PERGAMON

Information Processing and Management 39 (2003) 639–664

INFORMATION PROCESSING & MANAGEMENT

www.elsevier.com/locate/infoproman

Information landscaping: information mapping, charting, querying and reporting techniques for total quality knowledge management

Bor-sheng Tsai *

School of Information and Library Science, Pratt Institute, 144 West 14th Street, New York, NY 10011, USA

Received 9 April 2001; accepted 25 October 2002

Abstract

Information landscaping—an integration of information mapping, charting, querying and reporting techniques—has been developed to enable the construction of a total quality knowledge management system focusing on a particular subject information field. The techniques apply five major parameters of the Fuzzy commonality model (FCM) including unionization, quantity, continuity or stability, changeability, and critical probability, to construct a series of information maps (infomaps) and a set of chronological-statistical charts (infocharts). The infomaps and infocharts are used as the blueprints and navigation agents for building and developing a web-based subject experts depository and query–report system. Focusing on the subject experts/expertise, this system enables a researcher to expedite a query search through infomaps (qualitative reference) and infocharts (quantitative reference). The entropy measurement and the entropy constant (the square root of the average entropy measure) are calculated to compare with the critical probability of the FCM. This leads to the finding of a set of regression straight lines and the establishment of an information oscillogram. The tropics (upper limit, middle range, lower limit), and the potential/solstitial population and its growth rate within a subject information domain during a particular time period can be determined. They can effectively and efficiently guide librarians and information professionals towards the construction and the continuous development of an electronic collection. The cultivation of a virtual learning and referencing environment can also be created by utilizing this data.

Keywords: Bibliometrics and informetrics; Citation mining; Information landscaping; Information mapping; Knowledge management

* Tel.: +1-212-647-7297; fax: +1-212-367-2492.
E-mail address: btsai@pratt.edu (B. Tsai).

0306-4573/03/$ - see front matter
doi:10.1016/S0306-4573(02)00098-5

1. Research goals and objectives

This research merges total quality management (TQM) and knowledge management (KM) into total quality knowledge management (TOKM) and uses it as a conceptual model to direct and develop information landscaping techniques through the coordination of information mapping, charting, querying, and reporting. TQM has been a major force that has influenced business operations and organizational management since the 1970s. KM, on the other hand, has been consistently playing a key role in many information business platforms and has continued to gain popularity and momentum in recent years. It is logical that it would better serve information researchers if the above two key managerial concepts and techniques were merged. The word "total," particularly deserves attention and reemphasis. It indicates the importance of providing or foredrawing the big picture (i.e., a perspective total information landscape) in order to guide information researchers in their searches. The total big picture (namely, the information landscape) may be represented by a set of information maps (infomaps) and information charts (infocharts). The word "quality" concentrates on selected components, such as highly selected authors and their publications or information products (e.g., infomaps and infocharts, and infomap/chart-coordinated query–reports) that have undergone a filtering process (e.g., citation analysis and data mining-based information landscaping). More specifically, this research intends to explore and recognize important scholars in the study of cognitive information processing and coordination (hereafter, cognitive coordination). As an alternative information retrieval approach, cognitive coordination helps both information system managers and users alike in filtering quality information resources and reducing the tensions and stresses caused by information overload. The construction and the development of a total citation-based *information mapping–charting system* and a web-based *query/report-centered* database management *system* become essential and possible. With minimal coordination, these two systems can cultivate an expandable virtual learning and referencing environment.

The goals of this research are to: (1) integrate TQM and KM; (2) merge citation analysis (CA) and data mining (DM); and (3) apply data visualization (DV) and information architecture (IA) in drafting a series of directed citation maps (infomaps) and chronological-statistical charts (infocharts) as blueprints and navigation agents. These blueprints and navigational tools will be used to construct, navigate, and assess a web-based subject experts depository that is equipped with query–report capabilities.

The objectives of this research are to develop information landscaping techniques for constructing an information retrieval system that enables subject-focused researchers to instantly and effectively: (1) visualize a set of 2-D and 3-D infocharts that show a chronological flow of a special subject information field; (2) review a series of electronic infomaps that display a group of leading researchers in a special subject information field; (3) apply the structured query language (SQL)—with the aids of infomaps and infocharts—in retrieving quality subject information, storing useful query instructions, and obtaining an integrated query–report; and (4) continue to monitor the development of this web-based subject collection, as well as to assess the overall structural changes of the citation networks by utilizing the key Fuzzy commonality model (FCM) parameter P_c, the entropy measurement H_c and entropy constant E_c, and the linear regression approach (Y_L).

2. Research questions

Based on the above research goals and objectives, this study focuses on the following questions:

(1) *How can a subject infomap series be efficiently and effectively developed?* In other words, how can the scholarly communication activities among some leading information researchers in the same field, and (coincidentally in this case) in the same institution, be quickly, easily, and accurately captured and represented on an electronic infomap series?

Before we can answer the first question, a related background-check question is: *Why is the development of such a subject infomap series important to librarians and information professionals?* In other words, why is it necessary to build a technical model that enables librarians and information professionals to construct such a subject infomap series? This background question needs immediate attention and an explanation as follows.

Explanation: Infomapping is a concept and technique that applies citation analysis and quantitative methods in monitoring and mapping scholarly communication activities in a particular subject information field. The activities cover mainly formal communications in official written publication forms such as journal publications, conference proceedings, etc. This approach has been one of the important and significant collective successes achieved by a sub-discipline of information science called bibliometrics, informetrics, and scientometrics (Brookes, 1990; Tague-Sutcliffe, 1992; Sengupta, 1992; Hood & Wilson, 2001). The researchers in this sub-discipline mainly apply citation analysis and quantitative methods in studying information properties, structures, representation, storage, and retrieval, as well as in studying information flows, transfer, distribution, and user's information seeking behaviors. Regardless of its successes, the infomapping approach needs a major promotion in the library and information professional environment. In previous experimentations, the author designed a machine readable mapping (MARM) system for the construction of a series of 2-D electronic infomaps (Tsai, 1992, 1993), also applied were, respectively, the hDC Desktop (Tsai, 1994a, 1994b) and the Java script (Tsai, 1995, 1997) for constructing various sets of animated infomap series. Additionally, the application of VRML-based 3-D infomap blocks capable of showing data city skyline and linking and coordinating these data blocks with HTML-based web documents and websites was employed (Tsai, 2000a, 2000b). To date, this approach and its operations continue to significantly assist the author's teaching. Still, there are a few drawbacks: (a) the completion of the site construction in each technical development stage has been time-consuming; (b) the continuously expanding infomaps require constant filtration and timely revision; (c) the system demands continuous technical support/maintenance and technological upgrade/advancement; and (d) the computer programming knowledge and techniques are not always readily available and/or easily applicable for a lay user (e.g., a librarian or a library client) (Kirschenbaum, 1999). Inasmuch as librarians and information professionals can use the infomapping technique to expand their information service domain, it is the author's intention to find a simple formula so that they can conduct such infomapping in a way that meets the following six principles: (i) The end-product must be very easy to use. (ii) The system must follow the top principle of lowest common device (LCD) accommodating global affordability and accessibility. (iii) It must require a minimum of computer skills and lab time to make an infomap program. (iv) It must require the least human effort in learning the task of infomap-making. (v) It must require a minimum of financial support/investment from the library or information organization's administration and management level. (vi) Constant and frequent maintenance and modification must be easily achieved. In a nutshell, this new technical scheme for infomapping must be one that allows many librarians and information professionals to do their best with ease and accuracy, taking into consideration that many of them have busy schedules, limited budgets, and moderate-levels of computer skills/equipment. As such, to make the whole idea work, this new approach to infomapping must be a convenient way that requires minimal effort and time in the making of the e-products. All they should concentrate on is the mapping of the interrelated queues that can guide researchers to the expert contents of the desired subject information resources.

(2) *How can the TQKM, the merger of CA and DM, and the combination of DV and IA, serve the purpose of developing the information landscaping techniques?* In other words, how can we follow the guidance of the TQKM and apply the citation data mining to record, measure and represent the continued scholarly communication relationships of leading researchers in a special subject information field? In addition, how can we at the same time reveal and represent the cumulative advantage (success breeds more success) (Price, 1976) and continual dynamics (Garfield, 1979) that are derived from such intercommunication (or, informational recurrence/overlapping) (Goffman & Warren, 1980) relationships—as one of the main focus points applying the FCM—in the making of a set of infocharts?

Explanation: A city map mainly provides directional guidance for a tourist. Similarly, an infomap provides an overall picture for a researcher to self-explore around the mapped information environment. Although both maps can provide useful source information, they do not lead an anxious person to the desirable destinations (e.g., major sites revealing the beauty of the city, or major files containing the wanted expertise). To compensate for this missing link, the information landscaping technique is developed. As a conceptual and technical model, the information landscaping process cannot be performed without help from TQKM (i.e., the integration of TQM and KM), CA–DM (citation data mining—the merger of CA and DM), and DV–IA (i.e., the combination of data visualization and information architecture). The use of TQKM, CA–DM, and DV–IA allows the infochart to be derived, synchronized, and used as a quantitative mirror for simultaneously reflecting its qualitative origin—the infomap. The continual dynamics can be revealed on an infochart by measuring the change of the amount of information that each infomap produced and evolved in an infomap series within a specific time period.

(3) *How can a web-based subject experts depository—equipped with infomap–infochart and query–report capabilities—be constructed?* In other words, how can the drafting of subject infomaps and the making of infocharts be used as blueprints for constructing a web-based subject information resource management system that provides infomap, infochart, experts depository, and query–report functions?

Explanation: An experts depository is an information warehouse that stores bibliographic citations and summaries of the most frequently cited publications of a group of targeted experts in a special subject information field. The depository can provide an instant information package that is the result of using the navigational directions provided by the infomaps and the infocharts, and by applying the integrated query–report function embedded in the depository.

(4) *How can the TQKM-guided infomaps and the FCM-based infocharts be effectively used as navigation agents for information querying and reporting?* In other words, how can new knowledge be discovered from the visual directions of the infomaps and the infocharts, and from the operations and the functions of the constructed subject experts depository that is equipped with query–report capabilities?

Explanation: There exists a missing link between the infomap–infochart's navigation function and experts directory's query–report function. As the customer–supplier connection is one of the most important attributes of the TQKM (Taylor, 1982; Laszlo, 1999) that can lead an anxious information searcher to the desirable destination (i.e., to learn subject experts' major intellectual contributions and achievements), it requires the system to mingle the infomap–infochart function with query–report functions.

(5) *How can the major FCM parameters, the entropy measures, and the linear regression approach help in continuously monitoring the structural change of citation networks and assessing the development of this web-based subject collection?* In other words, how can the FCM derived critical probability P_c, the entropy measurement H_c, the entropy constant E_c, and the regression straight line Y_L be properly interweaved and interpreted to illustrate the dynamics of various information elements (authors in this case) that flow and recur in the citation networks and in the web-based subject collection?

Explanation: The FCM is a data-mining instrument developed and devised by the author to detect the dynamics of a growing database collection, namely the experts depository. The critical probability P_c derived from the FCM can be used to indicate the youth of this database. Other detecting devices such as the entropy measurement H_c, the entropy constant E_c, and the regression straight line Y_L are also equipped in parallel to show and compare the expansion and the progression rates of the database. The FCM, entropy measurement, and the linear regression approach are detailed in Sections 3.2, 3.3 and 7.

In sum, it is noticeable that the first three research questions target more on the *system development* and the last two questions target more on the *system assessment*. Therefore, the five research questions must be tackled in a "two target" sequence. As a TQKM system requires continuous quantity control and quality improvement, the synchronization of both qualitative referencing (infomapping) and quantitative referencing (infocharting) is significant and useful.

3. Research methodology

3.1. Concept groups

This research applies the following six concept groups:

(1) *TQM* (total quality management), a popular guideline for organizational management, is adopted for developing strategic infomaps and infocharts for an information organization (St Clair, 1997; Cronin, 2000; Gregory, 2000). Both infomaps and infocharts empower a researcher to instantly visualize a total scene and focus on a well-landscaped portion of a subject information field. The TQM has many features that can guide the management of this research project. These features include customer-oriented focus (Koenig & Srikantaiah, 2000); reduction of variation and R&D cycle time (Chua, 2001); value-added approach, customer–supplier connection, commitment from the top, organizational culture, and quality of management (Taylor, 1982; Laszlo, 1999); customer management, and continuous improvement of the quality of systems consisting of social system, technical system and management system (Bryan, 1996); employee participation and change intervention (Coyle-Shapiro, 1999); teaching–learning assessment (Killingsworth,

Harden, & Dellana, 1999; Vazzana, Elfrink, & Duane, 2000); and critical factor analysis, and constructs correlation (Ravichandran & Rai, 2000). The application of TQM in libraries (e.g., Harvard College Library and Oregon State University Libraries) prompted an inquiry into Deming's 14 steps or principles (Bryan, 1996) to TQM, particularly in the areas of continual quality and productivity improvement, continuing education, staff training and retraining, and leadership (Masters, 1996; Aldakhilallah & Parente, 2002); and TQM in KM (Johannsen, 2000).

(2) *KM* (knowledge management), a rising star in the information business world, provides important guidelines for the refinement of information resource management. To enhance this project, we focused on some major KM themes such as knowledge economy and information ecology (Davenport & Prusak, 1997, 1998); intellectual capital (Brooking, 1999; Kanter, 1999); knowing organization (Choo, 1996); heuristic knowledge (Tuthill, 1990); supporting software, coordinated interactive learning and reading aid, and experts summary (Haur, 1999; Levine, 2001); directory of expertise (Stromquist & Samoff, 2000); knowledge mapping, OLAP (online analytical processing) chart and graph, and intelligent agent (Stewart, 1995; Kanter, 1999; Vail III, 1999); architecture of organizational memory, knowledge repository, and repackaging and measuring (Hackbarth & Grover, 1999; Morrow, 2001; Townley, 2001); and information resource management (Van Den Hoven, 2001; Holsapple & Joshi, 2002).

(3) *CA* (citation analysis), a legitimate metric instrument for describing/revealing the activity of a scholarly communication network, is applied to enable the visualization and identification of a variety of significant contributors within a subject information field (White & McCain, 1998; Small, 1999). For this project, the focused areas in CA include: information communication studies (Griffith, 1989); digitized maps and tabular files (Davies, 1990); author citation mapping (Vickery & Vickery, 1987; McCain, 1990; White & McCain, 1998; Small, 1999); and mapping of science (Price, 1965, 1979; Small & Garfield, 1985).

(4) *DM* (data mining), a quantitative tool for knowledge management, is used to discover and represent knowledge from a large volume of data (Fayyad, Piatetsky-Shapiro, & Smyth, 1996; Trybula, 1997; Vickery, 1997; Jurisica, 2000). This project focused on DM principles, methods, and algorithm (Weir, 1998; Chou & Chou, 1999; Guenther, 2000; Chen, Sakaguchi, & Frolick, 2000; Chopoorian et al., 2001); DM and KDD (knowledge discovery in databases) (Brodley & Lane, 1999); DM and customer relationships (McCarthy, 2000; Thearling, 2000); DM and visualization (Jern, 1998; Wong, 1999; Thearling et al., 2001); and DM for libraries (Banerjee, 1998; Guenther, 2000).

(5) *DV* (data visualization), a base for constructing infomaps and infocharts, has received more attention and has been gaining in momentum steadily since the mid-1980s. With the advent of the Internet, particularly with the creation of the web environment, the applications of DV started to grow impressively. As more navigational and interface programs are needed for web searching, DV undoubtedly is becoming very instrumental. This research project mainly focuses on the making of infomaps and infocharts and the linking of them with the experts depository. The main concentration on DV includes: geographic visualization and cognitive mapping (Chen, 1999); multiple query results and theme changes (Havre, Hetzler, Whitney, & Nowell, 2002); data signatures and association rules (Wong, 1999, 2000); and visual map display (Tufte, 1983; Lin, 1997).

(6) *IA* (information architecture), a new way of developing information design and landscaping concepts and techniques for constructing mostly web-based information resource management systems, is applied to enable the construction of a web-based experts depository that is capable of

providing infomapping, infocharting and query–report functions (Nicotera, 1999; Tsai, 2000b; Zwies, 2000; Gallupe, 2001). Some IA concepts and techniques used for the depository construction include: knowledge portal (Adams, 2001); database-driven website design (Roberts, 2000); electronic core collection development (Black, 2001; Lord & Ragon, 2001); information design for distance learning (O'Neill, 2001; Tsai, 2001); global virtual library (O'Leary, 2002); DM, KDD, IT and IA (Periasamy & Feeny, 1997; Thompkins, 1998).

The above six concept groups are closely interrelated. The integration of them is critical for the construction of a networked information resource management system such as the targeted web-based experts depository (Guenther, 2000; Myburgh, 2000; Smith, 2000).

To answer the proposed five research questions presented earlier, a special subject information field focusing on *cognitive coordination* is detected. The FCM is developed to detect, capture and outline the images of significant elements (authors in this case) in this subject field. A series of infomaps is drafted. A conceptual model derived from the FCM describes and explains the formation of this set of infomaps. A data model derived from the FCM devises a software program for: (1) citation data mining and information landscaping; (2) infomapping and infocharting; and (3) information navigation, retrieval and repackaging (Roesler & Hawkins, 1994). These three key components of a TQKM process are operated in sequence (Schwarzwalder, 1999; Chen et al., 2000; Johannsen, 2000).

The above conceptual and data models are used to display, measure and visualize the dynamics of inter-communication activities among the subject experts. The computer programs utilize Microsoft PowerPoint, Excel, and Access to display and represent citation data in two and three dimensions. The continuing development of the electronic collection also utilizes three sets of computer programs with formulas for conducting entropy measurement, linear regression operation, and reviews of the overall infomap structures of the citation networks.

3.2. Fuzzy commonality model

The author has developed a FCM for identifying synthetic elements (authors in this case). The identified authors are then used for the construction of an intellectual genetic paradigm that is formed by a group of researchers who, as information generators (Tsai, 2002) of a subject knowledge field, "share an intellectual perspective regarding this subject matter" as a "consensual structure of concepts in a field" (Small, 1980) that can be represented by a network diagram (historiograph, or historiogram) (Garfield, 1979). By periodically monitoring the occurrence of these synthetic elements or cutpoints and their configurations (Goffman & Katz, 1981; Shaw, 1983) and by consistently converging the power series of diversified thinking (Zipf, 1972) into the cumulative advantage (Price, 1976) of logical thinking (collection accumulation of statements, facts, and/or figures) (Hoffman, 1980), the synthetic points (neurons) (Kosslyn, 1983) of a subject knowledge organism can be mapped and charted. Owing that "the graph could be regarded as a totality of 'atomic' statements" (Brookes, 1980), a life pattern of such synthetic elements and their related configurations can be recognized through long-term observation and elaboration, and via a parallel processing from several sub-systematic network graphs and databases. The FCM can be used to show that cognitive ability, tendency and habit may be cultivated and cumulated through a long-term forceful and influential information bonding process (Fig. 1). This constantly continued information bonding process is detailed below.

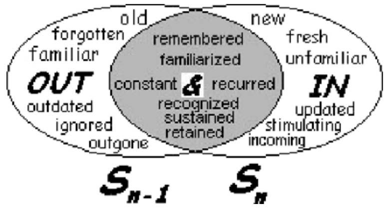

Fig. 1. Continued information bonding.

The two circles of Fig. 1 represent two consecutive sets of information within a special subject information field (a collection of a set of population), denoted as S_{n-1} and S_n, respectively. The area labeled OUT in S_{n-1} represents an *old/forgotten/familiar/outdated/ignored/outgone* informational subset (old preceding population). The area labeled IN in S_n represents a *new/fresh/unfamiliar/updated/stimulating/incoming* informational subset (new succeeding population). The overlapped (darkened) area labeled '&' represents the number of intersection or conservation of a *constant/familiarized/remembered/recognized/sustained/retained/recurred* informational subset (informational overlap) (Goffman & Warren, 1980). The diagram portrays a continuing relationship (frequently having something in common) between the two adjacent sets in a long subject information series. It is the *frequency* and the *continuity* of the *commonality* that helps maintain and strengthen the substance and continuation of such a long-term bonding relationship.

The information bonding relationship is complicated. A set of mathematical formulas can help consolidate the logical sequence and clarify the thought process of the FCM:

$$U = \mathrm{OUT} + \& + \mathrm{IN} = S_{n-1} + S_n - \&$$

$$S_{n-1} = \mathrm{OUT} + \&$$

$$S_n = \& + \mathrm{IN}$$

$$\mathrm{d}S = \mathrm{IN} - \mathrm{OUT} = S_n - S_{n-1}$$

$$\mathrm{OUT} = (U - \& - \mathrm{d}S) \div 2$$

$$\mathrm{IN} = (U - \& + \mathrm{d}S) \div 2$$

$$\& = U + \mathrm{d}S - 2 * \mathrm{IN} = U - \mathrm{d}S - 2 * \mathrm{OUT}$$

$$P_c = \& \div U$$

U denotes the number of the union of the two encountered consecutive informational sets (i.e., a unified statement of two adjacent sets of populations). S_{n-1} and S_n, denotes the two consecutive sets of information within a special subject information field. OUT denotes the number of the old/forgotten/familiar/outdated/ignored/outgone informational subset. IN denotes the number of the new/fresh/unfamiliar/updated/stimulating/incoming informational subset. & denotes the number of intersection or conservation of the constant/familiarized/remembered/recognized/ sustained/retained/recurred informational subset. dS denotes the number of difference or change between the two consecutive sets of information (i.e., the cognitively directed change); P_c denotes the critical probability of two encountered consecutive informational sets and indicates the density, intensity, sensitivity, maturity, stability, and general tendency of an information communication network's curvature.

The change or the difference (dS) between the two consecutive informational sets, resulting from the balance or coordination among IN, OUT, and the commonality bond & is the critical point at which the information generating process takes place. This common bonding, or overlapping point &, is the intersection or potential commonality of the two adjacent informational sets. The reason that the process of information bonding or overlapping is fuzzy is because it is continuously and dynamically shifting and changing. Further, it tends to chain, stretch or branch inward or outward, thus turning the process into a function of connectivity or a state of multiplicity and diversity. The process eventually becomes fuzzy and incomprehensible regardless of all connections still being tightly bonded and transparently functioned.

The FCM explains this phenomenon of informational overlapping and fuzziness and helps answer the first research question: *How can a series of subject infomaps be efficiently and effectively developed?* The second research question is: *How can the TQKM and the merger of CA and DM serve the purpose of developing the information landscaping techniques?*

In order to answer this second question, a description on how the FCM-based statistical instrument was made and how the data were gathered and analyzed, is included in the following two sections.

3.3. The making of the FCM-based statistical instrument

To respond to the aforementioned research goals, objectives and questions, the above profiled FCM data model was used to build an instrument for data analysis utilizing the MS-Excel program. Eight mathematical equations in the FCM were programmed. The correlations among the equations mutually support and double-check related cells on the Excel spreadsheet. The program can incorporate a 2-D or 3-D graphic presentation (Figs. 2 and 3). The instrument is also programmed to perform linear regression analysis, citation data mining tasks that show the variation of citation relationships and the state of continuing stability in a particular subject field (Frappaolo & Capshaw, 1999; Tsai, 1999, 2000a) and entropy measurement (detailed in Section 7). The practical use of this instrument is illustrated in the next paragraph. Using the FCM-based instrument, a set of author citation data of 1988–1995 was collected, analyzed and represented (Table 1).

The number of annual citation members (S) and the total number of all citation members that appeared in the eight-year period (1988–1995) were recorded. The annual citation data from two consecutive years were then compared to determine the annual population of recurring citation

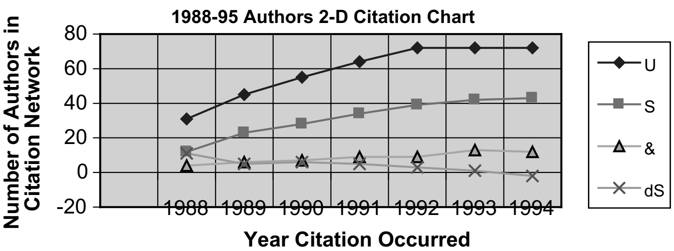

Fig. 2. FCM-based 2-D statistical data representation.

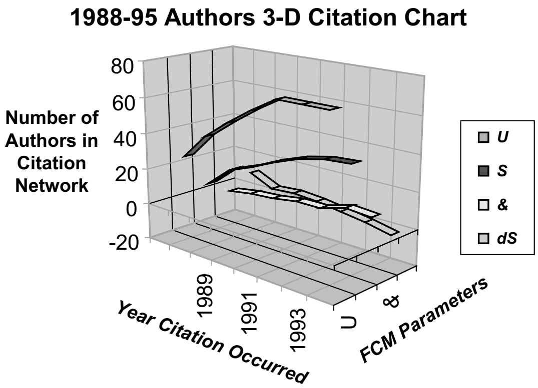

Fig. 3. FCM-based 3-D statistical data representation.

Table 1
FCM-based statistical data

Year	U	S	&	dS	P_c	IN	OUT
1988	31	12	4	11	0.1290	19	8
1989	45	23	6	5	0.1333	22	17
1990	55	28	7	6	0.1273	27	21
1991	64	34	9	5	0.1406	30	25
1992	72	39	9	3	0.1250	33	30
1993	72	42	13	1	0.1806	30	29
1994	72	43	12	−2	0.1667	29	31
1995	–	41	–	–	–	–	–
Average	58.71	32.75	8.57	4.14	0.1432	27.14	23.00

members (&). Only data sets for S and & were necessary for placement into the FCM statistical instrument. The following operations, figures, and graphs were automatically assembled, calculated, discovered, and displayed on infocharts: (1) the number of the annual new citation members (IN) and the old citation members (OUT); (2) the number of change in the annual population of citation members (dS); (3) the number of the union population of the citation members for two consecutive years (U); (4) the critical probability of two consecutive years ($P_c = \& \div U$) which indicates the strength of common bonding (i.e., the population of recurred citation members &) in the two-year union population of citation members U (Table 1, and Figs. 2 and 3). Take an example from Table 1. The values of S (i.e., the total number of citing and cited authors) in the years of 1993 and 1994 are $S_{93} = 42$ and $S_{94} = 43$ respectively. In comparing the 1993 and 1994 citation maps, thirteen members ($\& = 13$) were found to have co-occurred in these two years. Since the programmed statistical instrument uses the FCM formulas $S_{93} = \mathrm{OUT} + \&$ and $S_{94} = \& + \mathrm{IN}$ respectively among other parameters (discussed in Section 3.2 and this section earlier), the finding and input of the S and & values into the instrument immediately and automatically reveals and displays the values of $\mathrm{IN}_{94} = 30$ and $\mathrm{OUT}_{93} = 29$ on table charts as well as on 2-D and 3-D graphs. Other parametric values such as dS, U, and P_c are also instantly and simultaneously displayed. A total citation network curvature is thus formed (see Figs. 2 and 3). Any slight changes from the input sources S and & will automatically and immediately affect the entire chart and graphic representations. The P_c $(= \& \div U)$ values are particularly indicative and associative with the later entropy measurement in Section 7 (see Figs. 9 and 11). As such, both charts and graphs are useful and powerful for conducting quantitative information navigation.

This section answered the second research question: *How can the TQKM and the merger of CA and DM serve the purpose of developing the information landscaping techniques?*

4. Data gathering and analysis

By applying the FCM, a software program can be developed for performing: (1) citation data mining and information landscaping, (2) infomapping and infocharting, and (3) information navigation, retrieval and repackaging. These three sequential processes are the key components of a TQKM-based operation. Both infomaps and infocharts were used for the construction of a citation-based subject experts map-chart/depository/query–report system. This system is capable of linking a researcher to a web-based document depository (Nicotera, 1999; Tsai, 2000b; Zwies, 2000). A group of experts and their areas of expertise can be retrieved through a query–report system. Three criteria established for the selection of these subject experts are prescribed as follows: (1) quantity (S)—the selected authors must contribute a large number of publications to the subject literature during the designated observation period; (2) continuity or stability (&)—contributions of selected authors must constantly recur in the subject literature from year to year; and (3) changeability (dS)—the numbers of selected authors' populations must change from time to time, allowing dynamic shifts of population membership to occur and consequently maintain on advantageous and competitive edge. In other words, the selected authors must be highly cited, recur frequently from year to year, or play the role of "bridge" authors (Loizou, 1979) who link together two or more research clusters. The idea of this project is to apply the FCM to expedite time and effort in gathering and analyzing the citation data. By judging ranks that are based on the frequency counts of cited or citing authors, and according to the bridge authors' positions in the annual citation networks, a primary pool of authors can be determined.

As a result of the FCM operation, a set of statistical data was gathered from a citation analysis based on four prominent information researchers. These four researchers' areas of research are within the domain of cognitive information processing and coordination (or simply, cognitive coordination). Three of them are working in the same graduate school of communication, information and library science in the USA. These four researchers were selected for their outstanding research works. They are selected as seed information researchers, or synthetic cutpoints (or information landmarks in the notion of information landscaping) for mining and growing a citation network for this project. The synthetic cutpoint (Shaw, 1983), or the information landmark, may be defined as the component in a map that causes the number of components (i.e., the network clusters) in that map to increase when the component itself and its incident (or networked) lines are removed leading to a weaker relationship or a total collapse or chaos of the network. These scholars are: Nicholas J. Belkin, Brenda Dervin, Carol C. Kuhlthau, and Tefko Saracevic. Their citation data were collected from the *Social Science Citation Index (SSCI)*, 1988–1995. Derived from the citation data of these four researchers, the author citation maps reveal a group of experts' collective contributions to this subject field on cognitive coordination and its related areas (Haythornthwaite et al., 1999). The author citation data were analyzed to study its evolutionary processes. The mathematical equations formulated in the aforementioned FCM software program were applied for automatically mining the cleansed citation data (Frappaolo & Capshaw, 1999). The statistical analysis applied a frequency-filtering device, which used a set of threshold values from one to six (representing the frequency of citation) to filter the citation data. As a result of the analysis and filtration, a list of 430 authors was initially selected. The filtering/ranking mechanisms

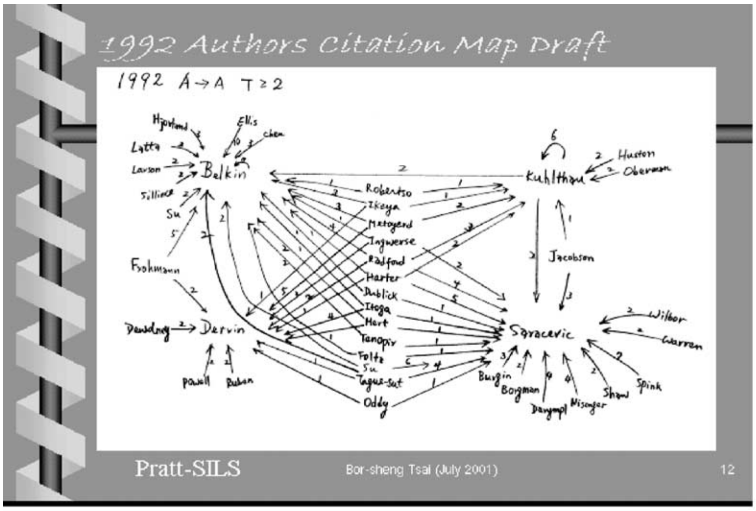

Fig. 4. A review of an annual (1992) authors citation map.

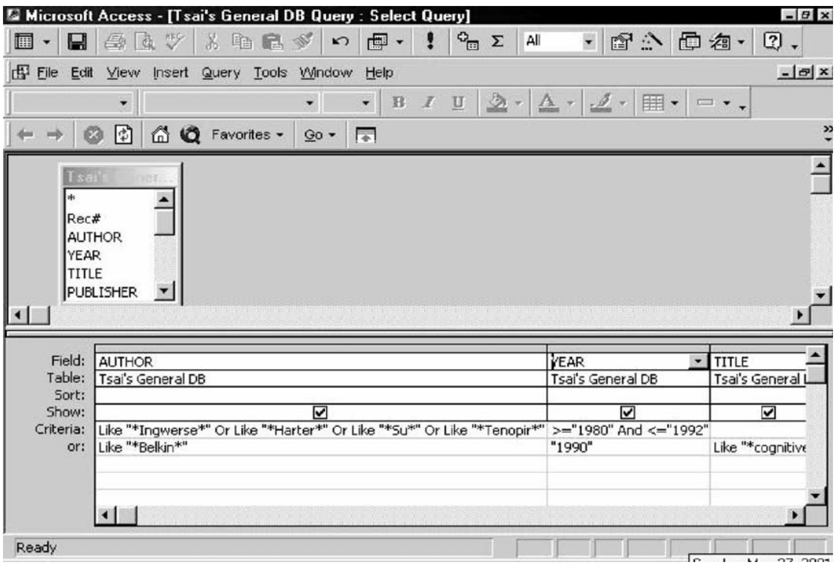

Fig. 5. An example of an SQL-based information retrieval.

and the principles of continuing recurrence, change and stability—as discussed above and elaborated in the previous section on the FCM—were the guideposts for selecting a second tier of 181 contributors from the original pool and for finalizing a third tier of 80 most significant contributors from the second tier. These 80 contributors formed the basis for the creation of the subject experts map-chart/depository/query–report system (Figs. 2–7) (Hodgson, 1999).

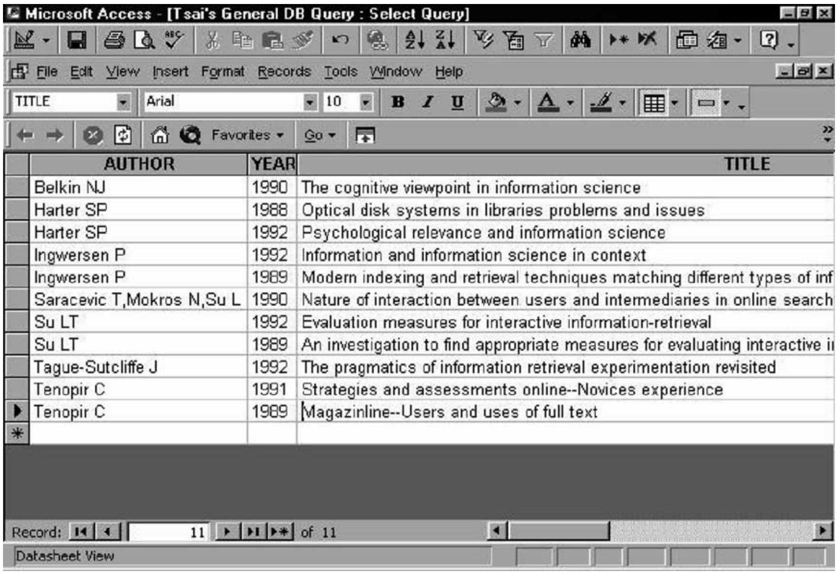

Fig. 6. A result of an SQL-based information retrieval.

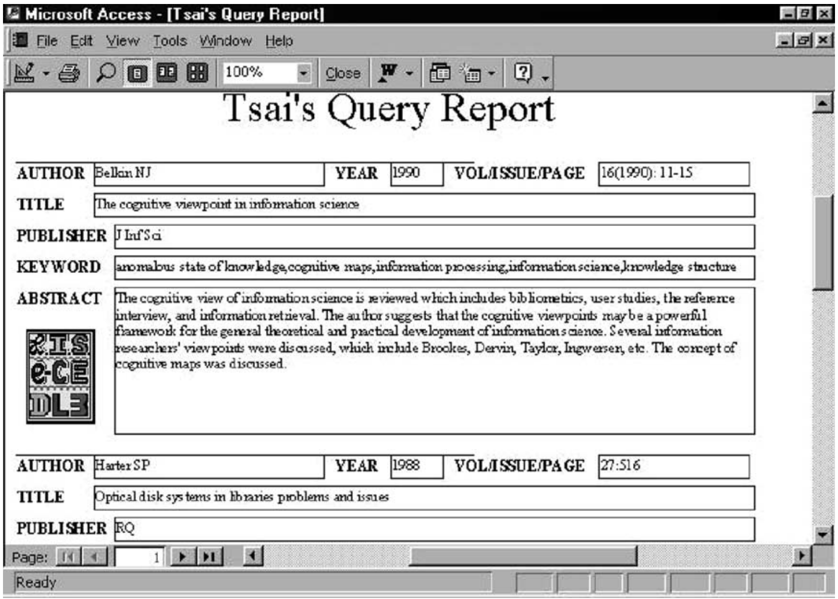

Fig. 7. Integrated query–report function for instant information repackaging and delivery.

An assessment was conducted to evaluate the infomap structure of the citation networks for electronic collection development (see Section 6 on information navigation, retrieval and repackaging, and Section 7 on assessing infomap structures for electronic collection development).

This section answered to a greater extent, the second research question: *How can the TQKM and the merger of CA and DM serve the purpose of developing the information landscaping techniques?*

5. Constructing a subject information architecture

5.1. The drafting and creation of digital citation maps as IA blueprints

As discussed earlier in Section 2 (research question 1) on the development of subject infomaps, this project has utilized a hand-drafting technique for an easy manual drawing of citation maps. Based on the previously completed citation analysis, each one of the eight annual citation maps (1988–1995) was manually drawn and completed within a few minutes. The only tool needed is a 2B pencil and a few drafting papers. After the completion of the drawing, the drafted papers were scanned by an ordinary image scanner and converted into digital forms (in this case, *.gif or *.jpg). The digitized maps were then represented in MS-PowerPoint format for further linking and displaying on the web platform. An eight-year infomap series was thus easily made without involving much computer skills or spending much work time or computer memory (only a few kilobytes were consumed for each map draft). The cost of the drawing and the scanning operations was extremely minimal. Therefore, this infomapping technique allows interested librarians or information professionals to easily perform this task. Additionally, a regular and inexpensive 3.5-in. floppy disk can sufficiently store many (years of) infomaps and allows librarians or information professionals to transfer the data easily, conveniently and cost-effectively.

As discussed in Section 3.1 (particularly, column 6 on IA), this research project has also focused on the design and the construction of a TOKM system—a knowledge portal that is capable of operating a 24-7-365 web-supported database management system carrying and continuously advancing and expanding the core of an electronic subject experts collection for distance learning. For continuous capturing and depositing of the scholarly communication patterns of these expert researchers, the building of a subject-focused electronic research collection is essential and crucial. A project for constructing a web-based special subject experts map-chart/depository/query–report system focusing on cognitive coordination was conducted. The aforementioned 80 selected significant contributors and their citation networks were the blueprints for the construction of such a system. By combining the HTML scripts, local web links, PowerPoint and Excel-based infomaps and infocharts, and by integrating the storage function with the SQL (structured query language)-based querying-reporting functions of Microsoft Access, this system is equipped with instant navigation, searching, displaying and repackaging capabilities (Notess, 2000a, 2000b; Roberts, 2000; Rogers, 2000).

To devise the directing and switching function, three MS-Excel, PowerPoint, and Access-based programs were designed and installed in the web server located at http://rand.pratt.edu/~btsai site. The PowerPoint and Excel-based program (located at http://rand.pratt.edu/~btsai/ESlideShow/Au88-95CitaChartMaps_files/v3_document.htm) provides citation data, infocharts, and infomaps in 2-D and 3-D formats representing citation frequency (S), critical probability (P_c), entropy measure (H_c), and entropy constant (E_c) in 1988–1995 authors citation networks. The Access-based program (located at http://rand.pratt.edu/~btsai/e-depot.html) provides the Access 2000 version (97 version also available) for users to connect with the electronic depository and to access the electronic collection (located at http://rand.pratt.edu/~btsai/Tsai2000DBMS/Tsai2000.mdb for 2000 version, or http://rand.pratt.edu/~btsai/Tsai2000DBMS/Tsai97.mdb for 97 version).

To use this system for general review purposes, a researcher can open an infomap to identify a desired expert as a starting point and connect him/her with a few neighboring experts in the nearby networked clusters who are also known to the researcher. By using the name search in the MS-Access SQL-based query unit, researchers can efficiently retrieve an electronic package of records related to the desired author(s). A researcher can also open an infochart to view the data flows in terms of quantity (S), change of quantity (dS), continuity and stability of sub-population (&), and the subtotal union population (U). By identifying a particular time period, a researcher can use the same MS-Access SQL-based query function for retrieving the records of the desired time span. The experts depository also allows keyword searching, if a researcher does not know any particular experts. The idea of knowledge portal architecture is realized by the use of the infomap and infochart's arrow pointing and the screen-switching mechanisms, as well as by the MS-Access SQL-based query function. Although the direct web-linking mechanism from an infomap or an infochart with the data records was possible, it cannot compete with the integrated query–report function of web-supported MS-Access.

This section answered the third research question: *How can a web-based subject experts depository—equipped with infomap–infochart and query–report capabilities—be constructed?*

6. Information navigation, retrieval and repackaging

The operations for information navigation, retrieval and repackaging are illustrated below.

(1) By concurrently applying the computer's multitasking function to review the 2-D or 3-D infocharts (Figs. 2, 3, 8 and 9) and the infomap series (Figs. 4 and 10) and tables (Tables 1 and 2), a researcher can immediately view a general picture, glance at the symbols, numbers and the changing waves of citing authors, and instantly learn and recognize the critical connecting points of the targeted subject field in an eight-year (1988–1995) time series (Havre et al., 2002).
(2) Based on the numerical data that were presented on the citation data charts (Tables 1 and 2), an infomap draft of a particular year—for example, 1992, the year when the amount of information seems to start to stabilize—may be chosen for further review (Fig. 4) (see Section 8 (6): Assessment of ranking mechanisms).
(3) After reviewing the chosen annual (i.e., 1992) infomap, five co-citers were selected (e.g., Harter, Ingwersen, Su, Tague-Sutcliffe, and Tenopir) who all served as bridge authors that sufficiently connected together all four co-cited seed information researchers (i.e., Belkin, Dervin, Kuhlthau, and Saracevic) (see Section 8 (4): Bridge authors).
(4) By applying the web-based MS-Access SQL function (particularly using its term truncation to retrieve co-authors) and query–report functions, the desired significant contributors' works can be instantly retrieved in chronological order (Figs. 5 and 6).
(5) The retrieved information can be modified, repackaged, reported, and delivered to a distant researcher (Fig. 7).

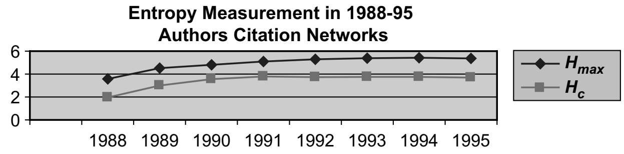

Fig. 8. Entropy measurement in 1988–1995 authors citation networks.

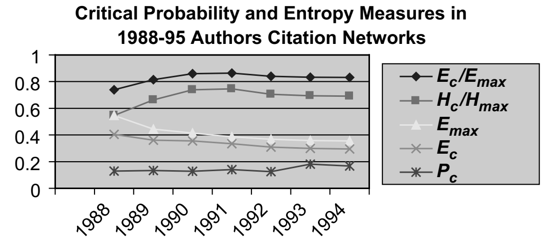

Fig. 9. Critical probability and entropy measures in 1988–1995 authors citation networks.

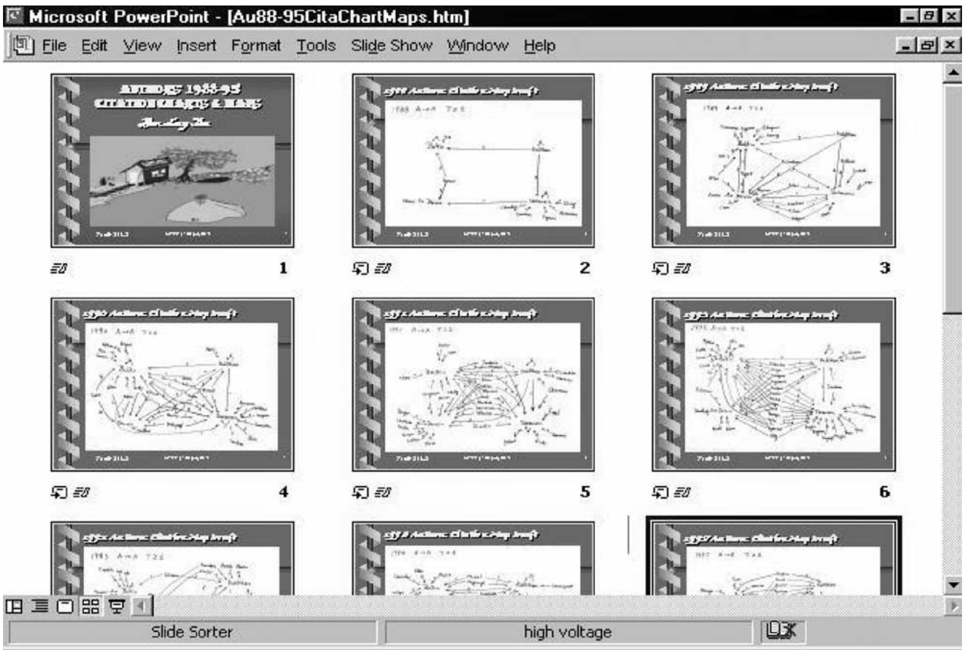

Fig. 10. An overview of the 1988–1995 authors citation map series.

(6) The contents of the information packages can be analyzed to discover new ideas from the elaborations on the interactions and correlated views and commonly agreed upon insights (co-citations) among the five selected bridge authors and the four seed information researchers (Garfield, 1979).

Table 2
Critical probability (P_c), entropy measure (H_c), and entropy constant (E_c) in 1988–1995 authors citation networks

Year	H_{max}	H_c	E_c/E_{max}	H_c/H_{max}	E_{max}	E_c	P_c
1988	3.5849	1.9591	0.7393	0.5465	0.5466	0.4041	0.1290
1989	4.5236	2.9967	0.8140	0.6625	0.4435	0.3610	0.1333
1990	4.8074	3.5481	0.8591	0.7381	0.4144	0.3560	0.1273
1991	5.0875	3.7929	0.8635	0.7455	0.3868	0.3340	0.1406
1992	5.2854	3.7229	0.8394	0.7044	0.3681	0.3090	0.1250
1993	5.3923	3.7439	0.8334	0.6943	0.3583	0.2986	0.1806
1994	5.4263	3.7486	0.8314	0.6908	0.3552	0.2953	0.1667
1995	5.3576	3.6981	0.8307	0.6903	0.3615	0.3003	–
Average	4.9331	3.4013	0.8219	0.6840	0.4043	0.3323	0.1432

Note: H_{max} represents the maximal amount of entropy measured. E_{max} represents the constant of the maximal amount of entropy H_{max}.

(7) The subtotal and the maximal amount of information that the information packages and the annual (1992, in this case) citation network generated, may be calculated and shown to the researcher. By applying the entropy measurement, which is depicted in the next section, this step can be completed.

This section answered the fourth research question: *How can the TQKM-guided infomaps and the FCM-based infocharts be effectively used as navigation agents for information querying and reporting?*

7. Assessing infomap structures for electronic collection development

In order to continuously develop and improve the electronic subject collection, it is necessary to evaluate the potential impact, potency, energy concentration, or density (Rifkin & Howard, 1980; Avramescu, 1980) that the citation networks may generate. For this purpose, an effort is made to assess the infomap structures. A feedback (looping mechanism) analysis helps in recognizing and understanding patterns that subject experts unconsciously cultivated during the time of their reference seeking and citation networking within a closely related subject field.

The evaluation of the looping mechanism adopts the *entropy measurement*, which is capable of assessing the "informational uncertainty" (Guy & Zunde, 1985) and can be encoded in "a set of probabilities assigned to the set of possibilities" that may cause "an adjustment in a probability assignment" (Tribus & McIrvine, 1971). From the viewpoint of Shannon and Weaver (1963), information is treated as "a measure of absence of uncertainty." Their application of entropy measurement stimulated chaos theorists to "redefine chaos as maximum information" (Hayles, 1989). Hicks and Essinger (1991) valued the reduction or closure of "cognitive load," which they considered as "the burden being placed on the brain at any one time," and "the cause of stress, fatigue and making mistakes." Similarly, Goffman and Warren regarded information as a measure of "removal of uncertainty" of the finite scheme. This means that the amount of information of a finite scheme can be measured once the uncertainty of the finite scheme is removed. The finite scheme here represents “a set of mutually exclusive and exhaustive events whose probability (P_i) of occurrence is known” (Goffman & Warren, 1980). The uncertainty of the finite scheme can be measured by the function at a critical point that is represented as $H_c = -\lambda \sum P_i \log P_i$ where λ is a constant ($= 3.32193$), “i” is ranging from 1 to n (“n” is the population of authors in this case), and P_i is the probability of occurrence of authors in a citation network at a given point of a specified time line (1988–1995 in this case). The *maximal amount of entropy* in each year is represented as $H_{max} = -\lambda \log(1/n)$. Another *entropy constant* that was measured in this assessment complies with the following formulas: the annual *entropy constant* $\boldsymbol{E}_c$ = *square root of the measured entropy divided by the total number of authors in that year, namely, the square root of the average entropy measure of that year*. Similarly, the annual *maximal entropy constant* $\boldsymbol{E}_{max}$ = *square root of the measured maximal entropy divided by the total number of authors in that year, namely, the square root of the average maximal entropy measure of that year*. The results of the calculations are presented in Table 2.

To illustrate how the data were derived, an example taken from Table 2 is provided. Still using 1993 as an instance, the $S_{93} = 42$ indicates that a basic pool of 42 citation members was recorded in 1993. According to the formula for the entropy measure $H_{max} = -\lambda \log(1/n)$ where $\lambda = 3.32193$, the maximum entropy H_{max} for 1993 will be equal to 5.3923. Accordingly, the value of H_c ($= -\lambda \sum P_i \log P_i$) for 1993 (with $\lambda = 3.32193$) is calculated to be equal to 3.7439 accounting for the fact that 42 citation members were regrouped into five clusters with 7, 10, 3, and 5 citing members respectively surrounding each one of the four seed (cited) members (namely, Belkin, Dervin, Kuhlthau, and Saracevic), and 17 independent bridge (co-citing) members who linked together all other four clusters. As a result, H_c/H_{max} for 1993 is equal to 0.6943 (from 3.7439 ÷ 5.3923). On the other hand, the annual *maximal entropy constant* E_{max} and the annual *entropy constant* E_c can be easily calculated by taking the square root of the annual average entropy amount of H_c and H_{max} respectively. Therefore, E_{max} for 1993 is equal to 0.3583, which is the square root of 0.1284 (from 5.3923 ÷ 42). Similarly, E_c is equal to 0.2986, which is the square root of 0.0891 (from 3.7439 ÷ 42). Hence, E_c/E_{max} for 1993 is equal to 0.8334 (from 0.2986 ÷ 0.3583). As the association goes, it is noticed that E_c/E_{max} (such as 0.8334) is the square root of H_c/H_{max} (such as 0.6943).

A further comparative analysis of entropy data among P_c, H_c, H_{max}, E_c, E_{max}, and the correlation coefficients among P_c, H_c/H_{max}, and E_c/E_{max} showed parallel data flow directions of the citation networks (Figs. 8 and 9). A set of linear regression formulas was determined (Table 3). They indicated that the average of the critical probability P_c was between 0.1139 and 0.1400, and the average entropy density H_c/H_{max} was between 0.6170 and 0.6800. The correlation coefficients of the regression straight lines for both the critical probability P_c and the entropy density H_c/H_{max} were 0.0073 and 0.0165 respectively (Fig. 11). It was also observed that starting in 1992, sufficient data showed a stable data flow continued to occur within the citation networks. Thus Table 2 and Figs. 8 and 9 indicate the *potential impact of citation networks* and can function in the same way as the previous FCM-based 2-D and 3-D infocharts (Table 1, and Figs. 2 and 3). They indicate the amount and the change of information and assist information navigation and retrieval activities. These tables and figures can also provide both quantity and quality management guidelines for electronic collection development.

This section answered the fifth research question: *How can the major FCM parameters, entropy measures, and the linear regression approach help in continuously monitoring the structural change of citation networks and assessing the development of this web-based subject collection?*

Table 3
Critical probability (P_c) vs. entropy density (H_c/H_{max}) and regression straight lines $(P_c)_L$ and $(H_c/H_{max})_L$

	Year	$(H_c/H_{max})_L$	$(H_c/H_{max})_L$	$(P_c)_L$	P_c	X^2	XY_1	XY_2	
	20	0.62	0	0.11	0	0	–	–	
	1	0.63	0.5465	0.12	0.1290	1	0.1290	0.5465	$B = \overline{Y} - m\overline{X}$
	2	0.65	0.6625	0.13	0.1333	4	0.2666	1.3250	$B_1 = 0.1139$
	3	0.67	0.7381	0.14	0.1273	9	0.3819	2.2143	$B_2 = 0.6170$
	4	0.68	0.7455	0.14	0.1406	16	0.5624	2.9820	$m_1 = 0.0073$
	5	0.70	0.7044	0.15	0.1250	25	0.6250	3.5220	$m_2 = 0.0165$
	6	0.72	0.6943	0.16	0.1806	36	1.0836	4.1658	$\overline{Y} = mX + B$
	7	0.73	0.6908	0.17	0.1667	49	1.1669	4.8356	$\overline{Y}_{1L} = 0.0073X + 0.1139$
									$\overline{Y}_{2L} = 0.0165X + 0.0617$
Total	28	5.40	4.7821	1.12	1.0025	140	4	20	
Average	4.00	0.67	0.68	0.14	0.14	–	–	–	

X = Time (year) series. $(Y_1)_L = (P_c)_L$ (regression straight line for critical probabilities). $(Y_2)_L = (H_c/H_{max})_L$ (regression straight line for entropy densities).

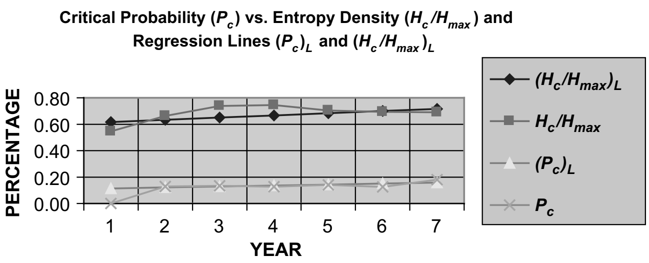

Fig. 11. Critical probability (P_c) vs. entropy density (H_c/H_{max}) and the regression straight lines $(P_c)_L$ and $(H_c/H_{max})_L$.

8. Discussion

(1) Approachability—Information landscaping techniques allow environmental scanners, who are monitoring a particular subject information field, to quickly capture, draft, scan, and reconstruct the images of a virtual citing scene (namely, an *imaginary citation space*) at a particular moment and to instantly represent these images in electronic form within a web environment (Vail III, 1999). A series of directed informaps and a set of dynamic infocharts are easily constructed using MS-Excel spreadsheets (Feicht, 1999). They are effective information indicators and navigators. The supporting depository and query–report functions immensely enhance information retrieval activity and electronic collection development (Van Den Hoven, 2001).

(2) Applicability—The infocharts and the infomaps greatly boost retrieval speed and quality management since the selected experts are highly filtered and condensed elements (many of them are accomplished scholars) of a subject research field. They enhance the study of collaboration relationship and structure among leading researchers (or key story characters) in the same field (or story board) (Shanahan & Shanahan, 1997), and in this case, in the same

institution as well. They also serve as the blueprints for information architecture construction and electronic collection development.

(3) Scalability—A virtual learning and referencing environment is instantly created and represented. This system's knowledge base (i.e., infomaps and infocharts) and database (i.e., e-depository) can be easily and continuously revised and expanded (Monahan, 2001). The techniques applied in this project enable the instantaneous measurement and display of the dynamics of a special subject information field. The web-based subject experts map-chart/depository/query–report model may be applied to other subject fields as well (Withers, 1999).

(4) Bridge authors—Co-citers are bridge authors. They are the joints of citation networks and the backbones and shafts of the subject information architecture. They "cross-validate" the quality of a group of subject experts' scholarly works, social relationships and information communication networks (McCain, 1989). Continued electronic collection development relies on the ongoing monitoring, filtering and association with these bridge authors.

(5) Interrelationships among seed information researchers—Interactivities among four seed information researchers can be easily viewed from the eight-year infomap series. Three researchers (Belkin, Kuhlthau, and Saracevic) are faculty members in Rutgers University. While some bridge authors indirectly connected all of the four seed researchers, Kuhlthau was the only seed researcher who directly and frequently made contact with each one of the other three seed researchers. These three scholars (Belkin, Dervin, and Saracevic) occasionally communicated among themselves during the observation period. All four seed researchers except Dervin frequently self-cited their own publications.

(6) Assessment of ranking mechanisms—The results of the analyses of infomap structures indicated that the amount of information that the citation networks generated might have reached the level of stability in 1992–1995. However, it is also possible that a retreat might occasionally occur due to low scientific validity or low distinctiveness of the group's interests and its work (Griffith & Mullins, 1972). The follow-up 1996–2002 citation data are to be studied to verify this point.

(7) Linear regression approach—It was found that the correlation coefficients for critical probability ($P_c = 0.0073$) and the entropy density ($H_c/H_{max} = 0.0165$) in the regression straight lines were both at a low level of growth rate (between 0.7 and 1.7 percent). In other words, the number of commonly recurring authors or the frequency of intercommunications among authors in the citation networks may indicate that the field of cognitive coordination study is still in its infancy. This indication helps us in forecasting the drastic or mild changes of collection size or patterns for the next few years (Monmonier, 1999). Since the number of major subject experts in this field has been identified (approximately 80 authors), a population of quality citation membership can be gradually expanded. More experts in this field can be easily identified and efficiently added to the infomap series.

9. Significance and benefits

Citation data mining and information landscaping are useful informetric and information architectural techniques that can help in: building an information resource management system, developing an electronic collection, and cultivating a virtual referencing environment (Hiltz &

Wellman, 1997). High quality experts and expertise can be identified and selected as guideposts for electronic collection development. The application of entropy measurement, as an indicator of potential impact, is useful for information navigation and infomap structural analysis and electronic collection development. The constructed information architecture can enhance the sharing of information resources cumulated and the knowledge discovered. The automatic calculations from the statistical instruments based on the three metric devices—FCM, entropy measurement, and linear regression—can efficiently and effectively devise an informational oscillogram to help analyze the structure and the growth rate of a target population through infomapping and infocharting and to guide information retrieval and collection development activities (Goffman, 1971). Most of all, the simple but powerful hand drawing allows an information architect to quickly and accurately capture, sketch, and generate or update the electronic images of the subject experts' leadership (general direction) in a subject research field. The three metric devices help justify, navigate and predict the general directions of the leadership of a subject research field (Goffman & Harmon, 1971). Finally, the infomap and infochart-coordinated subject information architecture can enhance the accumulation, discovery, and sharing of information resources and knowledge (Van Den Hoven, 2001). The integrated query–report system can effectively provide instant electronic repackaging and delivery services.

10. Conclusion

This project applies citation data mining and information landscaping for the total and parallel mapping and charting of major subject experts' citation network relationships within a particular subject information field. The work is supported by a web presentation system and an integrated query–report system utilizing the popular Microsoft web-based PowerPoint, Access, and Excel software packages. The infomaps and infocharts were not only used as blueprints but also served as navigation agents in searching, building and operating a subject experts map-chart/depository/query–report system. The capability and the usability of informetric applications of the FCM, the entropy measurement, and the linear regression approach significantly contribute to the continuous development and improvement of the system's quality (Dobransky & Wierman, 1996; Tsai, 2000b; Tsai, 2002). Continued efforts will be devoted to the analyses of links, indexing, retrieval, and contents of internal/external web documents and images, as well as to the expansion of systems and databases from 1996 to the present date. Several software packages for image processing and multimedia are currently under examination. They will be added to enhance the total function and quality of the knowledge management in this subject field.

References

Adams, K. C. (2001). Representing knowledge in enterprise portals. *KMWorld, 10*(5), 18–21.

Aldakhilallah, K. A., & Parente, D. H. (2002). Redesigning a square peg: total quality management performance appraisals. *Total Quality Management, 13*(1), 39–51.

Avramescu, A. (1980). Coherent information energy and entropy. *Journal of Documentation, 36*(4), 293–312.

Banerjee, K. (1998). Is data mining right for your library? *Computers in Libraries, 18*(10), 1–5.

Black, S. (2001). Using citation analysis to pursue a core collection of journals for communication disorders. *Library Resources and Technical Services, 45*(1), 3–9.

Brodley, C. E., & Lane, T. S. (1999). Knowledge discovery and data mining. *American Scientist, 87*(1), 54–61.

Brookes, B. C. (1980). Measurement in information science: objective and subjective metrical space. *Journal of the American Society for Information Science, 31*(4), 248–255.

Brookes, B. C. (1990). Biblio-, sciento-, infor-metrics??? What are we talking about? In L. Egghe, & R. Rousseau (Eds.), *Informetrics 89/90* (pp. 31–43). New York, NY: Elsevier Science Publishers.

Brooking, A. (1999). *Corporate memory: Strategies for knowledge management*. New York, NY: International Thomson Business Press.

Bryan, W. (1996). What is total quality management? *New Directions for Student Services, 76*(Winter 1996), 3–15.

Chen, C. M. (1999). *Information visualization and virtual environments* (pp. 15–18). New York, NY: Springer.

Chen, L., Sakaguchi, T., & Frolick, M. N. (2000). Data mining methods, applications, and tools. *Information Systems Management* (Winter 2000), 65–70.

Choo, C. W. (1996). The knowing organization: How organizations use information to construct meaning, create knowledge and make decisions. *International Journal of Information Management, 16*(5), 329–340.

Chou, D. C., & Chou, A. Y. (1999). A manager's guide to data mining. *Information Systems Management, 16*(4), 33–41.

Chopoorian, J. A. et al. (2001). Mind your business by mining your data. *Advanced Management Journal, 66*(2), 45–51.

Chua, R. C. H. (2001). What you need to know about six sigma. *Productivity Digest* (December 2001), 37–44.

Cronin, B. (2000). Strategic intelligence and networked business. *Journal of Information Science, 26*(3), 133–138.

Coyle-Shapiro, J. A. M. (1999). Employee participation and assessment of an organizational change intervention: a three-wave study of total quality management. *Journal of Applied Behavioral Science, 35*(4), 439–456.

Davenport, T. H., & Prusak, L. (1997). *Information ecology: Mastering the information and knowledge environment*. New York, NY: Oxford University Press.

Davenport, T. H., & Prusak, L. (1998). *Working knowledge: How organizations manage what they know*. Boston, MA: Harvard Business School Press.

Davies, R. (1990). Generating new knowledge by retrieving information. *Journal of Documentation, 46*(4), 368–372.

Dobransky, M. K., & Wierman, M. J. (1996). Genetic algorithms: a search technique applied to behavior analysis. *International Journal of General Systems, 24*(1–2), 125–135.

Feicht, L. (1999). 3-D graphing, contour graphs, topographical maps, and matrices using spreadsheets. *Mathematics Teacher, 92*(2), 166–174.

Frappaolo, C., & Capshaw, S. (1999). Knowledge management software: capturing the essence of know-how and innovation. *The Information Management Journal* (July 1999), 44–48.

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD process for extracting useful knowledge from volumes of data. *Communications of the ACM, 39*(11), 27–34.

Gallupe, B. (2001). Knowledge management systems: surveying the landscape. *International Journal of Management Reviews, 3*(1), 61–77.

Garfield, E. (1979). *Citation indexing—Its theory and application in science, technology, and humanities* (pp. 81–86, 142–143). New York, NY: John Wiley & Sons.

Goffman, W. (1971). A mathematical method for analyzing the growth of a scientific discipline. *Journal of the Association for Computing Machinery, 18*(2), 173–185.

Goffman, W., & Harmon, G. (1971). Mathematical approach to the prediction of scientific discovery. *Nature, 229*, 103–104.

Goffman, W., & Katz, M. J. (1981). Performance of ontogenetic patterns. *Philosophy of Science, 48*, 438–453.

Goffman, W., & Warren, K. S. (1980). *Scientific information systems and the principle of selectivity* (pp. 22–25). New York, NY: Praeger Publishers.

Gregory, V. L. (2000). Knowledge management and building the learning organization. In T. K. Srikantaiah, & M. E. D. Koenig (Eds.), *Knowledge management for the information professional* (pp. 161–179). Medford, NJ: Information Today.

Griffith, B. C. (1989). Understanding science: studies of communication and information. *Communication Research, 16*(5), 600–614.

Griffith, B. C., & Mullins, N. C. (1972). Coherent social groups in scientific change. *Science* (September 1972), 959–964.

Guenther, K. (2000). Applying data mining principles to library data collection. *Computers in Libraries* (April 2000), 60–63.

Guy, A. G., & Zunde, P. (1985). The impact on ‘‘informational entropy’’ of disproof of the Boltzmann relation. In J. C. Agrawal, & P. Zunde (Eds.), *Empirical foundations of information and software science* (pp. 295–302). New York, NY: Plenum Press.

Hackbarth, G., & Grover, G. (1999). The knowledge repository: organizational memory information systems. *Information Systems Management, 16*(3), 21–30.

Haur, M. (1999). Three thousand years of knowledge management: What can we learn from science? *Information Services and Use, 19*, 37–44.

Havre, S., Hetzler, E., Whitney, P., & Nowell, L. (2002). ThemeRiver: visualizing thematic changes in large document collections. *IEEE Transactions on Visualization and Computer Graphics, 8*(1), 9–20.

Hayles, N. K. (1989). Chaos as orderly disorder: shifting ground in contemporary literature and science. *New Literary History, 20*(2), 305–322.

Haythornthwaite, C. et al. (1999). Mapping the dimensions of a dynamic field. *Journal of the American Society for Information Science, 50*(12), 1092–1094.

Hicks, R., & Essinger, J. (1991). *Making computers more human: Designing for human–computer interaction* (p. 58). Oxford: Elsevier Science Publishers.

Hiltz, S. R., & Wellman, B. (1997). Asynchronous learning networks as a virtual classroom. *Communications of the ACM, 40*(9), 44–49.

Hodgson, C. A. (1999). Creating an experts database. *EContent* (October/November 1999), 15–20.

Hoffman, E. (1980). Defining information: an analysis of the information content of documents. *Information Processing and Management, 16*, 291–304.

Hood, W. W., & Wilson, C. S. (2001). The literature of bibliometrics, scientometrics, and informetrics. *Scientometrics, 52*(2), 291–314.

Holsapple, C. W., & Joshi, K. D. (2002). Knowledge management: a threefold framework. *The Information Society, 18*, 47–64.

Jern, M. (1998). The advent of visual data mining on the World Wide Web. *Information Systems Management, 15*(3), 66–69.

Johannsen, C. G. (2000). Total quality management in a knowledge management perspective. *Journal of Documentation, 56*(1), 42–54.

Jurisica, I. (2000). Systematic knowledge management and knowledge discovery. *Bulletin of the American Society for Information Science* (October/November 2000), 9–12.

Kanter, J. (1999). Knowledge management, practically speaking. *Information Systems Management, 16*(4), 7–15.

Killingsworth, B. L., Harden, M. B., & Dellana, S. A. (1999). Total quality improvement in the classroom: integrating TQM in a systems analysis and design course. *College Student Journal, 33*(3), 465–476.

Kirschenbaum, M. (1999). Lucid mapping: information landscaping and three-dimensional writing space. *Leonardo, 32*(4), 261–268.

Koenig, M. E. D., & Srikantaiah, T. K. (2000). The evolution of knowledge management. In T. K. Srikantaiah, & M. E. D. Koenig (Eds.), *Knowledge management for the information professional* (pp. 23–36). Medford, NJ: Information Today.

Kosslyn, S. M. (1983). *Ghosts in the mind's machine: Creating and using images in the brain* (pp. 24–25). New York: W. W. Norton & Co.

Laszlo, G. P. (1999). Implementing a quality management program—Three Cs of success: commitment, culture, cost. *The TQM Magazine, 11*(4), 231–237.

Levine, L. (2001). Integrating knowledge and processes in a learning organization. *Information Systems Management, 18*(1), 21–33.

Lin, X. (1997). Map displays for information retrieval. *Journal of American Society for Information Science, 48*(1), 40–54.

Loizou, G. (1979). Graph models of complex information sources. *Information Processing and Management, 15*(3), 127–131.

Lord, J., & Ragon, B. (2001). Working together to develop electronic collections. *Computers in Libraries, 21*(5), 40–44.

Masters, D. G. (1996). *Total quality management in libraries* (6 pp.). Syracuse, NY: ERIC Clearinghouse on Information and Technology Syracuse University.

McCain, K. W. (1989). Mapping authors in intellectual space: population genetics in the 1980s. *Communication Research, 16*(5), 667–681.

McCain, K. W. (1990). Mapping authors in intellectual space: a technical overview. *Journal of the American Society for Information Science, 41*(6), 433–433.

McCarthy, J. (2000). Phenomenal data mining. *Communications of the ACM, 43*(8), 75–79.

Monahan, D. (2001). Explorations of the wilderness: making maps with GPS. *GPS (Global Positioning System) World, 12*(10), 32–38.

Monmonier, M. (1999). Mapping the storm. *Weatherwise, 52*(3), 26–33.

Morrow, N. M. (2001). Knowledge management: an introduction. In M. E. Williams (Ed.), *Annual review of information science and technology* (vol. 35, pp. 381–422). Medford, NJ: Information Today.

Myburgh, S. (2000). The convergence of information technology and information management. *The Information Management Journal* (April 2000), 4–16.

Nicotera, C. L. (1999). Information access by design: electronic guideline for librarians. *Information Technology and Libraries, 18*(2), 104–109.

Notess, G. R. (2000a). Up and coming search technologies. *Online* (May/June 2000), 75–77.

Notess, G. R. (2000b). Online storage opportunities. *EContent* (December 2000), 62–64.

O'Neill, L. M. (2001). Universal design for learning: making education accessible to all learners. *Syllabus* (April 2001), 31–32.

O'Leary, M. (2002). E-Global library advances the virtual library. *Information Today, 19*(3), 19.

Periasamy, K. P., & Feeny, D. F. (1997). Information architecture practice: research-based recommendations for the practioner. *Journal of Information Technology, 12*, 197–205.

Price, D. J. de S. (1965). Networks of scientific papers. *Science, 149*(3683), 510–515.

Price, D. J. de S. (1976). A general theory of bibliometric and other cumulative advantage process. *Journal of the American Society for Information Science, 27*(5), 292–306.

Price, D. J. de S. (1979). The revolution in mapping of science. *Proceedings of the American society for information science annual meeting* (vol. 16, pp. 249–253).

Ravichandran, T., & Rai, A. (2000). Total quality management in information systems development: key constructs and relationships. *Journal of Management Information Systems, 16*(3), 119–155.

Rifkin, J., & Howard, T. (1980). *Entropy: A new world view* (pp. 33–43). New York, NY: Viking Press.

Roberts, G. (2000). Designing a database-driven web site, or, the evolution of the infoiguana. *Computers in Libraries* (October 2000), 26–32.

Roesler, M., & Hawkins, D. T. (1994). Intelligent agents: software servants for an electronic information world [and more!]. *Online* (July 1994), 19–32.

Rogers, E. (2000). Designing a web-based desktop that's easy to navigate. *Computers in Libraries* (April 2000), 34–40.

Schwarzwalder, R. (1999). Librarians as knowledge management agents. *EContent* (August/September), 63–65.

Sengupta, I. N. (1992). Bibliometrics, informetrics, scientometrics and librametrics: an overview. *Libri, 42*(2), 75–98.

Shanahan, T., & Shanahan, S. (1997). Character perspective charting: helping children to develop a more complete concept of story. *Reading Teacher, 50*(8), 668–677.

Shannon, C. E., & Weaver, W. (1963). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

Shaw, W. (1983). Statistical disorder and the analysis of a communication-graph. *Journal of the American Society for Information Science, 34*(2), 146–149.

Small, H. (1980). Co-citation context analysis and the structure of paradigms. *Journal of Documentation, 36*(3), 183–196.

Small, H. (1999). Visualizing science by citation mapping. *Journal of the American Society for Information Science, 50*(9), 799–813.

Small, H., & Garfield, E. (1985). The geography of science: discipline and national mapping. *Journal of Information Science, 11*, 147–159.

Smith, L. C. (2000). Knowledge discovery, capture and creation. *Bulletin of the American Society for Information Science* (December/January 2000), 11–12.

St Clair, G. (1997). Defining quality management in information services. In G. St Clair (Ed.), *Total quality management in information services* (pp. 45–65). West Sussex, UK: Bowker-Saur.

Stewart, T. A. (1995). Mapping corporate brainpower. *Fortune, 132*(9), 209–212.

Stromquist, N., & Samoff, J. (2000). Knowledge management systems: on the promise and actual forms of information technologies. *Compare, 30*(3), 323–332.

Tague-Sutcliffe, J. (1992). An introduction to informetrics. *Information Processing and Management, 28*(1), 1–3.

Taylor, R. S. (1982). Value-added processes in the information life cycle. *Journal of the American Society for Information Science* (September 1982), 341–346.

Thearling, K. (2000). Data mining and customer relationships. In A. Berson, S. Smith, & K. Thearling (Eds.), *Building data mining applications for CRM*. New York, NY: McGraw Hill. Available: http://www3.shore.net/~kht/text/whexcerpt.htm (April 15, 2002).

Thearling, K., et al. (2001). Visualizing data mining models. In U. Fayyad, et al. (Eds.), Information visualization in data mining and knowledge discovery. Available: http://www3.shore.net/~kht/text/dmviz/modelviz.htm (April 15, 2002).

Thompkins, W. T., Jr. (1998). Information Analysis. White Papers: NCSA (National Center for Supercomputing Applications) PSP (Private Sector Program) Site (8 pp.), January 30, 1998. Available: http://archive.ncsa.uiuc.edu/Indices/Partnerships/whitepapers/iapaper/iapaper.txt.

Townley, C. T. (2001). Knowledge management and academic libraries. *College and Research Libraries, 62*(1), 44–55.

Tribus, M., & McIrvine, E. C. (1971). Energy and information. *Scientific American, 225*(March 1971), 179–188.

Trybula, W. J. (1997). Data mining and knowledge discovery. In M. E. Williams (Ed.), *Annual review of information science and technology* (vol. 32, pp. 197–229). Medford, NJ: Information Today.

Tsai, B. (1992). Building parallel subject knowledge navigation systems for database searching in network environments. In *Proceedings of the 1992 mid-year meeting of the American society for information science* (pp. 83–96). May 27–30, 1992, Albuquerque, NM.

Tsai, B. (1993). Mapping metrics for subject-object coordinated field recognition and representation. *Journal of Library Science with a Slant to Documentation and Information Studies, 30*(3), 91–102.

Tsai, B. (1994a). Formulas for constructing an animation-supported navigation system. *Journal of Library and Information Science, 20*(1), 1–26.

Tsai, B. (1994b). Information business: applying infometry (informational geometry) in cognitive coordination and genetic programming for electronic information packaging and marketing. In *Proceedings of the 1994 American society for information science annual conference* (pp. 54–63). October 17–20, 1994, Alexandria, VA.

Tsai, B. (1995). Detecting the formation orbit of virtual subject knowledge images. *The International Journal of Scientometrics and Informetrics, 1*(3/4), 247–262.

Tsai, B. (1997). Infomapping in information retrieval. *Journal of Information Communication and Library Science, 4*(2), 3–13.

Tsai, B. (1999). Assessing a computer competency program focusing on the operational trends in program activities and the behavioral patterns in student learning. In *Proceedings of the 1999 conference of the international society for scientometrics and informetrics* (pp. 498–512). July 5–8, University of Colima, Colima, Mexico.

Tsai, B. (2000a). Infomapping in information retrieval. In T. K. Srikantaiah, & M. E. D. Koenig (Eds.), *Knowledge management for the information professional* (pp. 297–318). Medford, NJ: Information Today.

Tsai, B. (2000b). Making a web-based subject expert directory through citation data mining. *Journal of Information, Communication and Library Science, 6*(4), 1–12.

Tsai, B. (2001). Constructing a web-based electronic classroom for distance learning. In M. E. Williams (Ed.), *Proceedings of national online 2001* (pp. 487–504). May 15–17, 2001, New York. Medford, NJ: Information Today.

Tsai, B. (2002). A theory of information genetics: How four subforces generate information and the implications for total quality knowledge management. *Library Trends (special issue on current theory in library and information science), 50*(3), 521–552.

Tufte, E. R. (1983). *The visual display of quantitative information*. Cheshire, CT: Graphics Press, p. 28.

Tuthill, G. S. (1990). *Knowledge engineering: Concepts and practices for knowledge-based systems* (pp. 13, 22). Blue Ridge Summit, PA: TAB Books.

Vail III, E. F. (1999). Knowledge mapping: getting started with knowledge management. *Information Systems Management, 16*(4), 16–23.

Van Den Hoven, J. (2001). Information resource management: foundation for knowledge management. *Information Systems Management, 18*(2), 80–83.

Vazzana, G., Elfrink, J. B., & Duane, P. (2000). A longitudinal study of total quality management processes in business colleges. *Journal of Education for Business, 76*(2), 69–74.

Vickery, B. (1997). Knowledge discovery from databases: an introductory review. *Journal of Documentation, 53*(2), 107–122.

Vickery, B. C., & Vickery, A. (1987). *Information science in theory and practice* (p. xiii). Boston, MA: Butterworths Science Company.

Weir, J. (1998). Data mining: exploring the corporate asset. *Information Systems Management, 15*(4), 68–71.

White, H. D., & McCain, K. W. (1998). Visualizing a discipline: an author co-citation analysis of information science, 1972–1995. *Journal of the American Society for Information Science, 49*(4), 327–355.

Withers, C. W. J. (1999). Reporting, mapping and trusting. *Journal of the History of Science in Society, 90*(3), 497–521.

Wong, P. C. (1999). Visual data mining. *IEEE Computer Graphics and Applications, 19*(5), 2–3.

Wong, P. C. et al. (2000). Data signature and visualization of scientific data sets. *IEEE Computer Graphics and Applications, 20*(2), 12–15.

Zipf, G. K. (1972). *Human behavior and the principle of least effort* (p. 302). New York: Hafner.

Zwies, R. (2000). Observations on the American society for information science summit 2000 meeting: defining information architecture. *Bulletin of the American Society for Information Science* (June/July 2000), 10–12.