Information Science and Information Systems: Converging or Diverging?

Ira A. Monarch

Carnegie Mellon University

Abstract

Recent studies have claimed a disconnect between the disciplines of information science and information systems even though, *prima facie*, there seems to be considerable overlap or potential overlap in their respective subject matter. The present study will target representative journals in the areas of information science and information systems and examine in more detail the overlap or lack of overlap between the two fields as reflected in the co-word analysis of the titles and abstracts of these journal articles. That the subject matters of the two fields can be combined in a discipline will be shown by a similar analysis of a third field, medical informatics, a new discipline in it its own right and a seeming subject matter hybrid of information science and information systems.

1. Introduction

In a recent article in JASIS, Ellis et al. (1999) present the results of their investigation on the relations between the fields of Information Science (ISci) and Information Systems (IS) using citation analysis. They claim that while the two fields seem to exhibit considerable overlap or potential overlap, they found very little evidence of links or overlap in the literature. They explain that the absence of such indicators stems from a desire on the part of the respective members of these disciplines to focus attention on and realize their own disciplinary interests. According to Ellis et al. the desire to maintain separateness is an attempt to gain and use power to advance individuals and groups within institutions and beyond. The present study will not challenge the viability of the co-citation approach nor the idea that research is politics by other means. However, it will examine the assumptions made in the Ellis study manifested in its characterization of ISci and

IS subject matters. The present study will show that these assumptions and the characterizations that followed from them do not provide the necessary insight into the subject matter needed for identification and mapping of links and common interests among disciplines. In addition, the present study will challenge the idea that gaining political power is always a matter of dividing and conquering. It will provide another view of political power - that of creating something new out of heterogeneous elements - using techniques of co-word analysis and the interpretive the theoretical backdrop of actor networks (Callon et al. 1986). Through a co-word analysis of representative periodical literature, a potential hybrid of ISci and IS - Medical Informatics (MI) - will be exhibited. If such a hybrid exists, then the linking of the subject matters of ISci and IS into a common set of interests is a reality. In any case, the aim of the study is to reach a better understanding of the possibility of joint or cooperative work in ISci and IS and also to reach a better understanding of how to characterize the subject matter of hybrid disciplines.

2. Initial Characterizations of the Subject Matter of ISci and IS

There are several examples, in the Ellis study, of characterizations of ISci and IS by indigenous groups that indicate essential commonality between the two. The Institute for Information Scientists (IIS) in the United Kingdom in their *Criteria for Information Science* (1996) specifically and prominently refer to

> broad concepts and theories of information systems and information and communication technologies insofar as they apply to the principles and practices of information management. (The Institute for Information Scientists, 1996, p.1)

Their criteria include specification, identification, analysis, implementation, evaluation and utilization of manual and electronic systems and tools; information generation, communication and utilization; information management and organizational context; information environment and policy (The Institute for Information Scientists, 1996).

These criteria for ISci strongly suggest common interests with IS. However, Ellis et al. point out that ISci covers such a wide variety of subjects that not all could be close to the interests of IS researchers and practitioners. For this reason the Ellis study narrowed the focus of their investigation to subdisciplines of ISci where they claim the highest overlap of subject matter can be expected. Without much analysis and relying on scant references and cursory review, the areas of Information Retrieval (IR) and User Studies (US) are selected. The Ellis study then cites Hjorland (1997) as underlining the centrality of IR and US to ISci and claims that even a cursory review of the American Society for Information Science Annual Review of Information in recent years would bear this out. This may be accurate at the level of granularity in which the claim is made, but is such a cursory review and analysis adequate for characterizing whether and how the subject matter of the IR and US disciplines converge or might converge? Even more important, how applicable is the claim that concern with users and systems in IR and US is recognizably similar to that of IS? Again at a fairly coarse level of granularity and precision, it might be, but this coarse level of analysis is not conducive to making productive connections between disciplines.

The Ellis study makes similar claims about common interests of IS with IR and US proposed by some in IS. In this regard they refer to the United Kingdom Academy for Information Systems (UKAIS). The latter group has provided an outline of IS’ s domain of study and an outline of its eight theoretical underpinnings. There is as follows.

> *Definition:* Information Systems are the means by which organizations and people, utilizing information technologies, gather, process store, use and disseminate information.
>
> *Domain of Study:* The domain of IS requires a multidisciplinary approach to studying the range of socio-technical phenomena which determine their development, use and effects in organizations and society.

Theoretical Underpinnings of IS

1. Data, information and knowledge management
2. Information in organizational decision making
3. Integration of IS with organizational strategy and development
4. Information systems design
5. Development and maintenance of IS
6. Management of IS and services
7. Organizational, social and cultural effects of technology-based IS
8. Economic effects of technology-based IS (UKAIS 1997)

Ellis et al. refer to several other IS researchers who have been active in debating the nature of the field. These include Checkland and Holwel (1998), who reiterate some of these themes.

> Nowadays we take the core concern to the field to be the orderly provision of data and information within an organization using IT, that information being relevant to the ever-changing activity of the organization and/or its members. (Checkland and Holwel 1998, 39)

The Ellis study claims that these definitions and delimitations have a great deal in common with those put forward by IIS, which they do. However, as shall be seen, the co-word analysis performed in the study described below does not show this kind of overlap without some major qualifications. In addition the above definitions and delimitations omit important elements of the IS knowledge network that co-word analysis of the IS journal titles and abstracts find. In general, the multiplicity of points of view represented in the numerous articles from ISci, IS and MI journals analyzed by the co-word approach, counters, somewhat, biases that can be attributed to any single individual or group. Though the interpretation of the maps is subject to bias, the words and phrases in the maps come from the authors themselves. Thousands of multiple journal articles represent a broader cross section of the field than can even a fairly diversified body of individuals attempting to define and delimit a discipline.

Ellis et al. (1999) conclude their paper by initially reiterating their thesis that ISci and IS are conjunct subjects, in terms of their focus of interests, but remain disjunct disciplines in terms of their disciplinary recognitions. Part of their explanation for this is that while there is a seeming similarity in methods providing an overlap of interest, it is a superficial similarity. They claim, in contradistinction to their previous claims and references, that IR research focuses almost exclusively on the information content of the system and deals predominantly with textual information. On the other hand, they claim that IS research focuses more on the formal modeling of relationships in data and on the organizational context of the system. Similarly, they maintain that US are predominantly occupied with the use of information services such as libraries or reference databases, or with the use of channels of communication such as journals, books and conferences, whereas IS research focuses on the individual's function or role, and on the demands that may be made of formal, data-based IS. In other words their view is that ISci research tends to be concerned with the information content of systems and with the development of more effective information services, while IS research is more concerned with formal organizational relationships to data and the development of more efficient computer-based systems.

In a curious way, the Ellis study's explanation of how ISci and IS can be conjunct subjects but disjunct disciplines undercuts their claim that they are conjunct subjects. Their flip-flop actually supports the co-word analysis approach to disciplines as knowledge networks by demonstrating that the multiple crisscrossing links of a network can seem to support contradictory points of view depending on which parts of a network are emphasized. While the co-word analysis also uncovers some of the differences between ISci and IS that Ellis et al. acknowledge at the end of their study, co-word analysis also shows how these disparities can be combined in the hybrid network of MI. While the study reported below stops short of actually confirming the hybrid character of MI - such a confirmation would also have to appeal to extensive citation and co-citation analyses - it does provide the beginnings of a blueprint of what elements and links would have to be co-cited.

3. Selecting Journal Articles for Co-Word Analysis

That journal articles were selected as the sole kind of document to be considered for co-word analysis was done, to a certain extent, for practical reasons. Journal articles were the one kind of document where the probability was quite high that all authors, titles and abstracts for every article could be fairly easily and quickly retrieved online. This was not true, for example of Proceedings from Conferences and Annual Reviews. That this course was followed is not to say that the study would not have been more complete if these and other kinds of documents were included. However, the aim of this paper, given time and resource constraints is not to do an exhaustive analysis but only to show the usefulness of the co-word analysis approach in studying the relationships among disciplines and their linkability for enabling new and interesting research and development. Even so, the highly cited journals in a field are one of the better sources of information in determining the subject matter of a disciplinary field.

Representative journals for the years 1990-2000 were identified iteratively based on published expert opinion, their inclusion in the Institute for Scientific Information's (ISI'S) most cited journal list[1] and their co-word similarity to journals most commonly acknowledged as representative of ISci and IS. In the case of ISci the representative journal chosen was the *Journal of the American Society for Information Science (JASIS)*.[2] The claim is not that *JASIS* is the most representative, only that it is one of several that are. The same claim of representativeness is made for *Management Information Science Quarterly (MISQ)* with respect to IS journals. In the case of MI only Medical Informatics journals included in ISI's most cited list were selected - three in all - the *Bulletin of the Medical Library Association, International Journal for Medical Informatics* and *Methods of Information in Medicine*.[3] The total number of articles analyzed for MI was 1079.

There were four other journals selected for ISci - *Information Processing & Management, International Forum on Information and Documentation, Journal of Documentation* and *Journal of Information Science*. These were selected based on Smith's (1999) discussion of journals competing with *JASIS*, their high citation rank according to the ISI, and their accessibility. The total number of articles analyzed for ISci was 2026. There were also four other journals initially selected for analysis of IS - *International Journal of Information Management, Journal of Management Information Systems, Information and Management* and *Information Systems*. These were selected based on the discussion of Ellis et al., their high citation rank according to ISI and their accessibility. In a preliminary round of co-word analyses of these journal, one, *Information Systems*, was found to be much more technically oriented with respect to databases and much different in subject matter from the other four in this group. The final round of co-word analyses on which the results of the present study are based was not applied to this journal. Without it no relationship with "database" made it into the leximaps for IS. The word "database" was used about 150 times in the corpus consisting of the four included IS journals, but the word did not have sufficient co-occurrence strength with other terms in this corpus to make it into any IS leximap. "Database" not appearing in any IS leximap is especially interesting because it appeared in 18 of the 32 leximaps for ISci and MI. "Data" appeared in 35 of the 47 leximaps of the combined set. In light of what we have found in this study, the *Information Systems* journal will be included in future co-word analyses of IS[4]. The total number of articles analyzed for IS was 1325.

4. Co-Word Analysis

There is a fairly extensive literature on co-word analysis (Callon, Law, & Rip, 1986; Callon, Courtial, & Laville, 1991; Courtial, 1994; Courtial & Law, 1989; Law & Whittaker, 1992; Turner et al., 1988; Whittaker, 1989). Co-word analysis reveals patterns and trends in technical discourse by measuring the association strengths of terms representative of relevant publications and sometimes other texts produced in a technical field. A main tenet of co-word analysis is that the identified patterns of representative term associations are maps of the conceptual structure or knowledge network of a technical field and that a series of such maps produces a fairly detailed representation of the subject matter of a discipline.

Co-word analysis is related to co-citation analysis (Small, 1973; Small & Griffith, 1974). Co-citation analysis provides a method of mapping the structure of a research field through pairs of documents that are jointly cited. Co-word analysis deals directly with sets of terms shared by documents instead of with shared citations. Therefore, it maps the pertinent literature directly from the interactions of key terms instead of from the interactions of citations. While this paper will concentrate on co-word analysis, it would be interesting to investigate how co-citation analysis combined with co-word analysis can be used to represent the actor or knowledge networks that determine a discipline.

5. Details of the Metric Used in Co-Word Analysis

Co-word analysis enables the structuring of data at various levels of analysis: (1) as networks of links and nodes (nodes hold terms); links connect nodes, thereby forming networks); (2) as distributions of interacting networks; and (3) as transformation of networks over time periods. Future studies of the relations between ISci, IS and MI should take into account the history of the changes of co-word association networks in different time periods. This might be done in 5 year increments starting around 1980 for ISci and IS and starting in the early 90s for MI. Seeing how the networks change in different time periods would another important indicator of possible cross-over links between ISci and IS.

Co-word analysis reduces a large space of related terms (words and phrases) to multiple related smaller spaces that are easier to comprehend but are also indicative of actual partitions of interrelated concepts in the literature under consideration. This analysis requires an association measure and an algorithm for searching through a term space. The analysis is designed to identify areas of strong focus that interrelate. This scheme allows us to construct a mosaic of ISci, IS and MI topics.

Metrics for co-word analysis have been studied extensively (Callon, Law, & Rip, 1986; Callon, Courtial, & Laville, 1991; Courtial & Law, 1989; Law & Whittaker, 1992; Whittaker, 1989). Two terms, *i* and *j*, co-occur if they are used together in the classification of a single document. Take a corpus consisting of N documents. Each document is indexed by a set of unique terms that can occur in multiple documents. Let c_k be the number of occurrences of term *k* ; i.e., the number of times *k* is used for indexing documents in the corpus. Let c_{ij} be the number of co-occurrences of terms *i* and *j* and (the number of documents indexed by both terms).

Different measures of association have been proposed. The inclusion index I_{ij} provides a association hierarchy metric (essentially a conditional probability) through the function:

$$I_{ij} = c_{ij} / \min(c_i , c_j)$$

I_{ij} is not symmetrical (or bi-directional) and tends to highlight mainly the central poles in a domain and depict their relations with terms that occur less frequently (Callon, Law, & Rip, 1986). This metric does not do a very good job of identifying inclusion hierarchies at least in anything that coheres with common semantic relations like part-whole and subcategorization. Given that it failed in this regard, there was no other good reason to use it.

The basic metric used for this study is *Strength* S_{ij}. The Strength of association between terms *i* and *j* and is given by the expression:

$$S_{ij} = c_{ij} . c_{ij} / c_i . c_j , 0 <= S <= 1$$

This metric provides an intuitive measure of the strength of association between terms indicating only that there is some semantic relationship or other. The metric is easier to understand and utilize in the production and interpretation of term association maps than is the so-called inclusion metric. It allows associations of both major and minor terms and is symmetrical in their relationships (Callon, Courtial, & Laville, 1991). S can be used as the basis for several complementary measures of interactions of terms and term networks in a unified manner.

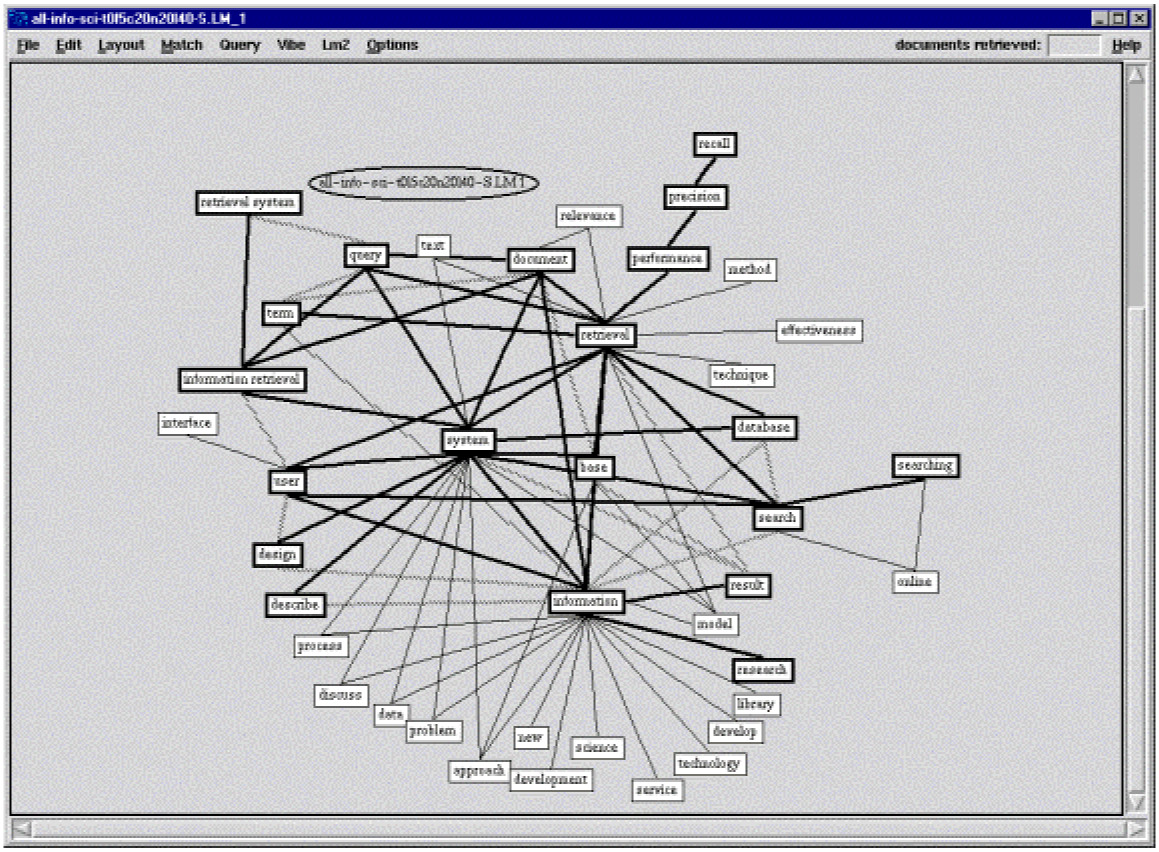
Two terms that appear many times in isolation but only a few times together will yield a lower S value than two terms that appear relatively less often alone but have a higher ratio of co-occurrences. Terms with relatively high S values form the networks' links. A term network consists of nodes (terms) connected by links. Each node must be linked to at least one other node in a network.

The co-word algorithm uses two passes through the data to produce pair-wise connections of terms in leximaps[5] which partially represent a term network (see Figure 1 which contains Map1 from the set of Information Science leximaps). Pass-1 builds networks that can identify areas of strong focus; Pass-2 can identify terms that associate in more than one leximap of a network and thereby indicate pervasive issues. This pattern of leximaps yields a mosaic of the data being analyzed.

The first pass (Pass-1) generates the primary associations among terms; these terms are called internal nodes and the corresponding links are called internal links. A second pass (Pass-2) generates links between Pass-1 nodes across leximaps, thereby forming associations among complete leximaps. Pass-2 nodes and links are called external ones.

Figure 1 illustrates the outcome of this process for an Information Science network. This figure displays a partial view of the term network connections as a leximap. As will be explained below, this map is quite important in organizing the interpretation of the rest of the Information Science leximaps for the years 1990-2000. Pass-1 links and nodes are represented by thick lines connecting thick boxes, respectively. Pass-2 nodes are in thin boxes, while Pass-2 links are shown as thin lines connecting Pass-1 and Pass-2 nodes.

Figure 1: Information Science Leximap 1



Without some minimum constraints, terms appearing infrequently but almost always together could dominate leximaps; hence a minimum co-occurrence c_{ij} value is required to generate a link. At the same time, some maps can become cluttered due to an excessive number of legitimate links (but of generally decreasing S values); hence, restrictions on numbers of nodes and sometimes links are required to help discover major partitions of concepts. However, in many cases only the number of qualifying nodes limits many term networks.

To generate maps for Information Science and Information Systems in the time period 1990-2000, the cij constraint was 20 co-occurrences and the node and link constraints were 20 and 40 respectively. The field of Medical Informatics over a similar time period seems to have been less interconnected. The cij constraint was 15 co-occurrences, the node constraint 20 but no link constraint was required. For these values the co-word algorithm generated 18, 15 and 14 term networks, respectively, for Information Science, Information Systems and Medical Informatics respectively.

6. Details of the Co-Word Algorithm Used

During Pass-1, the link that has the highest strength is selected first. These linked nodes become the starting points for the first term network. Other links and their corresponding nodes are then determined breadth-first. All nodes contained in the resulting Pass-1network are removed from consideration for inclusion in subsequent Pass-1 networks. The next network then starts with the link of highest S value of the remaining links (i.e., ones not containing nodes from any previous network).

The second pass (Pass-2) is designed to seek further associations among terms found in Pass-1. During Pass-2, networks are extended by the addition of Pass-2 links. To be a candidate for inclusion in Pass-2, both nodes (terms) of a Pass-2 link must be in some Pass-1 networks. A Pass-2 link connects a Pass-1 node in a given network to a node that had occurred as a Pass-1 node in another network (but is represented in the given network as a Pass-2 node).

As in Pass-1, candidate links are included in Pass-2 based on their strengths and co-occurrence counts. The order of Pass-2 links is by descending values for qualifying links. A node can appear in only one Pass-1 network, but can appear in more than one Pass-2 link.

The steps in the analysis procedure are:

1. Select a minimum for the number of co-occurrences, cij, for terms i and j.
2. Select maxima for the number of Pass-1 links and nodes;
3. Select maxima for the total (Pass-1 and Pass-2) links and nodes;
4. Start Pass-1;
5. Generate the highest S value from all possible terms to begin a Pass-1 network;
6. From that link, form other links in a breadth-first manner until no more links are possible due to the co-occurrence minima or to Pass-1 link or node maxima. Remove all incorporated terms from the list of subsequent available Pass-1 terms;

7. Repeat Steps 5 and 6 until all Pass-1 networks are formed; i.e., until no two remaining terms co-occur frequently enough to begin a network;

8. Begin P ass-2;

9. Restore all Pass-1 terms to the list of available terms;

10. Starting with the first Pass-1 network, generate all links to Pass-1 nodes in that network with any Pass-1 nodes having at least the minimal co-occurrences in descending order of S value; stop when no remaining terms meet co-occurrence minima or when total node or link maxima are met. Do not remove any terms from the available list;

11. Repeat Step10 for each succeeding Pass-1 network.

A maximum number of Pass-1 networks can be specified in cases where an excessive number of networks will be generated otherwise; this restriction was not necessary here.

Link and node limitations mostly determine how networks will be generated in concert with the corresponding co-occurrence minimum. If the co-occurrence minimum is too high, few links may be formed; if it is too low, an excessive number of links may result. In the former case, subspecialties in a field may not emerge; in the latter case, more representative and well connected themes and problem spaces will be harder to detect amidst the noise of less representative and less well connected ones. We experimented with many different sets of parameters during our research to identify values that yield a detailed, yet coherent, set of term networks. In any case, the choices made depend on the level of granularity one is interested in achieving and the time one has to sort out the information generated when the constraints are relaxed.

7. The Software Tools Used

Co-word analysis is an active research area at Carnegie Mellon University’ s Software Engineering Institute (SEI)[6]. A suite of software tools developed at SEI called CAIR (Content Analysis and Information Retrieval) performs all of the calculations described here. CAIR operates on documents represented by index terms that it derives from free-text documents, using a phrase clustering technique on automatically extracted stemmed noun phrases and verb phrases. Recently CAIR was employed to analyze a free-text database of best engineering practices to extract which best practices might be applicable in software engineering (Coulter & Monarch, 1996). CAIR has been used extensively at SEI to analyze various collections of free text to assist in understanding and categorization of software engineering management practices (Monarch, 1994; Monarch & Gluch, 1995; Monarch, 1996). Several new projects are beginning in the Spring of 2000 that include mining huge text bases with information on software failures, risks, computer security intrusions and lessons learned

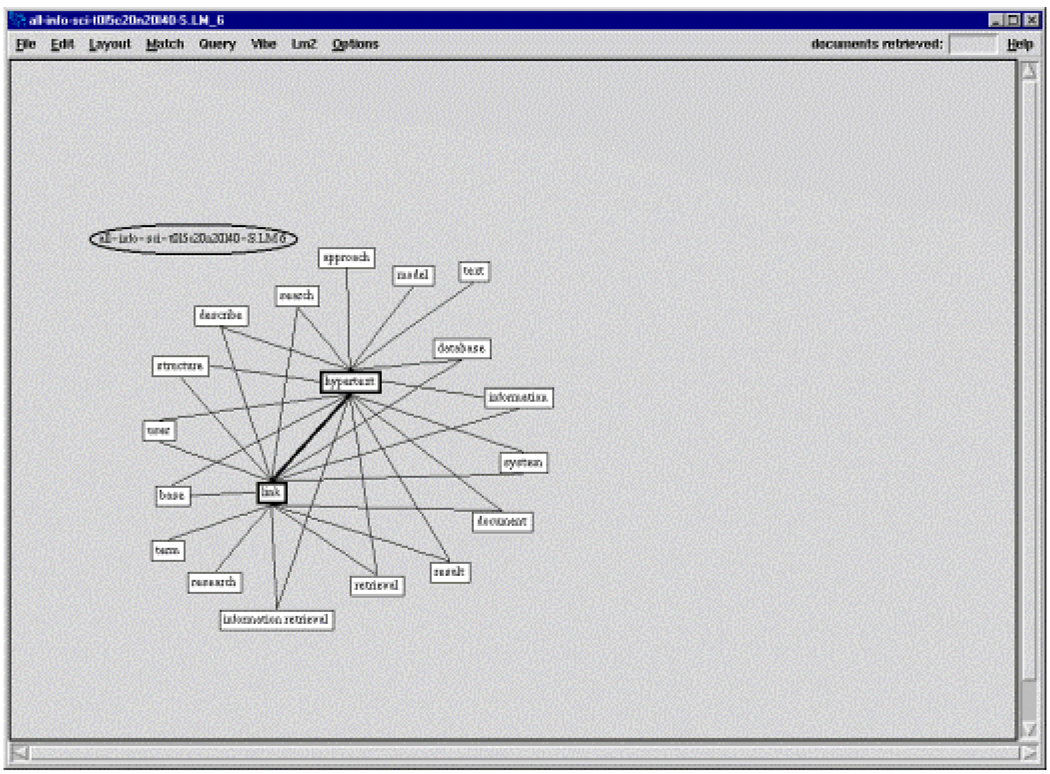
CAIR works with the probabilistic information retrieval system INQUERY[7] to allow users to query a database for specific documents whose key term associations are revealed through lexical maps. A graphical user interface (gui) connects term association maps to a document database and enables the formation of searches by selecting linked nodes or clustered document icons. The gui can be used to depict the semantic distance of documents with respect to one another and to the nodes on the leximap or term association map. It enables a selected map to be ranked against other maps for similarity and dissimilarity. It can be used to create a reduced document set derived from an ongoing search from which a new set of term association maps can be derived. It also provides other convenient features such as automating layout of networks and modifying them for enhanced presentation.

8. Interpreting Maps and Structures of Leximaps

Leximaps can be interpreted and primary focus points can be identified as intersections of important themes. This is not a precise activity. In some cases, leximaps have a single, highly connected pair of terms; in others, there are five or more terms exhibiting high connectivity, though only two or three stand out as primary. In the former cases, the theme can be a word or a short phrase. In the latter cases, several words and phrases are needed. In most cases, the main themes of leximaps can be can be identified using words and phrases contained in the map. Generally, these are term(s) from node(s) with the most links. In determining themes of maps, especially primary themes, most consideration is given to Pass-1 nodes.

For example in Figure 1, information retrieval system[8] is the primary focal point or theme. User, query, document, database, search, result and model are themselves well-connected, and all are linked to the primary theme so are also identified as part of the theme. However, in Figure 2 (leximap 6 of the Information Science set), hypertext link clearly captures the theme of that map.

Figure 2: Information Science Leximap 6



For reference in the following sections, the primary network themes for each set of leximaps, Information Science, Information Systems and Medical Informatics for the time period 1990-2000 are given in Table 1. Networks are listed in the order generated by co-word analysis algorithms for each set. The Table will be utilized in subsequent discussions.

Table 1: Network Themes

Information Science	Cpl/Coh	Information Systems	Cpl/Coh	Medical Informatics	Cpl/Coh
1.Information Retrieval System - User, Query, Document, Database, Search, Result and Model	1 2	1. Information System - Support - Decision Support System (DSS) - Group (GSS, GDSS), Organization, Technology; Executive (EIS); Process; Develop	1 3	1. Information System - Knowledge Representation Language (Semantics); Patient Data and Record (Electronic); Health Science & Library (Access - Internet)	1 1
2. New WWW, Communication Technology, Development; Model, User, Approach	2 3	2. User (Satisfaction; Design) Model (Test) and Data (Interchange, EDI Electronic), Analysis	3 5	2. Service (Library, Hospital, Professional, University, Health) - Internet (WWW), Access (National) Available Resources)	4 2
3. Automated Text Techniques/Methods and Indexing; Data	3 4	3. Group (Participant Member) - GDSS, GSS, Work, Task, Experiment	2 4	3. User - Interface, Librarian,	2 4

(Analysis, Structure, Model, and SemanticRepresentation)		(Subject, Laboratory, Effect, Finding), Meeting, Participant		Online Search (Medline - Index, Literature), Database; Support Decision - Support System, Diagnostic, Knowledge - Model	
4. Library (Service, Research, Academic, Librarian, Digital, Resource) Access (Online) and Role in Science (Literature, Journal, Field, Discipline Citation)	5 6	4.InformationTechnology (Information Use, Manager, Impact) - Business (Firm) Change (Implementation), BPR, Strategic Planning (Resource), Competitive Advantage	4 6	4. Hospital (University - Faculty, Student, Curriculum, Training) Information System - Development - New Technology (Communication - Electronic Medical/Patient Record)	3 3
5. Measure retrieval Similarity; Relevance Feedback and Effectiveness;	4 7	5. United States	15 1	5. Diagnostic - Expert (System) - Decision, Disease	6 7

information retrieval system Process (Application, Cognitive, Design, state, Required, Decision Support)				Classification, Interpretation (ECG), Test; Analysis Method: Computer Program & Software	
6. Hypertext Link	14 1	6. Internal and External	14 2	6. Health Care - Professional, Provider - Network (Time, User, Access, Environment), Informatics - Education, Cost; Technology Change; Conduct Survey (Service, Library, User)	5 8
7. retrieval Evaluation, Algorithm Test (what significant effects found) wrt a Collection and Experiment;	6 9	7. Survey (Conduct, Questionnaire) Significant (Performances, Relationships Variability - Measure) Differences	5 9	7. Knowledge Structure - Problem, Domain, Object, Terminology; Physician - Clinic, Finding, Question	9 10
8. Knowledge - Human	7 13	8. Software (Tool,	6 10	8. Document	11 5

Interaction Tasks, Expert, Domain, Conceptual (Theory); Survey Findings Suggest; Relevant Retrieval		Engineering, Productivity) Application - Expert (Problem Solving) System - Knowledge (Domain, Acquisition, Skill)		Delivery	
9. Illustrate Example	18 5	9. Methodological, Conceptual, Theoretical Framework; Computer Communication; Researcher- Practitioner wrt IS	8 13	9. Design - Research (Patient Care), Implementation (Information System), Tool, Issue - Current - Review, Practice); Healthcare Standard; Information System and Software Architecture and Integration	7 9
10.Issues; Work - Group, Computer - Network (National International)	9 14	10.New - Service, Customer, Industry, Technology, Market, Company, Cost, Experience, Environment	7 12	10. Quality (Improvement) Information Management - Academic, Organization; Natural Language	8 13

				Processing and Concept Representation	
11.Recent Year	17 8	11. Efficient & Effective	13 7	11.Digital Image	14 6
12.Impact Factor (Influence)	15 12	12. Research and Practice (Implications and Context) and Literature - Value, Current/Future, Review, Importance	9 14	12.Relevant Data, Major Challenge, Sources - Library, Access, Survey, Database, Data	10 14
13.Implementation (Implement)	13 10	13.Recent Year	11 8	13.Functions and Parameters	13 11
14.View Point	16 11	14.Evaluation (Evaluate)	12 11	14.Work Requirement	12 12
15.Subject -Abstract, Record, Select, Investigate, Classification; InformationSystem Management; Software Project (system)	8 16	15.Structue (Form) and Function wrt Activity;	10 15		
16.Tools and Interfaces	10 15				
17.Order - Determine; Indicate	11 18				
18.Frameworks and Contexts	12 17				

9. Attributes of Leximaps and Relating them in Networks

Leximaps are related in three types of networks: principal, secondary, and isolated. Principal networks are connected to one or more (secondary) networks. Secondary networks generally are linked to principal networks through a relatively high number of external links in the principal networks. Isolated networks have an absence (or low intensity) of links with other networks. Isolated leximaps often have links with high S values, usually accompanied by low co-occurrence cij values. While isolated maps are easy to recognize, principal and secondary maps may not be. Therefore, we will define and operationalize terms that characterize these functionalities.

Cohesion is defined as the mean of the Pass-1 S values of a leximap; coupling is defined as the square root of the sum of the squares of the Pass-2 S values of a map in order to distinguish among relatively close values. Cohesion represents the internal strength of the links of a map, while coupling represents a map's position in strength of interaction with other maps.

Plots of cohesion and coupling for each of the three sets of leximaps are shown in Figures 3, 4 and 5. The origin of these figures is the median of the respective axis values (the horizontal axis represents coupling; the vertical axis represents cohesion).

Not surprisingly, most leximaps with strong coupling scores also show relatively high Pass-1 node and link counts. Isolated maps show relatively low Pass-1 link and node counts (see Figure 2 Information Science Leximap 6, for example).

Figure 3: Information Science Leximap Distributions

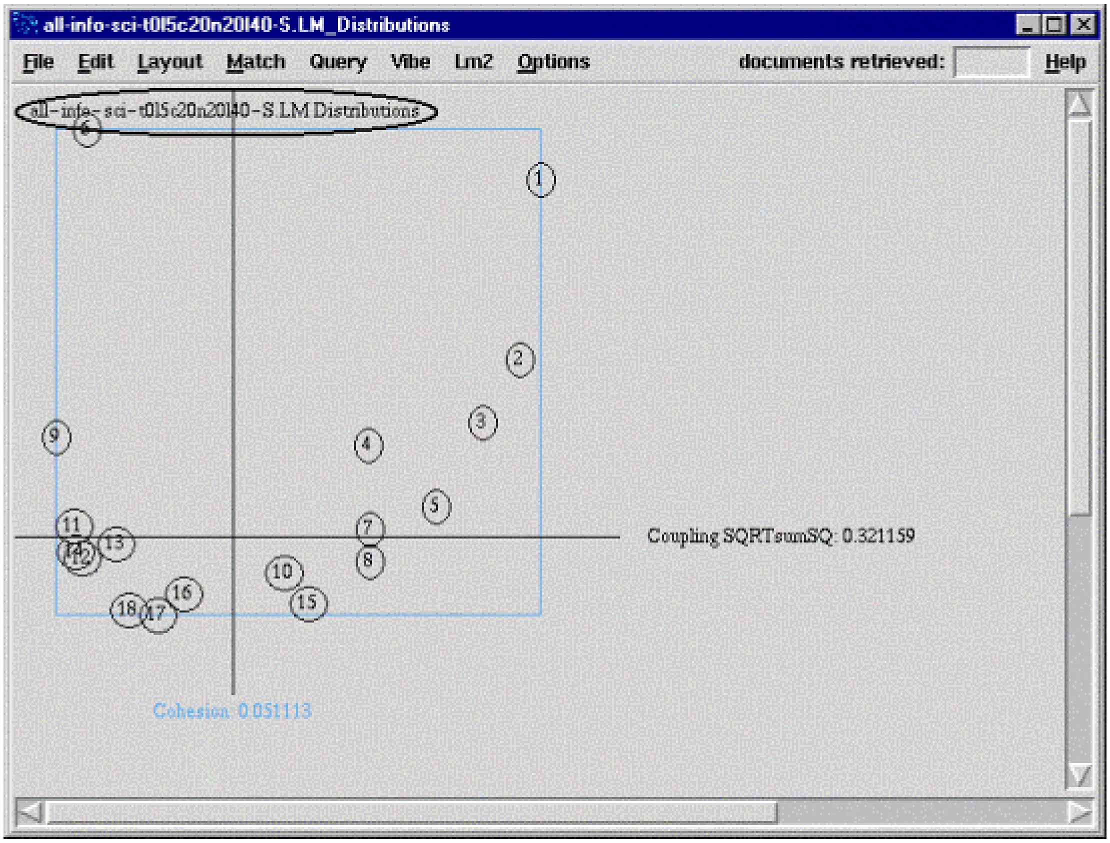


Figure 4: Information System Leximap Distributions

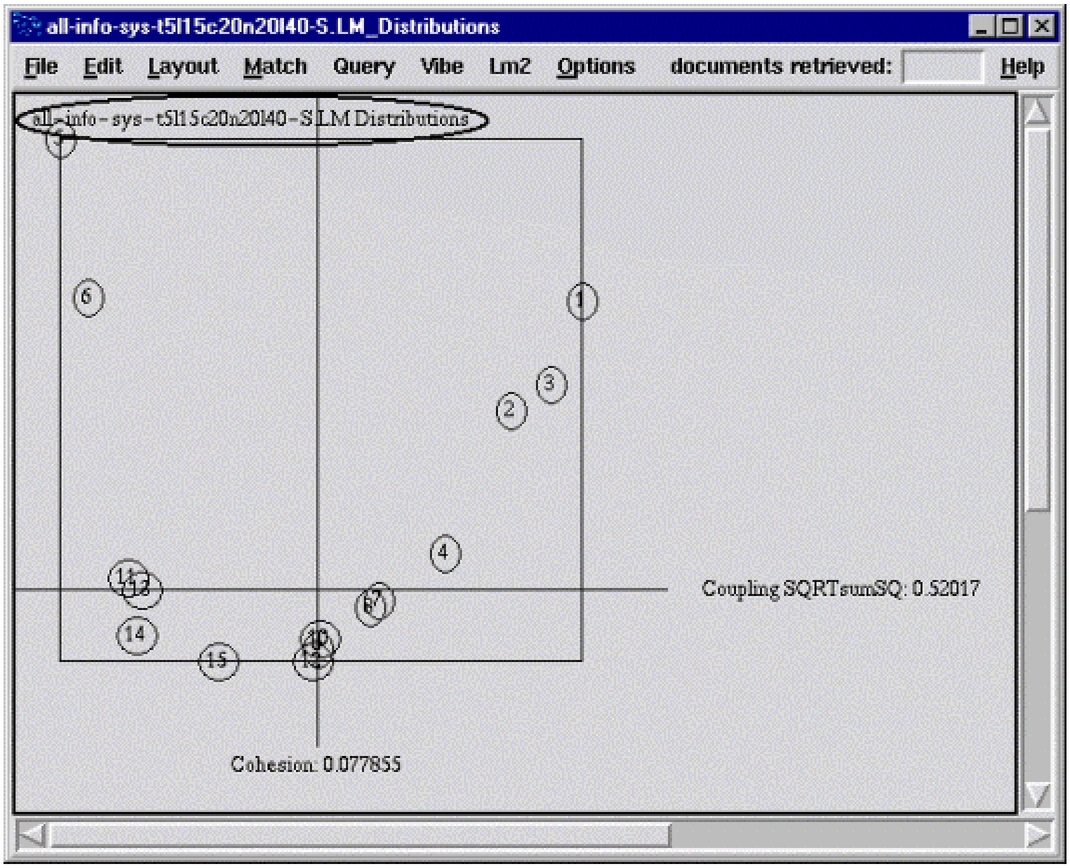
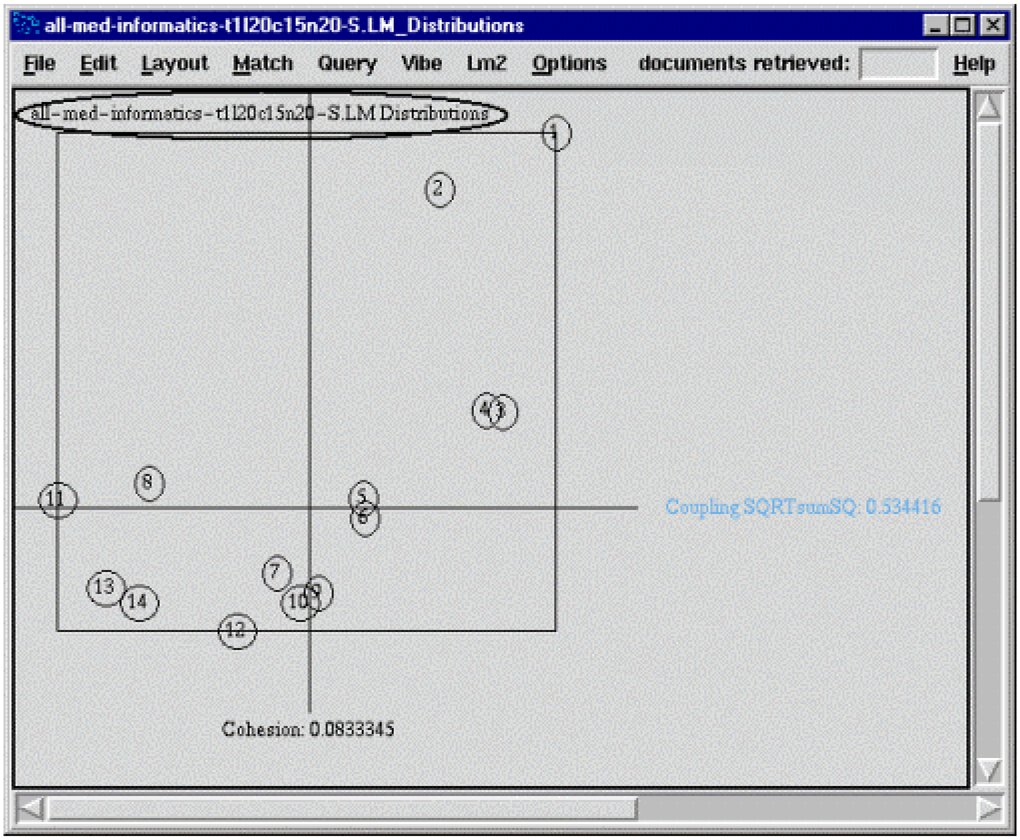


Figure 5: Medical Informatics Leximap Distributions



Perhaps the most interesting leximaps are the ones with both strong cohesion and strong coupling. Each of the three sets of leximaps has several examples of these indicating that Information Science, Information Systems and Medical Informatics all have definite focal points that are useful in organizing information in these fields. Information Science has six maps exhibiting both relatively high cohesion and coupling with map 1 in this set clearly showing the way for the other maps in the set. Map 6 of the Information Science set, though highest I cohesion is also one of the lowest in coupling.

Information systems has four such maps with its map 1 but also maps 3 and 2 charting the course of the other maps. Maps 5 and 6 in this set are highest in coherence, but also lowest in coupling.

Finally Medical Informatics has five of these maps with maps 1 and 2 separating themselves from the rest in both cohesion and coupling, which is quite unique.

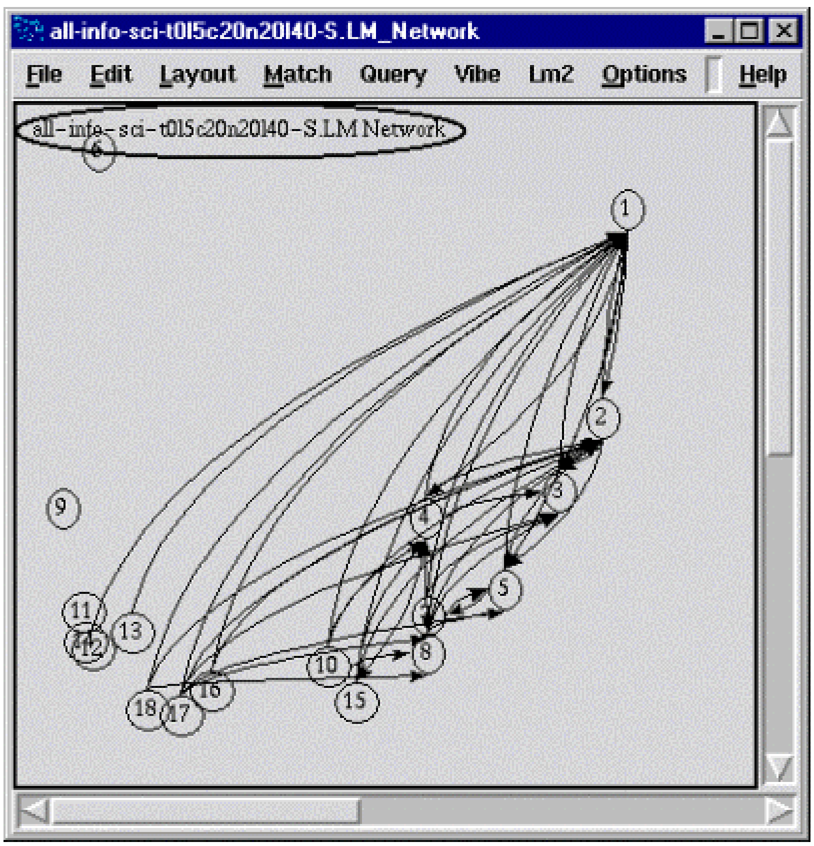
In addition to describing how leximaps compare within a set, we can be more specific in describing how they interact with other specific maps; this addresses coupling in a more focused fashion, but does not substitute for the general

coupling measure. Coupling is a composite measure of a leximaps' s intersection with all other leximaps generated in the same set; networks of leximaps describe pair-wise links of maps in a given set.
More formally, principal and secondary maps are defined as follows: If Map-A has internal nodes that are Pass-2 (thin-lined) nodes linked (by thin lines) in Map-B, then Map-A is a secondary map of Map-B. Networks of maps contain directed links to capture this situation. They are of the form:

Network-A →Network-B

An example of such a network of maps is shown in Figure 6. It is the Information Science network of maps.

Figure 6: Information Science Network of Maps



10. Applying Co-Word Analysis to the Journal Literature of ISci, IS and MI

The co-word metrics and algorithms were applied to noun and verb phrases extracted automatically from the titles and abstracts of representative journal articles from 1990-2000. It should be noted at the outset of this part of the discussion that the interpretation being offered is just one route through the maps. Once the reader has taken the trip, he or she will be able to take other trips on their own.

The highest ranking leximap in each of the three sets (ISci, IS and MI) was Leximap 1 (see Figures 1, 7 and 8). Each of these maps shared one very important point in common. All had information and system linked and these linked terms determined the primary theme of these maps. In ISci Leximap (LM) 1, information and system were joined by retrieval as part of the primary theme of the map. This was not surprising given the discussions of the ISci literature reviewed earlier in the paper. Also not surprising were the sub-themes linked to the primary theme in the ISci LM 1 such as term, query, document, user and search. The sub-themes of IS LM 1 were quite different from ISci LM 1 emphasizing management, organization and process along with two distinct but similar kinds of information technologies - decision support systems and executive information systems. MI LM 1 also introduces specific kinds of information systems, patient information systems and health care information systems. Its important sub-themes are knowledge representation and language semantics, access to library and information services and patient data and electronic records.

Figure 7: Information Systems Leximap 1

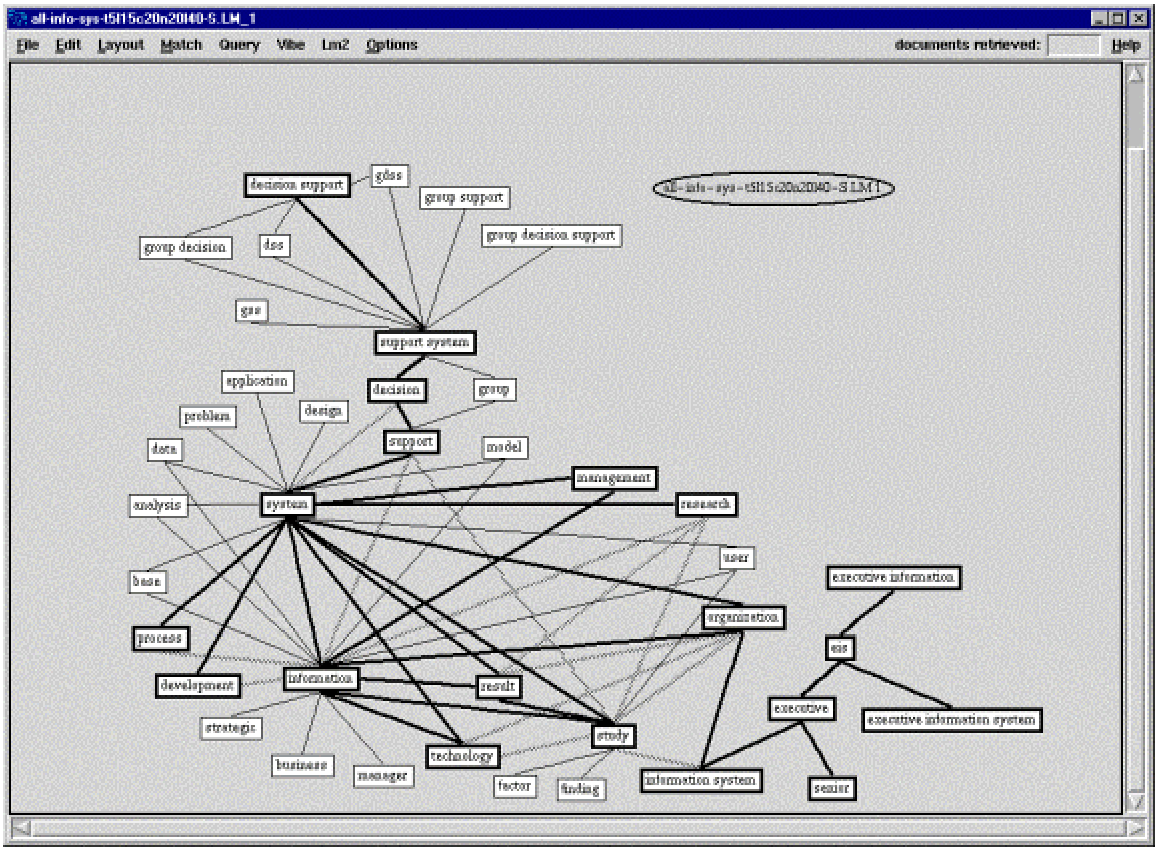
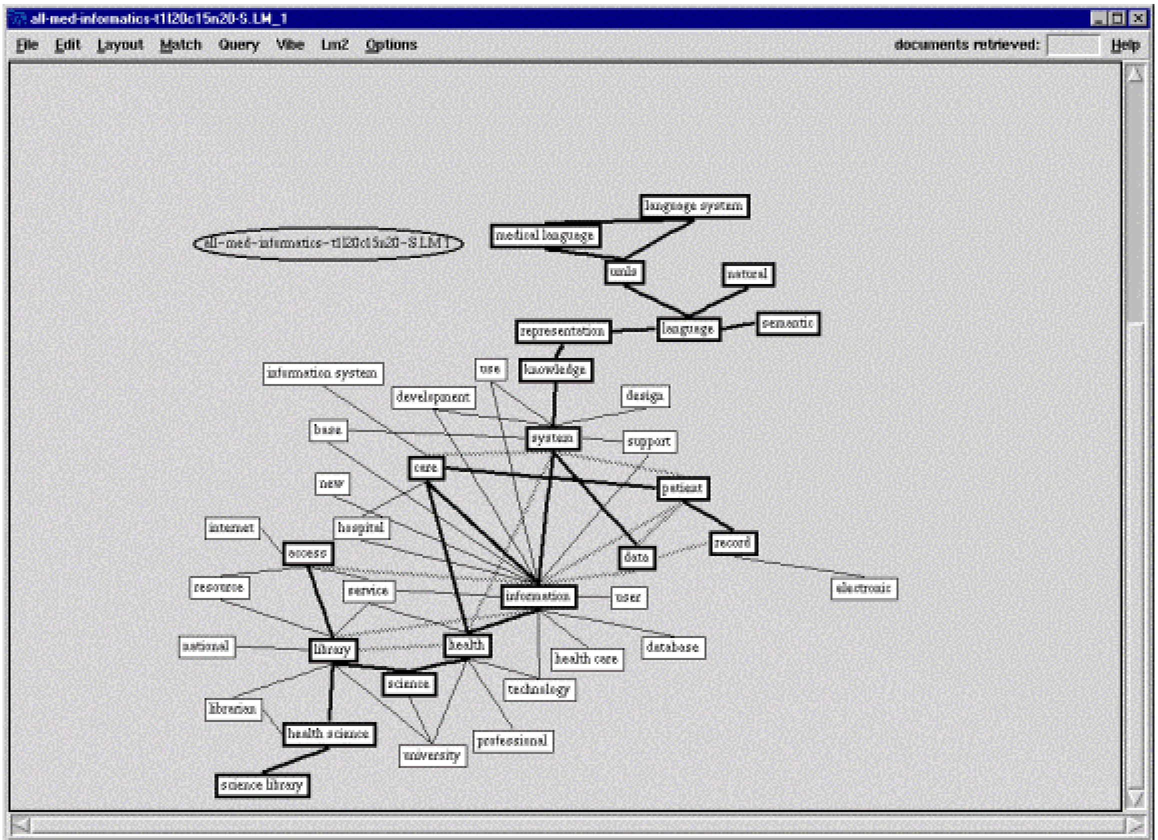


Figure 8: Medical Informatics Leximap 1



To a certain extent these three maps bear out the story that concluded the Ellis study. ISci is primarily concerned with the information content of systems and retrieval, and IS is concerned with the organizational context of systems and computer based decision support and executive information systems. However, in contradistinction to the Ellis study, MI LM 1 shows how these so-called disjoint elements can be combined in one subject matter. MI LM 1 shows an awareness of information content in its knowledge representation sub-theme and an awareness of organizational context in its health care, hospital and patient sub-themes. There is also indication of both an interest in health information technology and also access to library services.

In concluding this paper, interpretation of the cross-over themes of libraries, decision support technologies, knowledge and expert systems is carried several steps further.

11. Combining Themes of Information Science and Information Systems in Medical Informatics

The co-word analysis performed for this study does provide prima facie evidence that MI is indeed a hybrid of ISci and IS. More work will have to be done to establish that it actually is a hybrid, including co-citation analysis. But the co-citation analysis should be guided by co-word analysis or some kind of rigorous content analysis. Identifying and characterizing links among disciplines that underlie productive interdisciplinary or cross-disciplinary work or that determine a discipline to be in significant part the result of combining important building blocks from multiple disciplines requires that the subject matter of each of the disciplines is characterized at the necessary levels of granularity and scope. The section above took a first step in applying the kind of method and teasing out the kind of content at the level of detail and scope needed to establish MI as a hybrid of ISci and IS. The paper concludes with a discussion of three other terminological structures important in determining the kinds of subject matter whose connectivity would have to be established.

11.1 Knowledge in ISci, IS and MI

Figure 9 shows ISci Leximap 8. It shows, among other things the terminological structure in which knowledge is included. Not only are relationships to domain knowledge and expert systems shown, but also an interest in knowledge structure and models of knowledge including conceptual models. Information Systems Leximap 8 in Figure 10 also shows relationships between knowledge and domains and expert systems but focuses on acquisition rather than conceptual models and structures of knowledge. MI Leximaps 1, 3, 5 and 7 shown in Figures 8, 11, 12 and 13 respectively, shows relationships between knowledge on the one hand and representations, models and structures on the other. Some of latter are also related to decision support in diagnosis and expert systems. Clearly MI combines ISci and IS interests with respect to knowledge.

Figure 9: Information Science Leximap 8

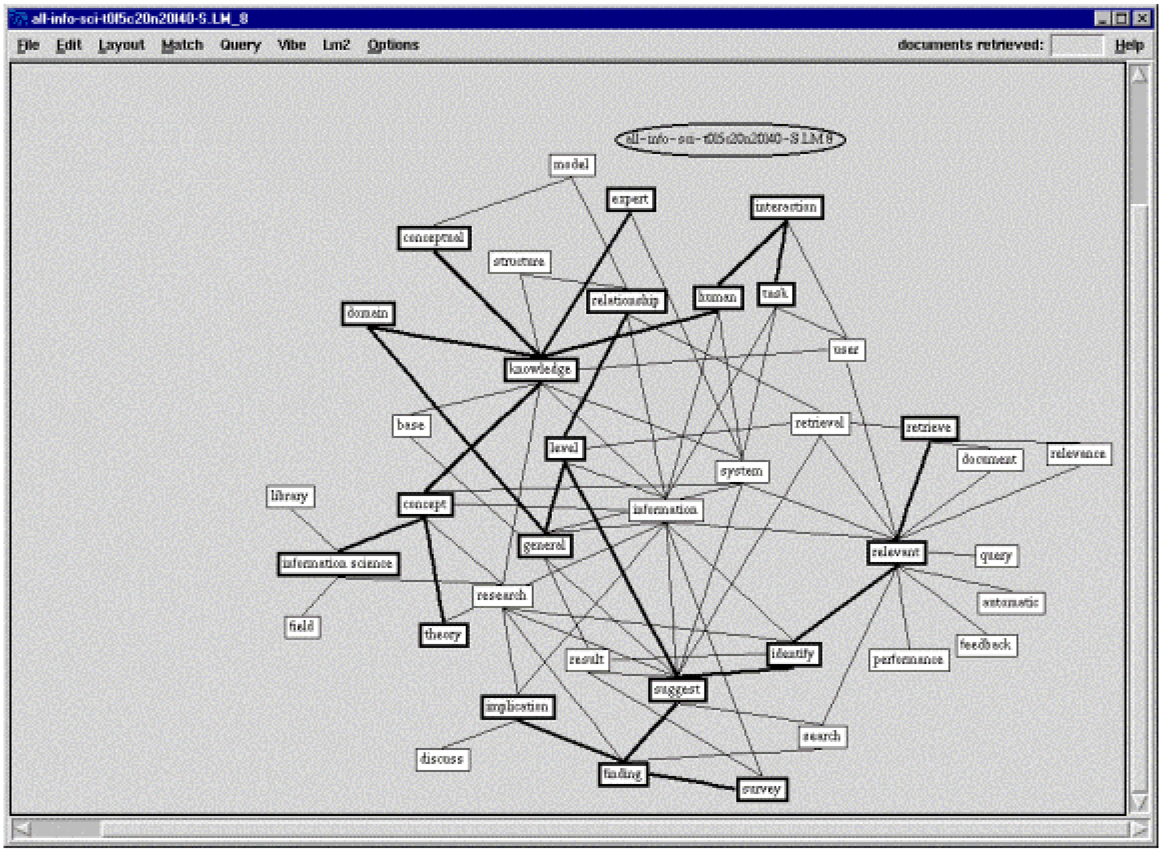


Figure 10: Information Systems Leximap 8

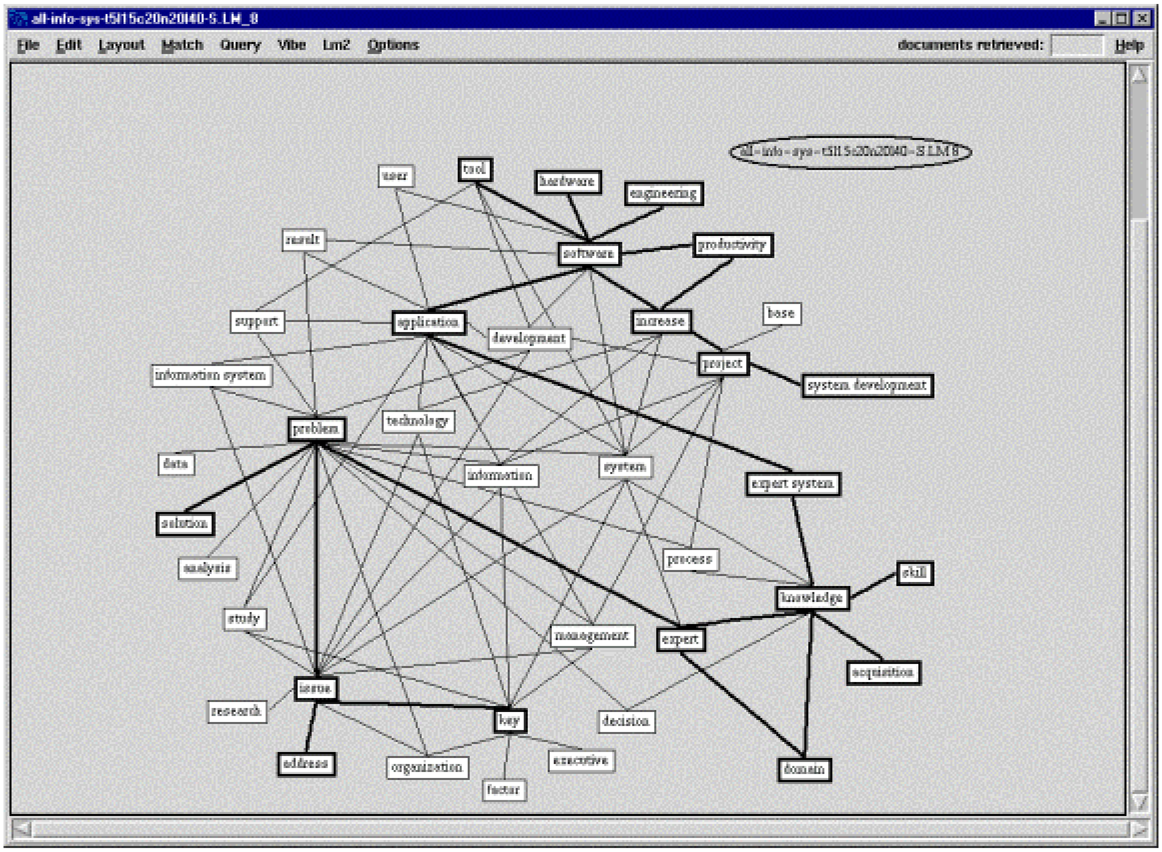


Figure 11: Medical Informatics Leximap 3

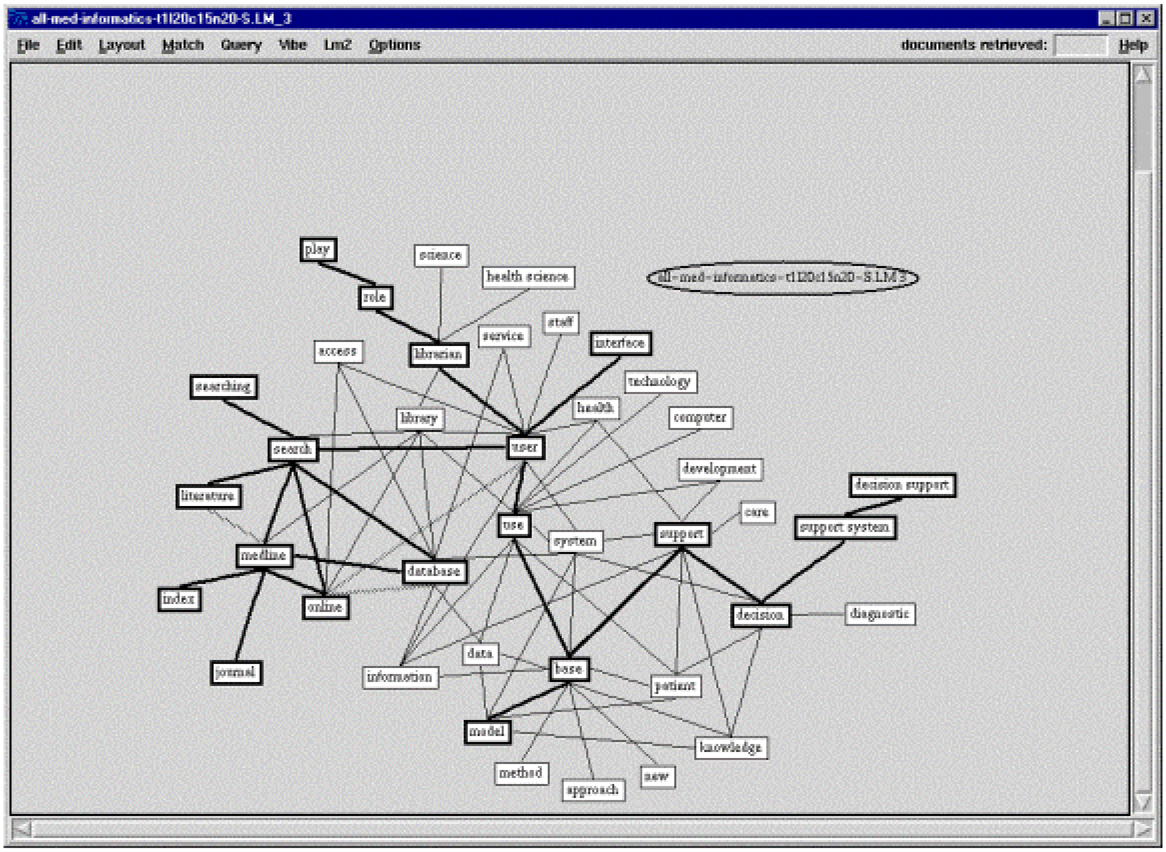


Figure 12: Medical Informatics Leximap 5

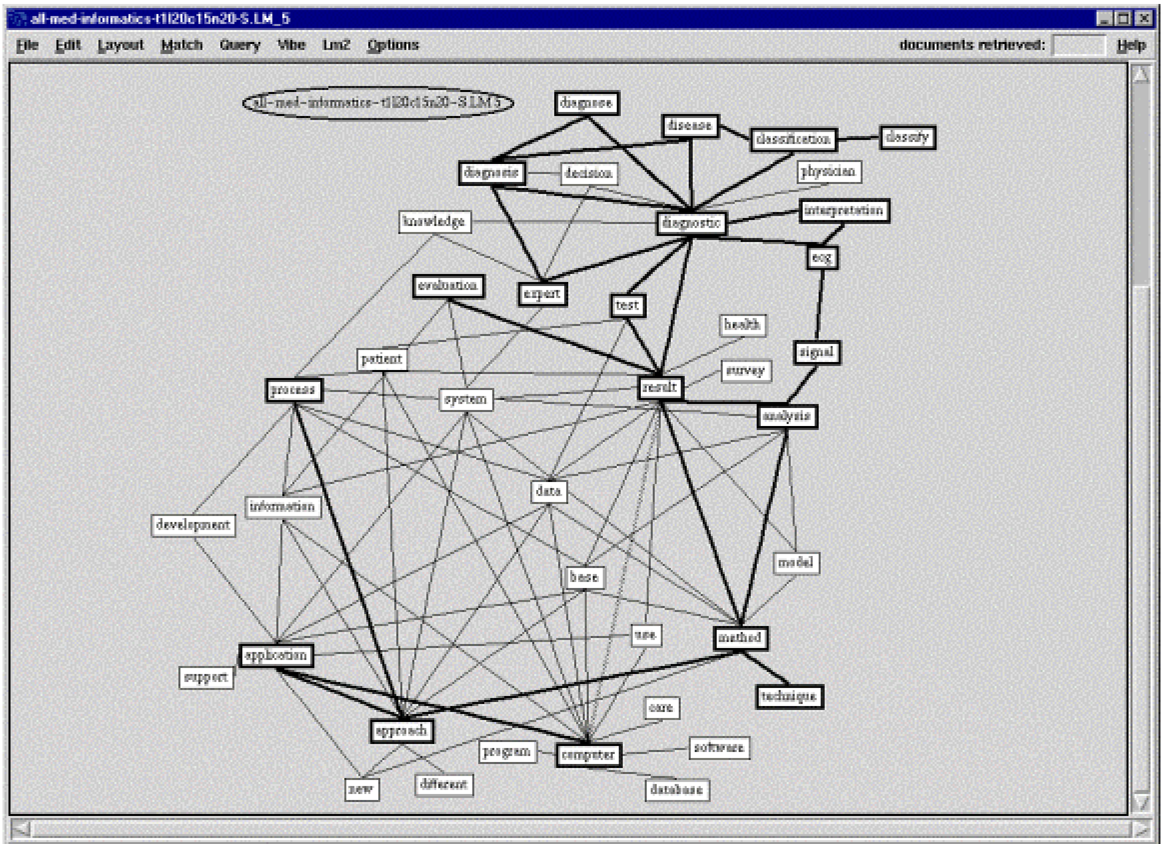
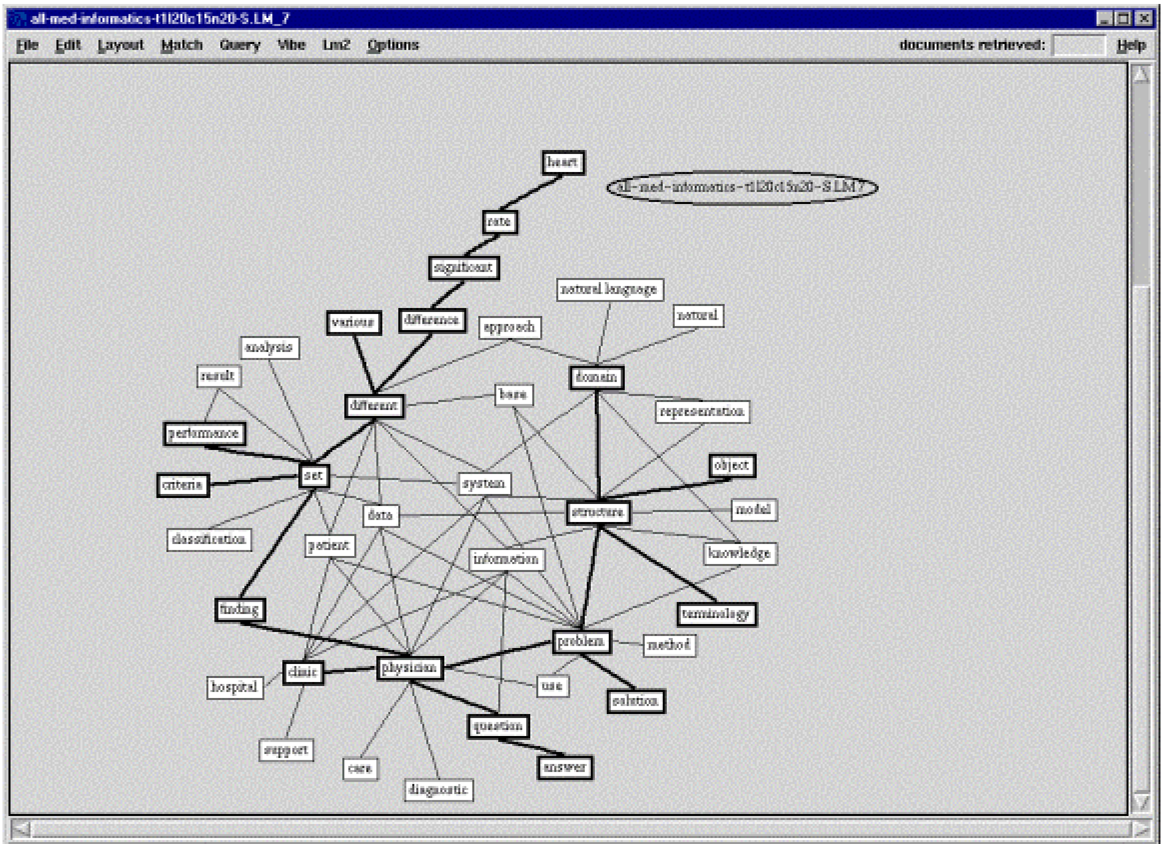


Figure 13: Medical Informatics Leximap 7



11.2 Decision Support Systems in ISci, IS and MI

IS Leximap 1 in Figure 8 shows a very significant relationship between **information systems** and **group/decision support systems**. This is further elaborated in IS Leximap 3 in Figure 14 where the significance of **groups** and **group work** and other **group** activities in IS is shown. There is some relationship shown in ISci Leximap 5, Figure 15, between **information systems** and **decision support**, but not nearly as specified in terms of types of systems as in the IS maps. With respect to **decision support systems**, MI is more similar to IS than ISci. This is shown Leximaps 3 and 5 in Figures 11 and 12 respectively. In MI, **decision support** systems are put in service of **diagnosis** and implemented in **expert systems**. The interest in **decision support** is as differentiated as that shown for IS and also makes the connection between **decision support** and **expert systems** as does IS. So MI does appear to side more strongly with IS in this case. But if MI sides with IS in decision support, it sides even more strongly with ISci with respect to libraries.

Figure 14: Information Systems Leximap 3

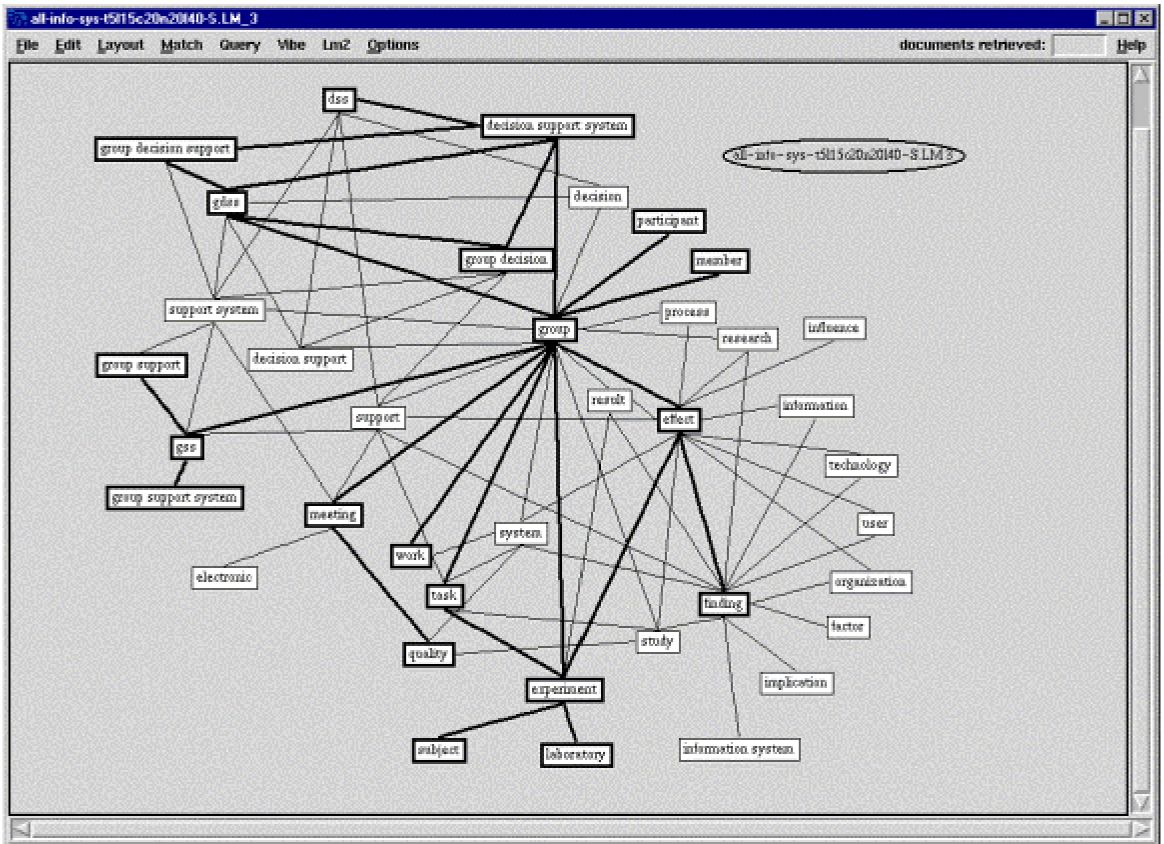
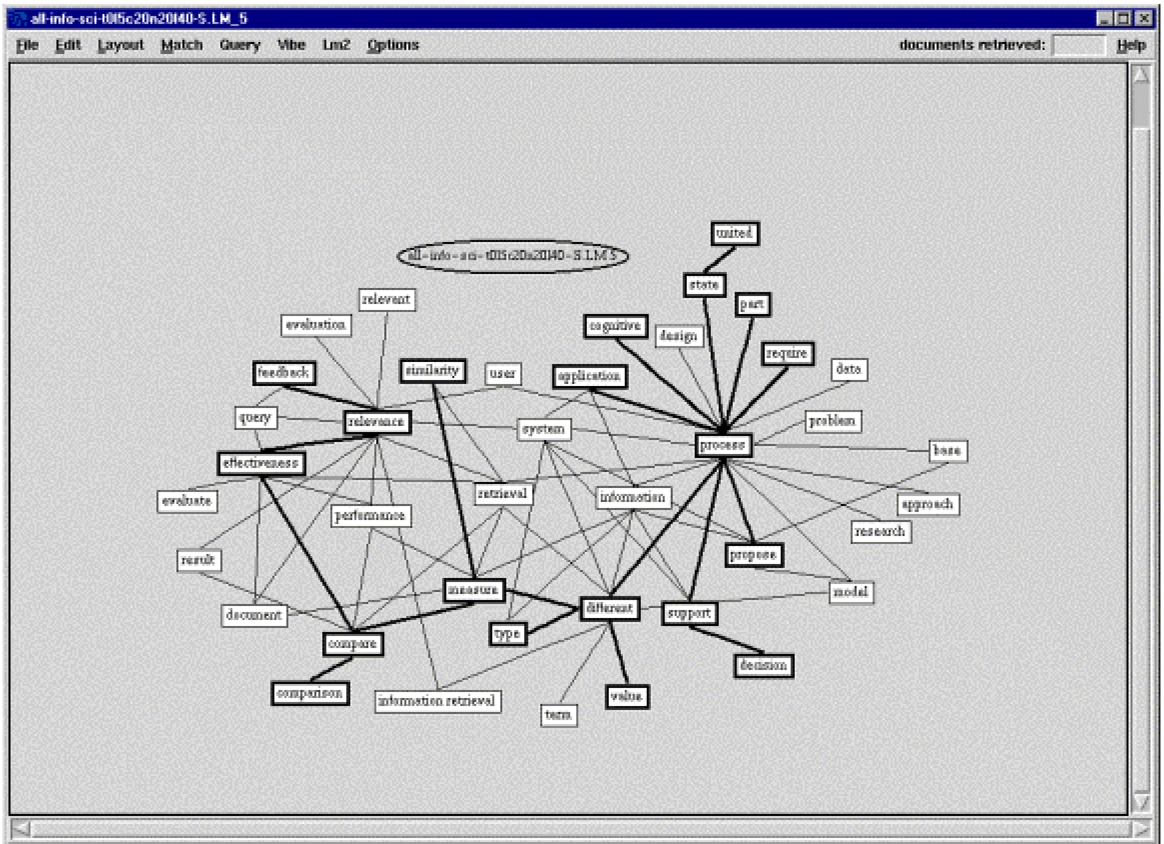


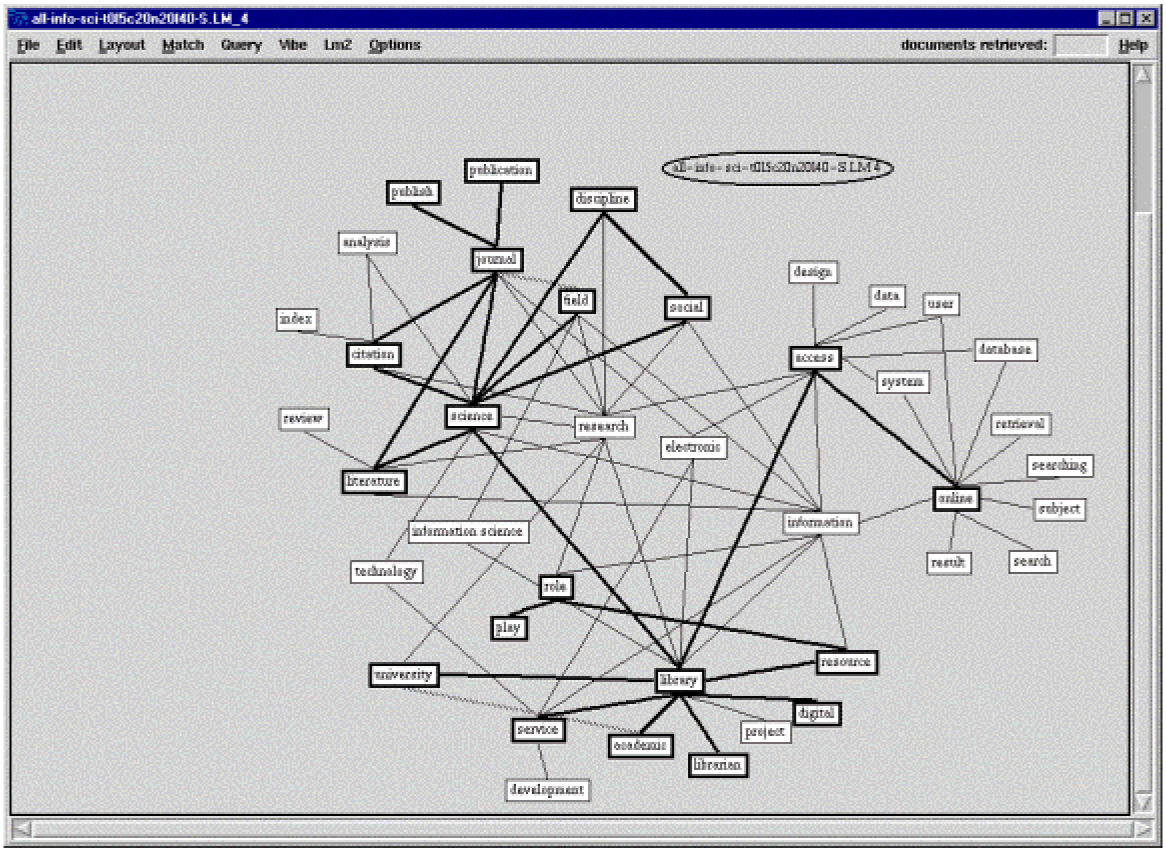
Figure 15: Information Science Leximap 5



11.3 **Libraries** in ISci, IS and MI

Libraries are obviously a central interest in ISci and this is shown in the ISci Leximap contained in Figure 16. The relationships of the **library** to the **university** and **science research** and for **academic service** are shown. Just as important is the emphasis on **library access** especially in **electronic** and **digital** forms and **online**. Almost the same interests are shown in MI Leximap 1, Figure 8. There is the relationships to the **university** and **science** (though in this case specifically to **health science**). There is also a concern with **library access** over the **internet**. Significantly there are no IS leximaps created based on this data. A typical view found in the few abstracts concerning libraries in the IS literature put forth the claim that the library is an inappropriate model for business. It would seem that this is a place for some cross-over work to be done.

Figure 16: Information Science Leximap 4



11.4 Conclusion

The paper only begins to tap the information in the leximaps created from the corpora of journal article titles and abstracts in ISci, IS and MI. Other areas of overlap and difference that are somewhat more intricate and too time-consuming to present have been identified. These include the respective fields' interests in **users**, **data**, databases and **methods** of **research** and **development**. Nevertheless, there is good reason to believe, even based on the sample analyses presented in the paper, that IM's interests and subject matter are a hybrid of ISci and IS. However, to show that IM is an actual hybrid would require not only more extensive co-word analysis, but also extensive co-citation analysis as well. It might also require making the co-occurrence constraint for the co-word analysis more lenient allowing in for analysis more terminological levels of detail. It would also be interesting to see how each of the fields developed by performing co-word analysis on the literature for each field in different time periods.

This paper has attempted to provide a more viable approach to studying the subject matter of the respective fields so that their actual and potential links and common interests can be surfaced and new kinds of significant inquiries can be constructed.

Notes

[1] The Institute for Scientific Information maintains lists of the most cited journal in both Science and Engineering and the Social Sciences.

[2] See the study by Nisonger (1999) on JASIS' ranking with respect to other library and information science journals.

[3] One MI journal was excluded because of its small number of entries due to its recent beginning.

[4] An added reason for including Information Systems in the next round of analyses is that it was ranked as more representative of IS in the Ellis study than some of the other journals that were more similar to MISQ.

[5] These were originally called Leximappes (Turner, Charton & Laville, 1988).

[6] The Software Engineering Institute is a federally funded research and development center sponsored by the U.S. Department of Defense and operated by Carnegie Mellon University.

[7]INQUERY is licensed from Applied Computing Systems Institute of Massachusetts, Inc. (ACSIOM).

[8] Terms in bold indicate they are contained in nodes on leximaps.

References

Callon, M., Law, J., & Rip, A. 1986. *Mapping of the dynamics of science and technology*. London: McMillian.

Callon, M., Courtial, J-P. & Laville, F. 1991. Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. *Scientometrics* 22, 1, 153-203.

Callon, M., Laredo, P. and Mustar, P. 1997. *The Strategic Management of Research and Technology: Evaluation of Programmes*, Paris, France: Economica International.

Checkland, P. & Holwel, S. 1998. Information systems and information systems: Making sense of the field. Chichester, England: John Wiley & Sons.

Coulter, N., Monarch, I., Konda, S., & Carr, M. 1996. An evolutionary perspective of software engineering re-search through co-word analysis. *CMU/SEI-96-TR-019*, Software Engineering Institute, Carnegie Mellon University, Pittsburgh.

Coulter, N., & Monarch, I. 1996. Best practices: What software engineering can learn from manufacturing engineering, The SEI Software Engineering Symposium, Pittsburgh, (slides published in conference proceedings)

Coulter, N., Monarch, I., Konda, S. 1998. Software Engineering as Seen through Its Research Literature: A Study in Co-Word Analysis. *JASIS* November 99, 49(13), 1206-1223.

Courtial, J-P., & Law, J. 1989. A co-word study of artificial intelligence. *Social Studies in Science* 19, 301-311, London: Sage.

Courtial, J-P. 1994. A co-word analysis of scientometrics. *Scientometrics* 31, 3 251-260.

Ellis, D., Allen, D. and Wilson, T. 1999. Information Science and Information Systems: Conjunct Subjects Disjunct Disciplines. *JASIS* 50(12):1095-1107.

Hjorland, B. 1997. *Information seeking and subject representation: An activity-theoretical approach to information science*. Westport, CT: Greenwood Press.

Institute of Information Scientists. 1996. *Criteria for information science*. London, England: Institute for Information Scientists.

Law, J., & Whittaker, J. 1992. Mapping acidification research: A test of the co-word method, *Scientometrics* 23, 3, 417-461.

Monarch, I., & Nirenburg, S. 1993. An ontological approach to knowledge acquisition schemes. In N. Bourbakis (Ed.) *Knowledge Engineering Shells: Systems and Techniques*. 29-56, New Jersey: World Scientific.

Monarch, I (1994). An interactive computational approach for building a software risk taxonomy. The Third SEI Conference on Software Risk, Pittsburgh, PA.

Monarch, I., & Gluch, D. 1995. An experiment in software development risk analysis, Software Engineering Technical Report CMU/SEI-95-TR-014, Software Engineering Institute of Carnegie Mellon University, Pittsburgh, PA.

Monarch, I. 1996. Software engineering risk repository. 1996 SEPG Conference, presentation, Atlantic City, New Jersey.

Monarch, I., Konda, S., Levy, S., Reich, Y., Subrahmanian, E., & Ulrich, C. 1997. Mapping sociotechnical networks in the making. In Bower, J., Gasser, L., Star, L., & Turner, W. (Eds.), *Social Science, Technical Systems, and Interactive Technology*. 331-354, Mahwah, New Jersey: Lawrence.

Nisonger, T. 1999. *JASIS* and Library and Information Science Journal Rankings: A Review and Analysis of the Last Half-Century. *Journal of the American Society for Information Science*, 50, 1004-1019.

Small, H. 1973. Co-citation in the scientific literature: A new measure of the relationship between documents. *Journal of the American Society for Information Science*, 24, 265-269.

Small, H. & Griffith, B. 1974. The structure of scientific literature I: Identifying and graphing specialties. *Science Studies* 4, 17-40.

Smith, L. 1999. *Journal of the American Society for Information Science (JASIS)*: Past, Present and Future. *Journal of the American Society for Information Science*, 50, 965-969.

Turner, W., Chartron, G., Laville, F., & Michelet, B. 1988. Packaging information for peer review: new co-word analysis techniques. In Van Raan, A., (Ed), *Handbook of quantitative studies of science and technology*. 291-323, Amsterdam: North Holland.

UKAIS. 1997. Newsletter of the UKAIS 3. Available at http://www.scs.leeds/ac/ukais/Newsletters/vol3no4.html. Accessed August 24, 1998.

Whittaker, J. 1989. Creativity and conformity in science: Titles, keywords, and co-word analysis. *Social Science in Science*, 19, 473-496.