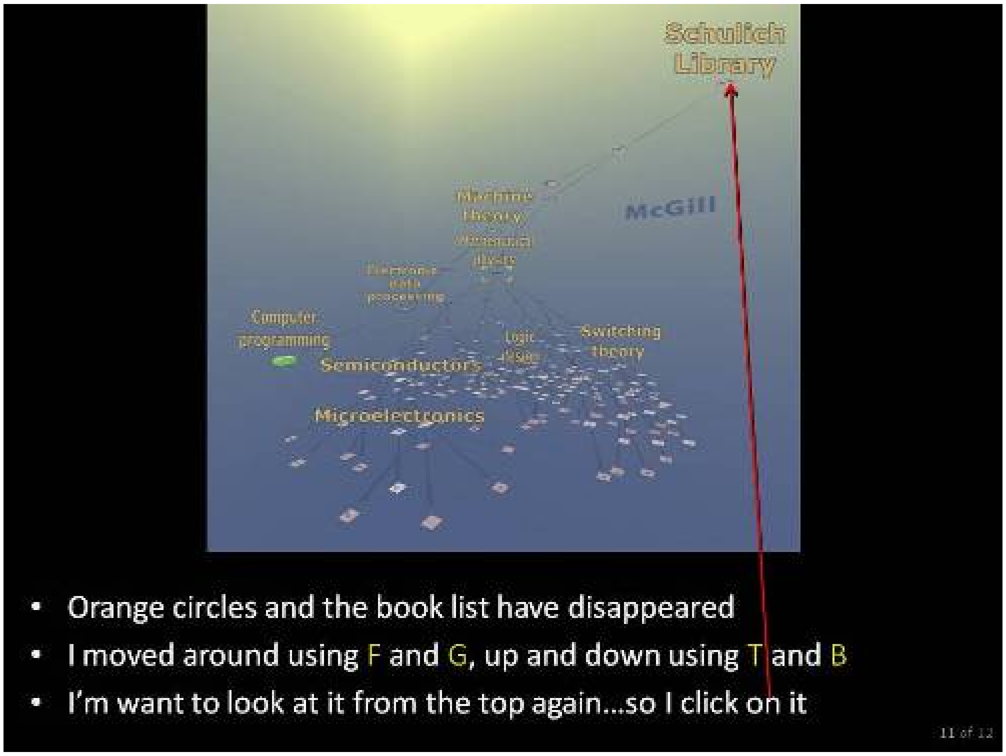
Annex Figure 13. SE-3D Training Slide 8



Annex Figure 14. SE-3D Training Slide 9



Annex Figure 15. SE-3D Training Slide 10



Annex Figure 16. SE-3D Training Slide 11

Annex Figure 17. SE-3D Training Slide 12

Annex 6. Stop List

List of words excluded from the search index. Stop terms are language-dependant and typically conjunctions, pronouns and articles which occur often but are usually of little value of topical searching.

a	it
about	la
an	le
and	of
are	on
as	or
at	that
be	the
by	their
com	this
de	to
du	und
en	was
for	what
from	when
how	where
I	who
in	will
is	with