

Educational Knowledge Domain Visualizations: Tools to Navigate, Understand, and Internalize the Structure of Scholarly Knowledge and Expertise

Peter A. Hook

Indiana University, SLIS
10th Street & Jordan Avenue
Bloomington, IN 47405, USA
E-mail: pahook@indiana.edu

Katy Börner

Indiana University, SLIS
10th Street & Jordan Avenue
Main Library 019
Bloomington, IN 47405, USA
Phone: (812) 855-3256 Fax: -6166
E-mail: katy@indiana.edu

Chapter 11. Educational Knowledge Domain Visualizations: Tools to Navigate, Understand, and Internalize the Structure of Scholarly Knowledge and Expertise

Peter A. Hook and Katy Börner
Indiana University

[M]ost cognitive scientists believe, learning best begins with a big picture, a schema, a holistic cognitive structure, which should be included in the lesson material—often in the text. If a big picture resides in the text, the designers' task becomes one of emphasizing it. If this big picture does not exist, the designers' task is to develop a big picture and emphasize it[.]

(West, Farmer and Wolff, 1991, p. 58).

1. Introduction

Today, we attempt to access all of humanity's knowledge and expertise using search engines such as Google. This works well for fact retrieval. However, search engines do not enlighten the user as to the inherent structure of the information being searched or give the user feedback as to its completeness. There is no 'up' button. The user is not able to see what dataset was queried, how the entries in a search result set relate to each other or how the retrieved entities relate to the entities that were not retrieved. Effective approaches to information access and management need to take into account the human user's perceptual and cognitive capabilities. Humanity is in true need of better tools to filter, navigate, understand, and utilize (scholarly) knowledge.

This chapter discusses domain maps as an alternative means to organize, navigate, and internalize scholarly knowledge. We first discuss the educational uses of maps and the benefits of information visualization and spatialization for education. Subsequently, we introduce thematic maps, cognitive and concept maps, knowledge domain visualizations, and information spaces employing the metro map metaphor. All four are visual representations of geographic or abstract semantic spaces. Given that our interest is in the access, management, and internalization of scholarly knowledge, knowledge domain visualizations are discussed at greater length. To this end, we discuss how the educational use of knowledge domain visualizations is supported by the semantic network theory of learning. We also discuss some of the elements of good knowledge domain map

design. These are drawn from visual perception principles and the study of human memory, and cognition. The final section projects a potential future of educational knowledge domain visualizations.

2. Educational Usage of Maps

This section discusses the utilization of spatial learning strategies, big picture views, and conceptual maps in educational settings. There has been a long history of spatial learning strategies in the field of education (Holley and Dansereau, 1984). With the development of the semantic network theory of learning, many educational theorists began creating and implementing spatial learning techniques. These spatial representations of knowledge are used as: (1) learning tools, (2) evaluation tools, (3) curriculum and instruction planning tools (via both macro and micro maps) and (4) tools to facilitate cooperative learning (Milam, Santo and Heaton, 2000) and cooperative scientific research or collaboratories (MacEachren, Gahegan and Pike, 2004).

In education, visual representations of the big picture view are applied for diverse reasons. First, they provide a structure or scaffolding that students may use to organize the details of a particular subject. In this fashion, information is better assimilated with the student's existing knowledge and the visualization enhances recall. Second, big picture displays make explicit the connections between conceptual subparts and how they are related to the whole. Third, big picture

representations help to signal to the student which concepts are most important for them to learn (West, Farmer and Wolff, 1991).

One commonly used instantiation of big picture views are concept maps. (See Figure 1 and also the subsequent section on Cognitive and Concept Maps.) A concept map is made up of four core elements: (1) shapes or nodes--representing core elements of a concept, (2) connectors or links between the shapes or nodes, (3) connecting words--that describe how two nodes are related, and (4) patterns--such as a hierarchical or circular ordering of the nodes (Milam, Santo and Heaton, 2000). There are at least nine different categories of connecting words between nodes: subsuming, similarity, quantity, enabling, causal, timing, dissimilarity, equivalence, and categorizing (West, Framer and Wolff, 1991). Concept maps have also been referred to as mind maps, pattern notes, brain patterns, spider maps, networks, semantic maps, semantic networks, and semantic webs (predating and different from the WWW Consortium's creation and promotion of the Semantic Web for the Internet) (Milam, Santo and Heaton, 2000).

Concept maps are used to evaluate a learner's understanding of a concept by having the learner display the key elements of the concept in a spatially structured, interconnected layout. This allows an instructor to detect fundamental inaccuracies of a student's conceptual schema for a particular piece of knowledge and to take steps to remedy the inaccurate schema. As teaching

tools, the concept maps of the teacher or other experts are presented to the student to use as cognitive scaffolding for assimilation of the topic being studied. Presently, it is common to find un-labeled cognitive map worksheets (still containing blank nodes and connecting lines), for students to fill in as they study a particular topic. These worksheets are contained in the pre-printed workbook that accompanies many grade school textbooks (Feather, Snyder and Hesser, 1993). Furthermore, a variety of tools exist that ease the creation of concept maps.

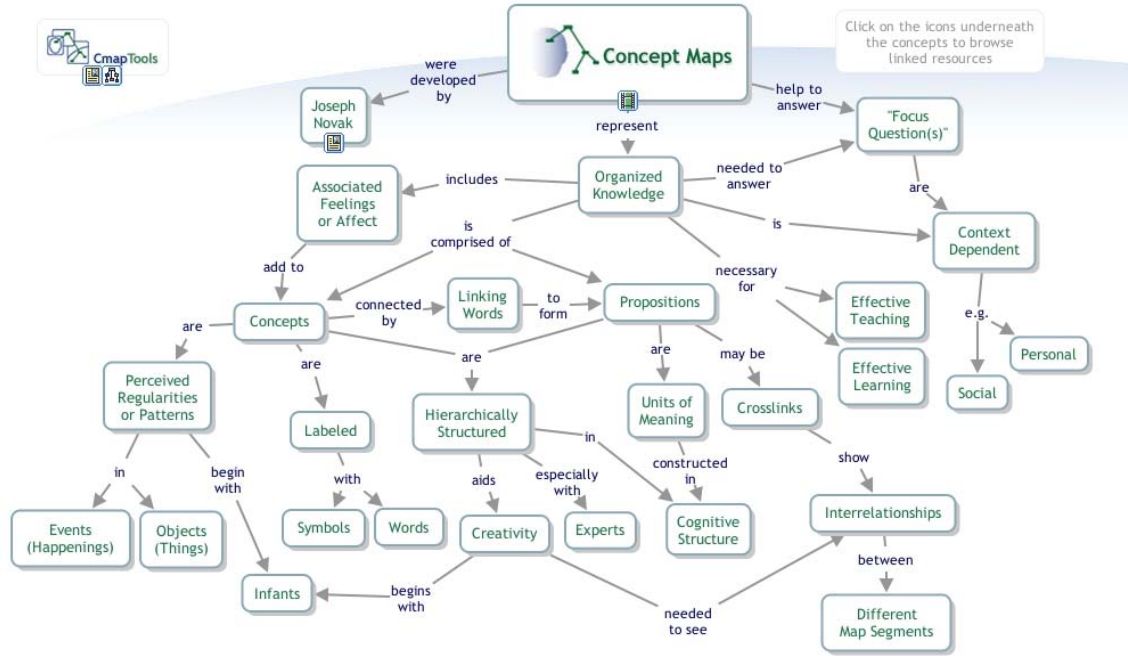


Figure 1. Exemplary Concept Map showing Information About Concept Maps. Created by Joseph Novak and Rendered with CMapTools. Copyright: Institute for Human and Machine Cognition (IHMC). Used with permission.

¹ See: CMap - <http://cmap.ihmc.us/>; Inspiration <http://www.inspiration.com/>; Mindmanager <http://www.mindjet.com/us/>; SMARTIdeas <http://www2.smarttech.com/st/en-US/Products/SMART+Ideas/>; Visimap - <http://www.visimap.com/>

The study of inherently text based subjects may be significantly enhanced by the benefits of information visualization and spatialization.² A concept or domain map responds to the human brain's preference for information that is organized spatially. Information is easier to remember if it can be stored and located spatially (Miller, 1968; Winn, 1994). In fact, the spatial dimension of memory and learning is more important than non-spatial features such as color, shape, or action (Miller, 1968). Additionally, spatial information is easily committed to memory (Allan, 1999).

A domain map of a subject adds a spatial component to a topic that is unavailable in a strictly linear presentation such as a table of contents appearing at the beginning of a textbook or a list of navigable topics in a database. Users of a domain map will have increased recall of the topic by having their memories stimulated by the spatial layout of the domain map. In part, this is due to the fact that recognition is faster and more accurate than recall. "[V]isual memory traces of objects and scenes are stored as part of the processing mechanism; thus it is not necessary for an object to be fully processed for recognition to take place" (Ware, 2004, 299). Additionally, humans store textual and visual information in different areas of the brain. This is known as the 'dual coding' theory of memory (Kulhavy and Stock, 1994; Ware, 2004).

The visual cues inherent in a domain map that the user stores in his or her nonverbal (non-textual) memory region should be highly effective in cuing memory stored in the user's verbal (textual) region:

² Spatialization is the spatial representation of non-spatial data.

“Intact map images retain the spatial characteristics of the objective stimuli, and the information within them becomes simultaneously available when they are brought into working memory. Such images have an advantage as far as the cuing of text retrieval is concerned. The advantage derives from the fact that attention can be shifted from location to location across the map image without exceeding the limits of working memory” (Kulhavy and Stock, 1994, p. 155).

Kulhavy and Stock (1994) ran a series of experiments to test if the associative recall of textual information stimulated by maps was a result of structural features (the spatial layout of the map) and not just non-spatial features such as landmarks, labels, drawings, topographic symbols, and the shape, size and color of depicted items. They concluded that the spatial arrangement of map content was a crucial element in cuing the textual information.

Newbern, Dansereau and Patterson (1997) confirmed Kulhavy and Stock’s co-joining of spatial and verbal memory hypothesis. The authors found that students learning a concept from a knowledge map (in this case a text heavy flow diagram of biological processes) had better recall of the concept than students who learned the same concept from a conventional textual write-up. Furthermore, students were better able to pinpoint where information was located on the knowledge map than with the textual write-up of the subject.

Furthermore, remembering that something is located in a specific place, and what that something is are two different memory tasks. However, they complement each other. When object and location become linked, they serve as memory cues for one another. The object will trigger recall of the spatial location, and the spatial location will trigger recall of the object (Allen, 1999). For all of these reasons, domain maps introduce important spatial and visual elements to subjects that are most frequently presented in a text-heavy manner. It is these spatial and visual elements that amplify cognition and enhance learning. In sum, spatial representations of knowledge appear to improve access to, retrieval, and management of knowledge. Research findings that point to the educational use of spatial learning strategies, big picture views, and conceptual maps should help improve the design and utilization of educational knowledge domain visualizations explained in the next sections.

3. Spatial/Visual Representations of Knowledge

Cartography has long used spatial representations of the planet as a substrate for additional variables. Maps imposing one or more variables onto a literal spatial substrate are referred to as thematic maps. Spatial representations of knowledge have also been used in the fields of wayfinding, psychology, and education to model an individual's understanding of particular concepts. These are known as cognitive maps and concept maps. More recently, the field of information science has begun to represent entire domains of knowledge using

spatial representation techniques. These are known as knowledge domain visualizations (KDV) (Börner, Chen and Boyack, 2003).

4. Thematic Maps

Cartography is the science dealing with the representation and transformation of spatial information (MacEachren, 1995; Skupin, 2004). There are two major types of maps: (1) general purpose or reference maps which display numerous features of a landscape and emphasize location, and (2) thematic maps, which display the geographical characteristics of select statistical phenomena and often focus on a single theme (Slocum, 1998). Thematic maps are designed to convey the structural characteristics of geographic distributions (Chen, 2003). “[T]hey attempt to show the spatial distribution of one or a few variables on the earth’s surface, variables that may not be directly perceptible in the environment at all (e.g., disease rates)” (Montello, Hegarty, Richardson and Waller, 2004, p. 255). Thus, the major design elements of thematic maps are a geographical, or spatial, substrate and an overlay of additional information. As will be discussed later, these same elements are present in most knowledge domain visualizations.

Figure 2 shows a typical thematic map. It depicts the results of the 2004 United States presidential election. The size of the less populous states won by George W. Bush might give the impression that the election was a lopsided victory in Bush’s favor. Figure 3 shows a cartogram of the same election results. A cartogram is a thematic map in which the spatial substrate has been intentionally

distorted to better reveal the characteristics of the information overlay (Slocum, 1998). In this case, the cartogram in Figure 3 provides a better indication of the competitiveness of the 2004 presidential race. It reflects the relative closeness of both the popular vote and the vote in the Electoral College. Such manipulation of the underlying spatial substrate to showcase desired relationships is also common in the production of domain maps.

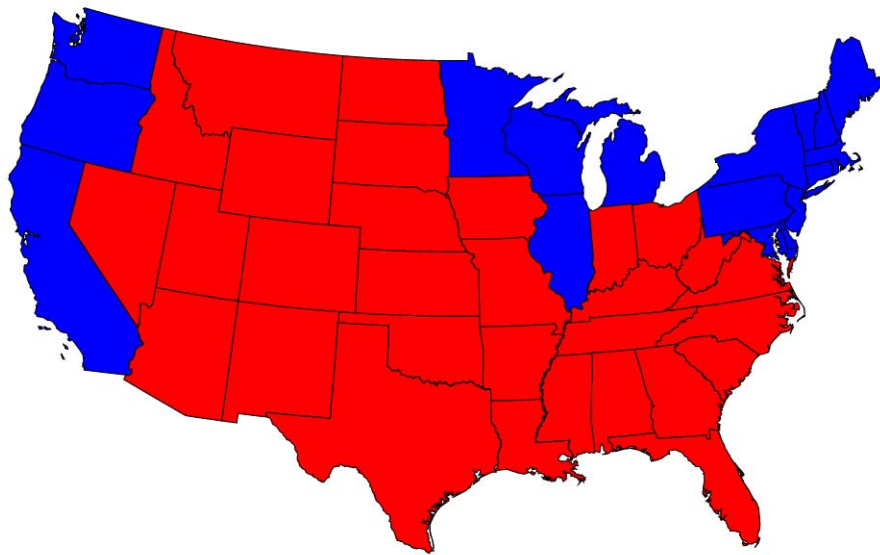


Figure 2. A Typical Thematic Map Displaying the Results of the 2004 Presidential Election. Red States Won By George W. Bush. Blue States Won By John Kerry. Created by: M. Gastner, C. Shalizi, and M. Newman. Permission to reproduce given on Website: <http://www.cscs.umich.edu/~crshalizi/election/>

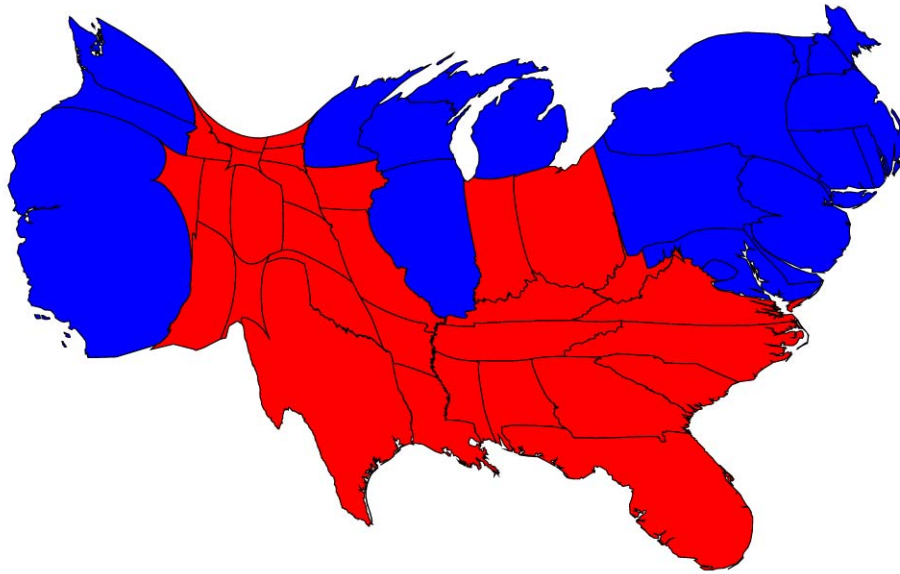


Figure 3. A Cartogram of the 2004 Presidential Election With the Geo-Spatial Substrate Distorted to Reflect the Population of Each State. Red States Won By George W. Bush. Blue States Won By John Kerry. Created by: M. Gastner, C. Shalizi, and M. Newman. Permission to reproduce given on Website: <http://www.cscs.umich.edu/~crshalizi/election/>.

5. Cognitive Maps and Concept Maps

Cognitive maps refer to the representation of spatial information inside the mind of a particular subject (Golledge, 1999). They are used in wayfinding, the science of how people navigate spatial environments. Cognitive maps used in wayfinding employ most of the graphic elements of concept maps used in the field of education. (1) Shapes or nodes are referred to as points. In wayfinding, points can be either landmarks or reference nodes. (2) Connectors or links between the shapes or nodes are lines (routes, paths, or tracks). (3) Connecting words used

in educational concept maps do not have an exact parallel in wayfinding or cartography. In wayfinding and cartography, the connecting relationships between nodes are usually implicit in the spatial arrangement of the map or are labels given to routes between nodes. The latter case includes such information as the name and surface features of the route. However, the labels almost never explicitly state notions of causation or other relationships. (4) Patterns are also present in the cognitive maps of wayfinding. Patterns are conceptualized as “regions, neighborhoods, and topological containment or inclusion,” and these patterns are also detectable in three dimensional notions such as density and changes in elevation (Golledge, 1999, p. 15). “Places and locations form spatial distributions; tracks, paths, and roadways form networks; landmarks and nodes form hierarchies. All of these combine to represent the total knowledge structure (cognitive map).” (Gollege, 1999, p. 20).

As introduced previously, concept maps are a representational tool used in the fields of education and psychology. They represent non-spatial, abstract concepts using spatial representations, see Figure 1 and 4. In wayfinding, the user’s internal cognitive map models actual spatial reality. Education and psychology use the spatially arranged, networked nodes of concept maps to represent ideas or processes that are not inherently manifest in three dimensional space. While there are subtle differences between cognitive maps, concept maps, and all of the other similar types of maps used in the fields of education, psychology and wayfinding, they will henceforth be collectively

referred to as 'concept maps'—internal spatial representations of external reality or ideas.

Figure 4 portrays a concept map with an underlying substrate that was created with procedural rigor. It reflects the consensus of the concept maps created by 13 graduate students in an education course in which the subject matter was integrating technology into the classroom (Kealy, 2001). It is unique in that it was created using the input of an entire class. It is typical in that it has all four elements common to concept maps used in the field of education: (1) nodes, (2) links, (3) connecting words that describe how the nodes are related, and (4) patterns (in this case a spatial ordering of the nodes suggesting their semantic proximity.)

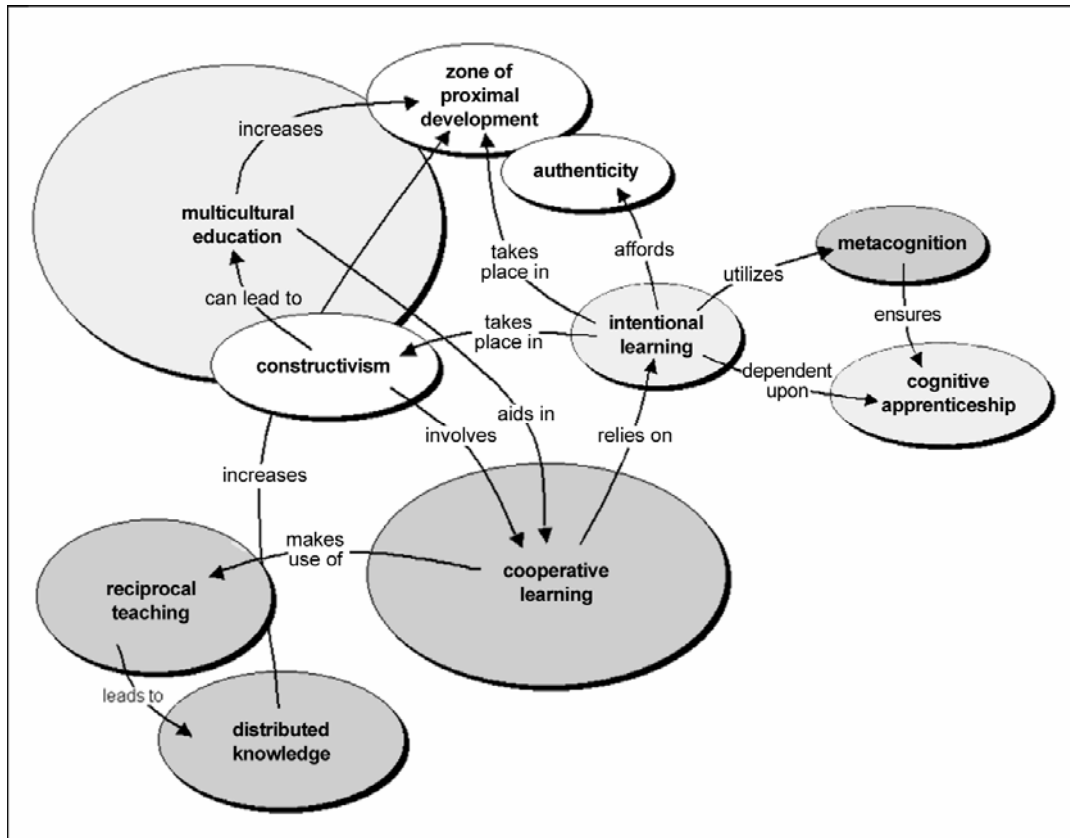


Figure 4. Exemplary Concept Map Showing the Subject Matter of a Graduate Level Education Course on Integrating Technology into the Classroom (Kealy, 2001, p. 345.) Copyright: Baywood Publishing Co., Inc. Used with Permission.

The concept map in Figure 4 is one of the very few known to the authors from the field of education that employs rigorous methods to derive both the spatial layout of the map and the characteristics of the nodes. (In contrast, such production techniques are commonly used in the creation of knowledge domain visualizations (see below)). The spatial layout, or substrate, of the concept map in Figure 4 was created using multi-dimensional scaling ('MDS') techniques (see generally Kruskal and Wish (1984)). These techniques also revealed that the two most important dimensions for the concept map were importance (important vs.

unimportant) and utility (theoretical vs. practical). Surveys of the class participants led to the creation of the nodes with size representing importance and hue representing utility.

6. Knowledge Domain Visualizations

Knowledge domain visualizations (KDV) (Börner, Chen and Boyack, 2003) are the graphic rendering of bibliometric data designed to provide a global view of a particular domain, the structural details of a domain, the salient characteristics of a domain (its dynamics, most cited authors or papers, bursting concepts, etc.) or all three, see Figures 5 and 6. KDV are also referred to as domain maps and the process of their creation as domain mapping.

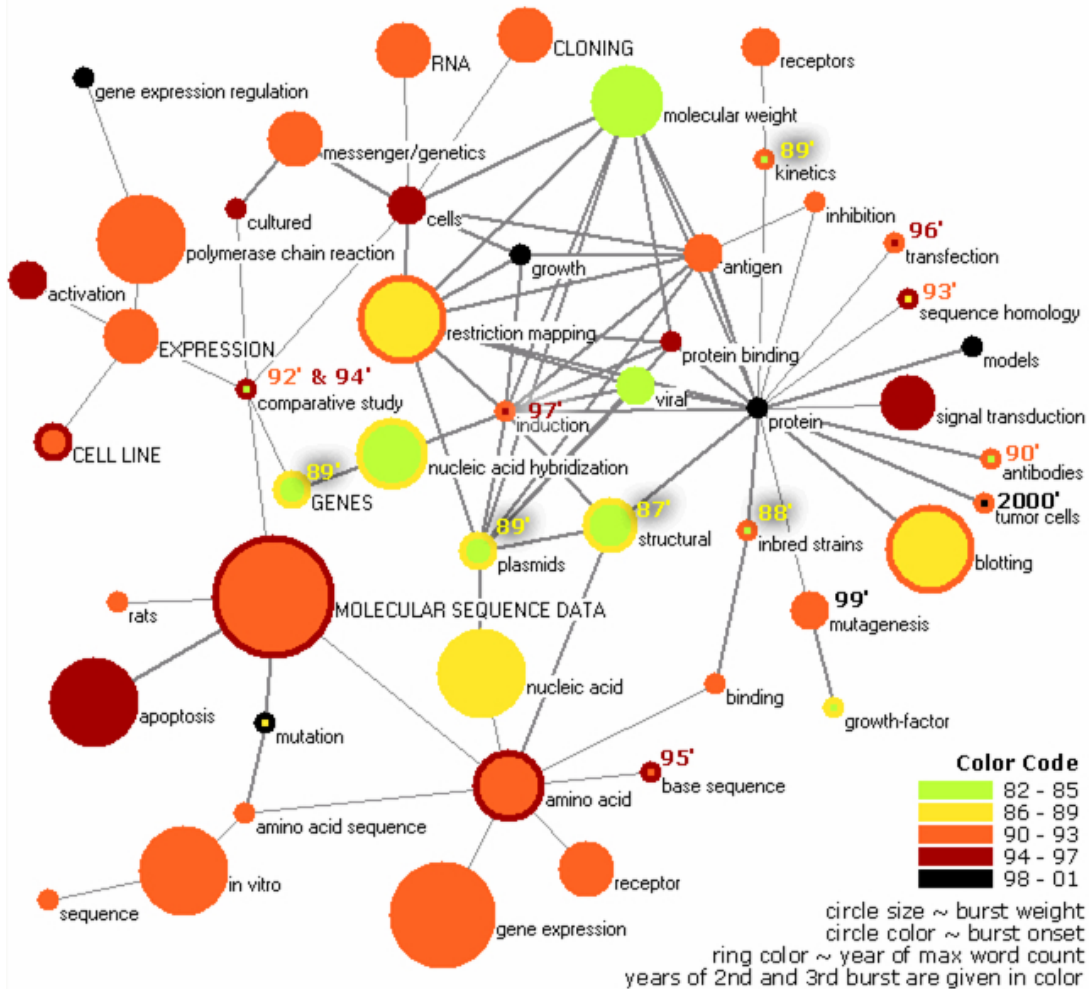


Figure 5. Exemplary Knowledge Domain Visualization Showing a Node Link Diagram of Keywords Appearing in PNAS Between 1982 and 2001. It Portrays the 50 Most Frequent and Bursty Words Used in the Top 10% Most Highly Cited PNAS Publications (Mane and Börner, 2004). Copyright: The National Academy of Sciences of the USA. Used with Permission.

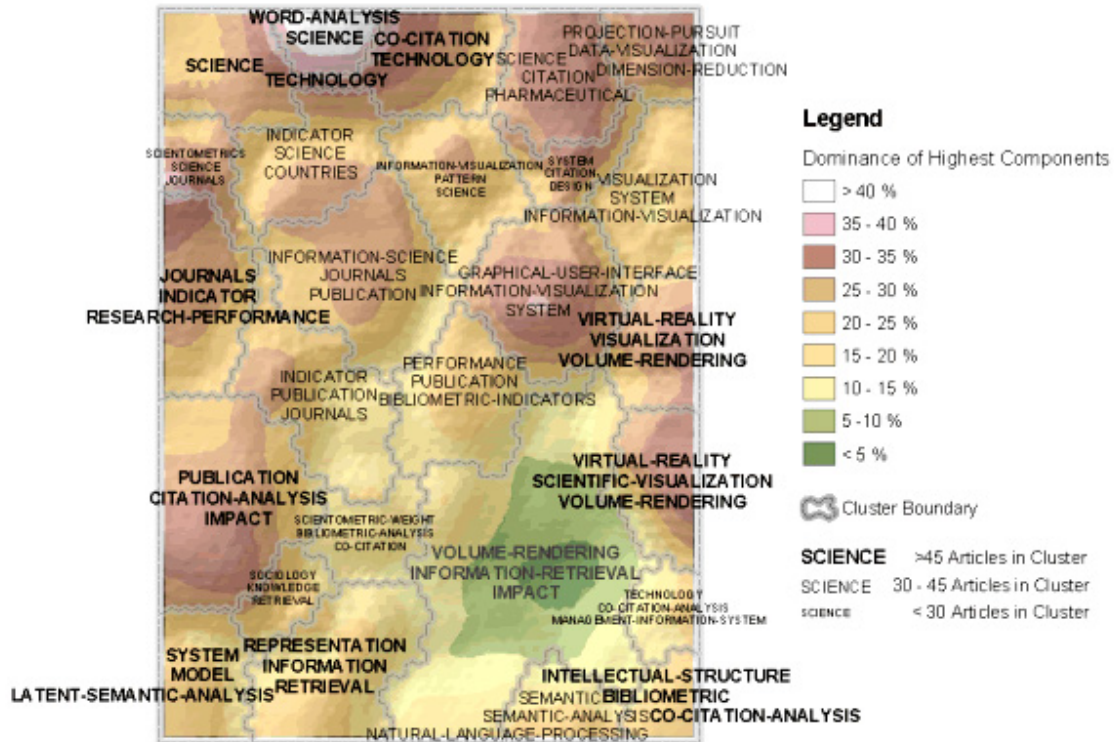


Figure 6. Another Exemplary Knowledge Domain Visualization Showing a Geospatial Rendering of Semantic Space Relevant to Articles Related to Knowledge Domain Visualizations (Börner, Chen and Boyack, 2003). Created by André Skupin. Copyright: American Society for Information Science and Technology. Used with Permission.

Research on KDV is conducted in the fields of information science (and its related sub-disciplines of scientometrics and bibliometrics) (Bernal, 1939; Börner Chen and Boyack, 2003; Braam, Moed and Van Raan, 1991; Callon, Law and Rip, 1986; De Solla Price, 1965; Garfield, Sher, Torpie and Torpie 1964; ISI, 1981; Marshakova, 1973; McCain, 1990; Small, 1973, 1999; White and McCain, 1998). Information science investigates the origination, dissemination, collection, organization, storage, retrieval, interpretation, and use of information (Rubin, 1998). Scientometrics aims at the quantitative study of science. Bibliometrics

studies scholarly production based on bibliographic data (Buter, Noyons and Van Raan, 2004). Consequently, KDVs generated in these three different areas serve different information needs and have a different appearance.

Today, KDVs are typically generated by KDV experts in close collaboration with domain experts using a rather time consuming, laborious process that requires extensive knowledge in terms of data sampling, data analysis, layout, interaction design and last but not least, the interpretation of the resulting maps. Few companies and institutions have the resources to pay for customized maps that reveal the key information (major experts, papers, inventions, emergence of new research frontiers) for their area of interest. This is unfortunate as such knowledge domain visualizations can help to analyze research productivity and lead to more informed decisions as to the allocation of scarce resources.

KDVs are seldom used in educational settings today. This is most likely due to the expense of their generation. In addition, many KDVs are hard to read for non-experts and frequently need to be translated into the jargon of the domain to be truly useful. This is unfortunate in light of the compelling evidence of the utility of maps stemming from the fields of cognitive science, educational psychology, cartography, citation analysis, scientometrics, complex systems, network studies, and bibliometrics. Well designed KDVs have the ability to facilitate understanding, recall, and to convey to the user the schematic, geo-spatial, temporal, semantic, or social organization of the underlying domain.

KDVs differ from concept maps currently used in the field of education in the scope of the domain being represented. In education, the maps are micro thematic—most often dealing with a single theme or concept within a field. In contrast, information scientists often create large scale knowledge domains focusing on entire fields or all of science.

Also, the educational community places more emphasis on the importance of drawing and labeling the connections between nodes with action phrases such as “gives rise to,” and “causes.” These explicitly labeled connections, linkages, or notion of causality are largely missing from KDVs. While suggested implicitly by the spatial proximity of related topics, such issues as causality would be hard to make explicit using automated techniques currently used by information scientists. Garfield's Histograms (Garfield, Sher, Torpie and Torpie, 1964) of the discovery of the structure of DNA have these explicit, temporally labeled connections. However, the use of labeled connections does not rise to the level of use in concept maps used in education.

7. Metro Map Metaphor

Henry Beck's 1933 redesign of the London Underground (subway) map was a revolutionary advancement in graphic design that has recently begun to be applied to non-geographic, information spaces (Nesbitt, 2004). Beck did three things. (1) He removed most of the surface features from the map, thereby

reducing clutter. (2) He decoupled the map from the need to show accurate scale distances between all stations. In other words, he distorted the spatial substrate to give the center, denser portion of the map more area so that its features could be disambiguated and made clearer to the viewer. Also, stations were portrayed at equal distances from one another, when in reality this was not the case. (3) Finally, he unraveled the sinuous nature of the map and made all routes orthogonal with only forty-five or ninety degree angles (with tiny rounded elbows) (Garland, 1994). These easy to comprehend features soon became common on metro maps used throughout the world, see (Ovenden, 2003).

Figure 7 shows the metro map metaphor applied to a non-geographic, information space. It portrays the interconnecting lines and nodes of thought for Keith Nesbitt's doctoral dissertation. The metro map metaphor has the potential to merge domain maps used in the field of education (concept maps) with those coming out of the field of information science (knowledge domain visualizations).

The metro map metaphor contains most of the four elements of concept maps: (1) nodes--representing core elements of a concept (the stations), (2) connectors or links between the shapes or nodes (the route lines) and (3) patterns--such as a hierarchical or circular ordering of the nodes. The missing element, explicit labels describing the connections between nodes, may also be added. Already, the route lines demark subject groupings and their interconnections. In fact, the metro map metaphor allows for more complex, intertwined interactions among

subject groupings than is possible with traditional groupings of bounded regions. Additionally, the layout of the nodes, route lines, and perhaps even clearly demarcated spatial regions may be rigorously performed using the same information science techniques that produce knowledge domain visualizations.

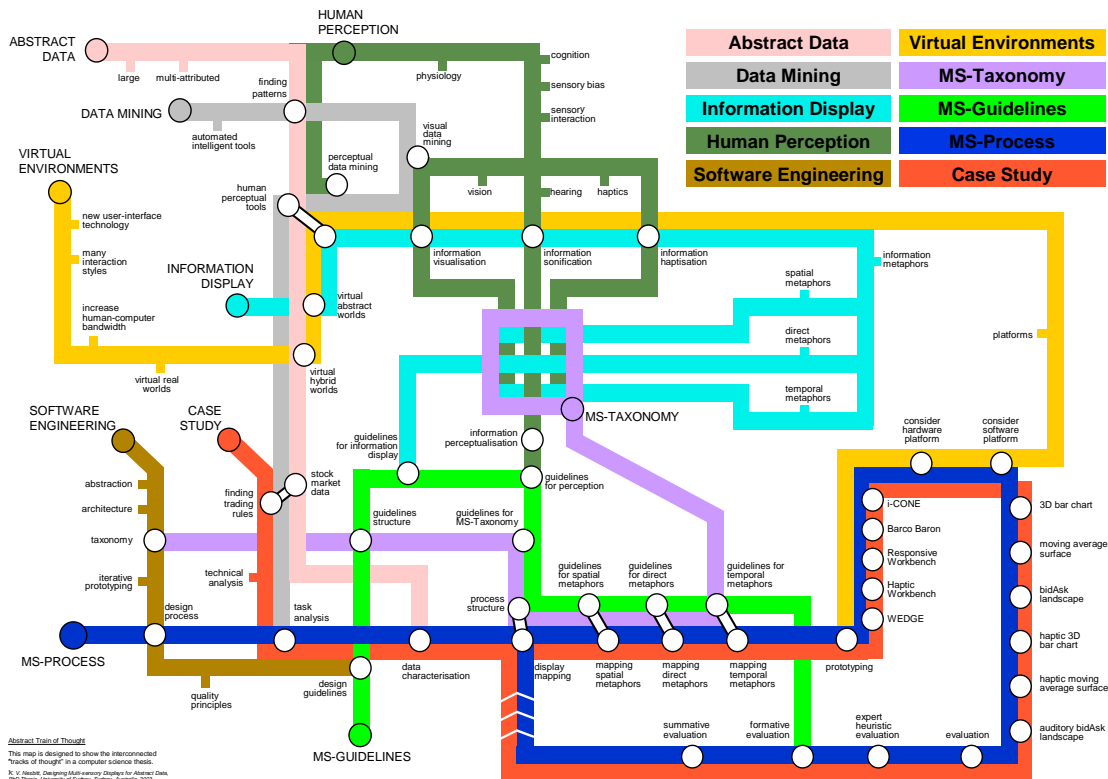


Figure 7. Exemplary Knowledge Domain Visualization Utilizing the Metro Map Metaphor. It Shows the Interconnecting Lines and Nodes of Thought for Keith Nesbitt’s Doctoral Dissertation (Nesbitt, 2004). Copyright: IEEE. Used with Permission.

8. Summary of Substrate/Overlay Information Spaces

Table 1 shows how KDV’s compare with thematic maps, concept maps, and information spaces employing the metro map metaphor. All four map types are composed of two parts: a geographic or base map, and a thematic overlay.

Thematic maps and KDV's use numerous techniques to structure the spatial component of the domain map. While concept maps have a spatial substrate, they are seldom created with rigorous methods. While none currently exist that we are aware of, metro map information spaces have the potential to be created with rigor as to the spatial substrate. Also, while thematic maps typically use a rigorously ordered spatial substrate, they rarely use node-link diagrams. On the other hand, nodes and links are a crucial component of most KDV's, and all concept maps and metro map information spaces. On top of the spatial substrate employed by all four types of maps are placed the thematic components of the map. In the case of knowledge domain visualizations these might be topic words, authors, specific works, contributing institutions, etc.

Map Type	Substrate			Overlay
	Usage of nodes & edges	Rigorous spatial layout	Labeled connections (causality, equivalence, similarity, etc.)	
Thematic Maps	rarely	yes	no	yes
Concept Maps	yes	very rare	yes	yes
Metro Map Information Spaces	yes	potentially	no	yes
Knowledge Domain Visualizations	yes	yes	rarely	yes

Table 1. Comparison of Thematic Maps, Concept Maps and Knowledge Domain Visualizations

Recent efforts aim to create tools that will ease the generation of highly usable maps of the sciences (Borner, Chen and Boyack, 2003). We believe that KDV's

can benefit enormously from user studies and case studies conducted for other types of maps.

9. KDVs Supported by the Semantic Network Theory of Learning

According to the **semantic network theory of learning**, human memory is organized into networks consisting of interlinked nodes. Nodes are concepts or individual words. The interlinking of nodes forms knowledge structures or schemas. Learning is the process of building new knowledge structures by acquiring new nodes. These new nodes are interrelated with existing nodes and with each other. When learners form links between new and existing knowledge, the new knowledge is integrated and comprehended (Jonassen, 1993). In other words, “[w]e learn new concepts by associating them with familiar ones” (Chen, 2003, p. 69).

Learning is the process of reconciling new phenomenon with existing frameworks of understanding. If reconciliation is not possible, then the existing framework(s) must be modified to accommodate the new knowledge. KDVs are an effort to explicitly convey the underlying structure of a domain to the user so the user can internalize the framework presented in the KDV and reconcile it with his or her existing framework. A good understanding of the structural organization of a domain is a better predictor of being able to problem solve in that domain than aptitude as measured by standardized test scores (Jonassen, 1993). Research has also shown that what separates expert and novice problem solvers is the

well-developed and interconnected knowledge networks of the experts that facilitate both the interpretation and solution of the problem (Jonassen, 1993). KDVs should assist in providing the structural knowledge of the domain that will allow novices to become better problem solvers.

Any **expert** attempting to convey information, will instinctively structure the information based on his or her knowledge schema for that particular domain. Readers who are able to discern the author's schema are better able to remember and internalize the information than those who do not recognize the author's organizational framework. Studies also show that learners who are regarded as having good reading comprehension instinctively search out the author's organizational structure and use it to recall content (Jonassen, 1993). Experts in a domain have fewer problems making spatial representations of the domain than non-experts (Breuker, 1984). It is thought that experts possess structurally unifying notions of the domain. It is these notions that can be brought out in a domain map and conveyed to a user.

A **novice user** benefits from a domain map as it expressly conveys the organizational structure of the domain. The user does not have to piece it together from such things as chapter headings, subheadings, and non-content signal words in the text that convey the relationship between ideas (for example, therefore, next, in contrast, on the other hand, etc.). Domain maps provide the user with the necessary visual scaffolding to assimilate and comprehend the

domain. A structural understanding of a domain is also the basis of a deeper, more conceptual, and abstract understanding of a domain (Jonassen, 1993). Furthermore, structured items are better recalled than unstructured items (Jonassen, 1993).

Hopefully, KDVs will become as central to learning as the periodic table of the elements is to chemistry, planetary maps are to geographers, star charts are to astronomers, and city maps are to out-of-town visitors. Like the periodic table, KDVs provide the scaffolding upon which a learner may organize and build new understanding. Analogous to the periodic table, a KDV's transmission of the underlying structure of the domain should also have predictive qualities. By seeing where a topic is placed on a KDV, the user may draw from his or her store of existing knowledge about adjacent topics to begin to understand what an unknown topic is about.

Additionally, much has been written about the existence of multiple learner types (visual, aural, tactile, etc.) (Hook, 2002). Most likely, KDVs will be most effective with visual learners. However, even if KDVs appeal more strongly to visual learners, it is good to add another beneficial tool to the overall package of teaching tools so that collectively they may be used to reach all learners.

10. The Design of Effective Knowledge Domain Visualizations

The design of effective KDVs should be based on a deep understanding of visual perception principles and human cognitive abilities. The utilization of KDVs in educational settings requires knowledge of how maps can be used to store, access, manage, and communicate information. Figure 6 sketches the process of domain map generation and usage. Map generation can be seen as a representational process aimed to convey the structure and dynamics of a knowledge domain.

The knowledge domain exists independent of the observer. Bibliometric data on the domain is harvested, cleaned, analyzed and displayed by the KDV expert. Once the domain map is created, it is presented to the viewer (typically a domain expert) for interpretation. The viewer first perceives the map using low level visual perception (Palmer, 1999; Ware, 2004). Next, the viewer employs higher level cognitive processes to internalize and understand the domain map by reconciling the new content with the viewer's previous knowledge.

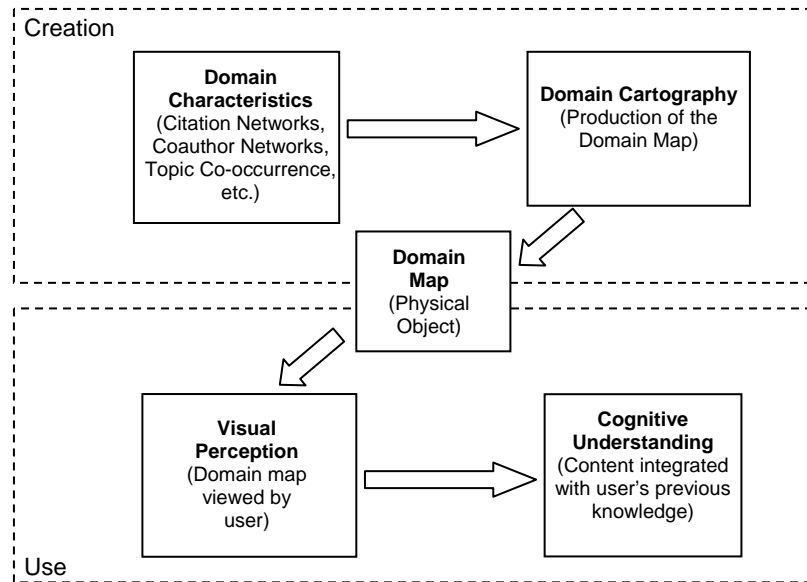


Figure 8. The Process of KDV Creation and Usage

Numerous articles detail the process of creating K DVs (Börner, 2003; Chen, 2003). Subsequently, we discuss some of the perceptual and cognitive issues involving K DVs. Our ultimate goal is the identification of visual perception and cognitive principles that inform the generation of K DVs that truly amplify cognition.

It is interesting to note that the validation of large scale maps of science is a serious problem. Domain *experts* have a very specialized and often subjective view of their area of expertise. Confronted with a map of all of biology they need to zoom 10 to 15 times (depending on their level of expertise) until they recognize papers or peoples' names with which they are familiar. We believe K DVs can play an important role in enhancing and fleshing out the concept maps of expert users by reminding them of implicit knowledge and pointing out related

research areas. Also, KDVs provide a cognitive structure that will influence and potentially rearrange a *novice* user's internal representation of the domain, see

Figure 9.

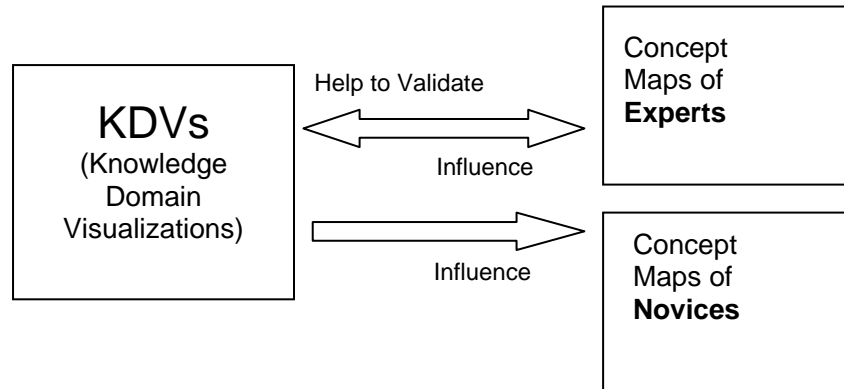


Figure 9. Interaction Between KDVs and Concept Maps of Experts and Novices

Subsequently, we review major visual perception principles and human cognitive abilities that might help guide the design of effective KDVs and their utilization for educational purposes.

11. Visual Perception Principles

A number of excellent textbooks exist on visual perception principles (Palmer, 1999; Ware, 2004). Here we discuss principles that are directly related to the design of highly informative and readable KDVs. Several of them are related to the perception of space and the grouping of objects in space.

Spatial proximity as an expression of semantic closeness is one of, if not the central, metaphors used for the design of KDVs. User studies have shown that subjects are able to equate distance with similarity. This is also known as the “distance-similarity metaphor” (Montello, Fabrikant, Ruocco and Middleton 2003). This ability appears to be independent of the user’s background or familiarity with spatial data. In a series of experiments, Skupin and Fabrikant (2003) have shown that users are able to associate “(1) interpoint distance with the concept of document similarity in a document collection; (2) graphic clusters representing the information content and structure of a digital collection with concentration of related documents; and (3) graphical change in resolution (zoom-in) with different levels of detail in a document collection (hierarchical order)” (p. 110). Spatial proximity and grouping is **processed preattentively** and hence rather fast (Ware, 2004).

Optical illusions. The perception of distances can be distorted by **vertical illusion** effects. Montello, Fabrikant, Ruocco and Middleton (2003) conducted studies in the context of document spaces and showed that for two items of equal distance from a referent third item, the item oriented along the vertical access from the referent item will be perceived as further away than the item oriented along the horizontal access from the referent item. Furthermore, the perception of spatial distance is affected by intervening items that lie between and adjacent to the items being compared for proximity. These are known as **emergent feature effects** (Montello, Fabrikant, Ruocco and Middleton, 2003).

For instance, once three or more items lie in a linear manner between Item A and referent Item C, Item A will be perceived as more proximate to the referent Item C than Item B even though Item B is closer in distance but has no intervening linear points.

Clustering is another visual perception phenomenon that impacts the perception of proximity. Given two items A and B, Item A will be perceived as closer to a referent Item C, if A and C are in the same cluster. This is true even if Item B is actually much closer to Item C, but not in the same cluster as A and C.

Gestalt principles can be employed to further support the perception of semantically related articles, authors, journals, keywords, domains of sciences, etc. See generally (Koehler, 1947; Moore and Fitz, 1993). For example, items that have a similar appearance tend to be perceived as a unit. For this reason, different types of objects depicted in a KDV should be distinguished by different shapes. It should be noted that items with the same shape are more readily perceived as belonging to the same category as items of the same color (Ware, 2004).

Connectedness. Connections can be “a more powerful grouping principle than proximity, color, size or shape” (Ware, 2004, p. 191). Linear connections (links) and separations (boundaries) help to offset the unwanted perceptual effects set out above. Also, they can help with overcoming the inevitable distortions that

arise in the automated conversion of high dimensional data into a low dimensional spatial layout (Skupin and Fabrikant, 2003). The educational psychology community has long felt that connecting lines are a crucial element in representing knowledge to a viewer. While difficult to derive using automated means, connecting elements are a quick and powerful way to convey information about a domain. Whenever possible, they should be included in a KDV. This is true even if they have to be added by hand based on the knowledge of domain experts. The addition of connecting elements derived from expert observation and not from bibliometric data is consistent with Buter's appeal for the use of both quantitative and qualitative bibliographic maps in understanding knowledge domains (Buter, 2004).

Visual or Organizational Hierarchy. Hierarchies are a crucial component of how humans process, remember, and utilize information. Likewise, the hierarchical structuring of information on maps is equally important. Not every piece of information on a map is of equal significance. Graphic techniques are able to signal the most important concepts of a map, or those elements that should be visually processed first. The most important items should be rendered with the greatest contrast to their surroundings. Less important items should be rendered with less contrast (Chen, 2003). Additionally, all seven of Bertin's visual variables (position, shape, orientation, color, texture, value and size) may be used to reveal the data hierarchy of the knowledge domain visualization (Bertin, 1983). Label sizes (letter size, the use of capital and lower case letters, and the

size of the map symbols) are good tools to convey the visual hierarchy.

Graduated color choices and differing line widths demarking boundaries are another way to convey the hierarchical structure of a map (Skupin, 2004).

Distinguishable Regions or Clusters. Studies show that good map learners divide a map into regions and focus their attention on the regions that they have yet to understand and internalize (Allen, 1999; Thorndyke, 1981; Thorndyke & Hayes-Roth, 1982). Consequently, different regions of the map should be well bounded, easily demarcated, readily distinguishable, and memorable.

Landmarks. Both cartography and the field of wayfinding have established the importance of landmarks in understanding maps. Landmarks help lighten the cognitive load on the user. They help orientate the user while viewing/ experiencing the spatial representation of a map. “Landmarks are often noticed and remembered because of dominance of visible form, peculiarity of shape or structure, or because of sociocultural significance” (Golledge, 1999, p. 17).

The significance of a landmark may be either objective or subjective. Some objects are identified as landmarks by most members of a group. Other landmarks have significance because of individual experiences with the object. Landmarks are analogous to important nodes in the schematic network theory of learning with which new concepts are associated. Landmarks provide anchor points from which the rest of the information on the map is organized. Landmarks

act as “primary organizing features in cognitive maps by dominating a spatial classification or clustering process to facilitate environmental knowing and understanding” (Golledge, 1999, p. 17).

Major landmarks on domain maps should be readily distinguishable from other nodes. It is not enough that they be labeled with bolder text, represented by larger dots, extend higher into space, or otherwise have a larger shape. They should have an idiosyncratic component. They should be unique and preferably represented by an image. The rationale for this may be drawn from Winn and Holliday’s principle number four for the design of diagrams and charts: “Include small pictures in diagrams to teach concept identification, especially with students of low verbal ability” (Winn and Holliday, 1982, p. 286). Visually prominent landmarks will also facilitate recall by being easy to find again. They will also serve as a mnemonic for better remembering the information on the map.

Colors. Color is often the best variable to use to code data categories (Ware, 2004). Color coding makes it easy for a viewer to categorize different objects (Ware, 2004), and presumably different concepts. Color may be used to differentiate bounded regions and contribute to the overall legibility of a KDV. Color choices also contribute to the viewer’s perception of the information hierarchy of the KDV.

Ware (2004) lists a set of criteria to consider when using color to encode data.

(1) Colors should be distinct. (2) Unique hues should be used (red, green, yellow, blue, black, and white.) (3) Colors should contrast with the background. (4) Color blindness should be taken into account. (5) Only about five to ten color codes may be rapidly distinguished. (6) Larger color coded areas make it easier to perceive contrasts in color coding. Extremely small color-coded areas should not be used as it becomes hard to distinguish between the colors. Furthermore, “[w]hen large areas of color coding are used, for example with map regions, the colors should be of low saturation and differ only slightly from one another. This enables small, vivid color-coded targets to be perceived against background regions” (Ware, 2004, p. 125). Finally, (7) the mapmaker should consider cultural color coding conventions (red = hot, blue = cold).

Ware recommends the use of the following twelve colors to encode data: red, green, yellow, blue, black, white, pink, cyan, gray, orange, brown, and purple. The first six are to be used before the second six (Ware, 2004). The use of black or white borders around colored symbols helps to ensure that the symbols stand out from the background colors (Ware, 2004). Additionally, colors may encode more than one variable. For instance, the hue of a color may represent one variable while the saturation of a color represents a second variable. However, such maps are extremely hard to read and the second variable should be represented using an additional information channel such as texture (Ware, 2004).

12. Cartographic Literacy and Caution

A human's ability to use maps is to some degree a learned behavior. Humans learn universal cartographic symbols and conventions that make understanding an unfamiliar map easier (Allan, 1999). For instance, some small children have difficulty with the concept that a truncated map boarder does not display the literal shape of countries that extend past the map's edge (Arnheim, 1969). In other words, students must learn that the items or concepts mapped have the potential to extend past the edges of the representational image.

Literacy with KDV's must also be cultivated. A user must first become familiar with the metaphor that spatial proximity, while seldom literal, implies proximity between the concepts being mapped. Once understood, the use of the next domain map should be easier. Also, the establishment of universal symbols and conventions should likewise increase the speed in which a domain map is understood. While efforts may be made to enunciate best practices and to establish conventions, these items will most likely emerge over time through the repeated creation and use of domain maps and user testing.³

There are also risks in using KDV's in education. KDV's may convey to the viewer a false sense of established proximities when in fact domains are fluid. In

³ Ware cautions: "standardization is the enemy of innovation and innovation is the enemy of standardization" (Ware, 2004 p. 386).

addition, there may be different paradigms or viewpoints for visualizing a domain. Also, different techniques used to produce KDVs will create different representations of a knowledge domain. In other words, there are no deterministic, exactly reproducible methods for creating KDVs. “Even with clear assumptions and good qualitative research methodology, there are a myriad of ways to create a single type of map of the same content. It is important to either involve a group of scholars in developing a map and/or to recognize that the resulting map is simply one of many possible patterns for documenting the links between complex ideas” (Milam, Santo and Heaton, 2000, p. 5). While Milam is referring to concept maps created without the more rigorous automated techniques employed by information scientists, the warning is still applicable. An educator will have to consider the risks and benefits of using a single KDV or presenting multiple KDVs of a particular domain created with different techniques or that encapsulate different paradigms.

13. Future Directions for Educational KDVs

The authors envision six phases in the use, adoption, and implementation of KDVs. The first two have already occurred. The remaining four are prospective—compelled by the cognitive benefits discussed above.

Phase one: bibliometricians realized that they could use bibliographic datasets and techniques such as author co-occurrence to provide maps of a particular

discipline. These early maps were graphically simple and painstakingly created by hand.

Phase two has been the implementation of automated techniques for data harvesting, processing, and information visualization. This has facilitated the mapping of larger domains.

Phase three: domain maps will become widely known outside of information science (Whitehouse, 2004). They will become popular with educators and will be used to enhance classroom pedagogy. In April 2005, a workshop at the annual conference at the American Association of Geographers brought together cartographers and information scientists to discuss creating better KDVs. See <http://vw.indiana.edu/aag05/>.

Phase four: the widespread use of domain maps will lead to steps that aim to harmonize and better preserve the scholarly data from which they are created (Börner, in press). This might include unique author identification numbers, better and standardized citation practices, and repositories containing information such as all of the sources cited in books.

Phase five: domain maps will routinely be used as one of the access options to digital libraries and online public access systems (OPAC's) (the tools library patrons use to search for materials). Even if a user chooses to do a keyword

search, domain maps will be dynamically displayed unobtrusively in the background, subtly conveying to the user the intellectual landscape of the particular domain and the specific neighborhood of the user's search interest.

Phase six: dynamic domain maps will capture and portray the diffusion of information. This diffusion of knowledge may occur from one author to the next, one journal to another, or among scientific disciplines, etc. Domain maps will have predictive elements that will forecast and model the spread of knowledge. They will be used widely for science forecasts in a similar fashion as today's weather forecast maps.

14. Conclusion

The colossal landscape of scholarly knowledge, growing at exponential rates, now requires representational maps utilizing advanced techniques to provide insight into the structure and dynamics of scholarly domains. Today we need intellectual cartographers to assist students and scholars in navigating, understanding and internalizing the structure and dynamics of scholarly bodies of knowledge. There is compelling evidence of the utility of KDVs stemming from the fields of educational psychology, cognitive science, cartography, and information science. Well designed KDVs have the ability to facilitate understanding, recall, and to convey to the user the schematic, geo-spatial, temporal, semantic, or social organization of the underlying domain.

Though educational knowledge domain visualizations are still in their infancy, we believe that they have a promising future in assisting with access to and the navigation, understanding, management, and communication of large-scale information spaces. Furthermore, when used as an interface for information retrieval, knowledge domain visualizations have the potential to convey the structural organization of the domain to the user. In turn, this structural knowledge of the domain provides the cognitive scaffolding with which the user may associate additional details about the domain.

15. References

- Allen, G. L. (1999). Spatial Abilities, Cognitive Maps, and Wayfinding: Bases for Individual Differences in Spatial Cognition and Behavior. In: R. G. Golledge (Ed.), *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes* pp. 46-80. Baltimore, MD: Johns Hopkins University Press.
- Arnheim, R. (1969). *Visual Thinking*. Berkeley, CA: University of California Press.
- Bernal, J. D. (1939). *The Social Function of Science*. London: Routledge and Kegan Paul Ltd.
- Bertin, J. (1983). *Semiology of Graphics: Diagrams Networks Maps*. Translated from the French by William J. Berg. Madison, Wisconsin: University of Wisconsin Press.

- Börner, K. (in press). Semantic Association Networks: Using Semantic Web Technology to Improve Scholarly Knowledge and Expertise Management. In: V. Geroimenko and C. Chen (Eds.), *Visualizing the Semantic Web*, Springer Verlag, 2nd Edition, Chapter 11.
- Börner, K. Chen, C. and Boyack, K. W. (2003). Visualizing Knowledge Domains. In: B. Cronin (Ed.) *Annual Review of Information Science and Technology* (pp. 179-255). Medford, NJ: Information Today.
- Braam, R. R. Moed, H. F. and Van Raan, A. F.J. (1991) (1). Mapping of Science By Combined Co-Citation and Word Analysis. I. Structural Aspects. *Journal of the American Society for Information Science*, 42(4) 233-251.
- Braam, R. R. Moed, H. F. and Van Raan, A. F.J. (1991) (2). Mapping of Science By Combined Co-Citation and Word Analysis. II. Dynamical Aspects. *Journal of the American Society for Information Science*, 42(4) 252-266.
- Breuker, J. A. (1984). A Theoretical Framework for Spatial Learning Strategies. In: C. D. Holley and D. F. Dansereau (Eds.) (1984). *Spatial Learning Strategies: Techniques, Applications, and Related Issues*. Orlando, FL: Academic Press.
- Buter, R. K. Noyons, E. C. M. and Van Raan A. F. J. (2004). A Combination of Quantitative and Qualitative Maps in an Evaluative Bibliometric Context. *Proceedings of the Eighth International Conference on Information Visualisation (IV'04)*, 1093-9547/04.

- Callon, M. Law, J. and Rip, A. (Eds.) (1986). *Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World*. London: Maxmillan Press Ltd.
- Chen, C. (2003). *Mapping Scientific Frontiers: The Quest for Knowledge Visualization*. New York: Springer.
- De Solla Price, D. J. (1965). Networks of Scientific Papers *Science*, New Series, 149(3683) (Jul. 30, 1965), 510-515.
- Feather, R. M. Jr. Snyder, S. L. and Hesser, D. T. (1993). *Concept Mapping, Workbook to Accompany, Merrill Earth Science*. Lake Forest, Illinois: Glencoe.
- Garfield, E. Sher, I. H. Torpie, R. and Torpie, J. (1964). *The Use of Citation Data in Writing the History of Science*. Philadelphia: Institute for Scientific Information.
- Garland, K. (1994). *Mr. Beck's Underground Map*. London: Capital Transport Publishing.
- Gastner, M. Shalizi, C. and Newman, M. (2004). Maps and Cartograms of the 2004 US Presidential Election Results. Available at:
<http://www.cscs.umich.edu/~crshalizi/election/>
- Golledge, R. G. (1999). Human Wayfinding and Cognitive Maps. In: G. R. Golledge (Ed.) *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes* pp. 5-45. Baltimore, MD: Johns Hopkins University Press.

- Holley, C. D. and Dansereau, D. F. (Eds.) (1984). *Spatial Learning Strategies: Techniques, Applications, and Related Issues*. Orlando, FL: Academic Press.
- Hook, P. A. (2002). Creating an Online Tutorial and Pathfinder. *Law Library Journal*. 92(4), 243-265.
- ISI (Institute for Scientific Information) (1981). *ISI Atlas of Science: Biochemistry and Molecular Biology, 1978/80*, Philadelphia, PA.: Institute for Scientific Information.
- Jonassen, D. H. Veissner, K. and Yacci, M. (1993). *Structural Knowledge: Techniques for Representing Conveying, and Acquiring Structural Knowledge*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kealy, W. A. (2001). Knowledge Maps and Their Use in Computer-Based Collaborative Learning Environments. *Journal of Educational Computing Research*, 25(4) 325-349.
- Koehler, W. (1947). *Gestalt Psychology* (Rev. Ed.). NY: Liveright.
- Kruskal, J. B. and Wish, M. (1984). *Multidimensional Scaling*. London: Sage Publications.
- Kulhavy, R. W. and Stock, W. A. (1994). Reference Maps as a Framework for Remembering Text. In: W. Schnotz and R. W. Kulhavy (Eds.) *Comprehension of Graphics* (pp. 153-162). Amsterdam: North-Holland.
- MacEachren, A. M. (1995). *How Maps Work: Representation, Visualization, and Design*. New York: Guilford Press.

- MacEachren, A. M. Gahegan, M. and Pike, W. (2004). Visualization for Constructing and Sharing Geo-Scientific Concepts. *PNAS*, 101(Suppl_1): 5279–5286.
- Mane, K. and Börner, K. (2004). Mapping Topics and Topic Bursts in PNAS. *PNAS*, 101(Suppl_1): 5287–5290.
- Marshakova, I.V. (1973). A System of Document Connections Based on References. *Scientific and Technical Information Serial of VINITI*, 6, 3-8.
- McCain, K. W. (1990). Mapping Authors in Intellectual Space: A Technical Overview. *Journal of the American Society for Information Science*, 41(6) 433-443.
- Milam, J. H. Jr. Santo, S. A. and Heaton, L. A. (2000). *Concept Maps for Web-Based Applications*. ERIC Technical Report. ERIC No.: ED445627.
- Miller, G. A. (1968). Psychology and Information. *American Documentation*. July, 286-289.
- Montello, D. R., Fabrikant, S. I. Ruocco, M. and Middleton, R. S. (2003). Testing the First Law of Cognitive Geography on Point-Display Spatializations. In: W. Kuhn, M. F. Worboys, and S. Timpf (Eds.) *Spatial Information Theory: Foundations of Geographic Information Science. Proceedings of the Conference on Spatial Information Theory (COSIT '03)*. *Lecture Notes in Computer Science 2825*. Berlin: Springer Verlag 316-331.
- Montello, D. R. Hegarty, M. Richardson, A. E. and Waller, D. (2004). Spatial Memory of Real Environments, Virtual Environments, and Maps. In: G. L.

- Allen (Ed.) *Human Spatial Memory: Remembering Where*. Mahwah, New Jersey: Lawrence Erlbaum Associates 251-285.
- Moore, P. and Fitz, C. (1993). Gestalt Theory and Instructional Design. *Journal of Technical Writing and Communication*, 23(2), 137-157.
- Nesbitt, K. V. (2004). Getting to More Abstract Places using the Metro Map Metaphor. *Proceedings of the Eighth International Conference on Information Visualisation (IV'04)*. Piscataway, NJ: IEEE.
- Newbern, D. Dansereau, D. F. and Patterson, M. E. (1997). Spatial-Semantic Display Processing: The Role of Spatial Structure on Recall. *Contemporary Educational Psychology*, 22, 319-337.
- Palmer, S. E. (1999). *Vision Science: Photons to Phenomenology*. Cambridge, MA: MIT Press.
- Rubin, R. E. (1998). *Foundations of Library and Information Science*. New York: Neal-Schuman Publishers, Inc.
- Skupin, A. (2004). The World of Geography: Visualizing a Knowledge Domain With Cartographic Means. *PNAS*, 101(Suppl_1): 5274–5278.
- Skupin A. and Fabrikant, S. I. (2003). Spatialization Methods: A Cartographic Research Agenda for Non-Geographic Information Visualization. *Cartography and Geographic Information Science*, 30(2), 95-119.
- Slocum, T. A. (1998). *Thematic Cartography and Visualization*. Upper Saddle River, New Jersey: Prentice Hall.

- Small, H. (1973). Co-Citation in the Scientific Literature: A New Measure of the Relationship Between Two Documents. *Journal of the American Society for Information Science*, 24, 265-269.
- Small, H. (1999). Visualizing Science by Citation Mapping. *Journal of the American Society for Information Science*, 50(9), 799-813.
- Thorndyke, P. W. (1981). Distance Estimation From Cognitive Maps. *Cognitive Psychology*, 13, 526-550.
- Thorndyke, P. W. and Hayes-Roth, B. (1982). Differences in Spatial Knowledge Acquired From Maps and Navigation. *Cognitive Psychology*, 14, 560-589.
- Ware, C. (2004). *Information Visualization: Perception for Design*. San Francisco, CA: Morgan Kaufman.
- West, C. K. Farmer, J. A. and Wolff, P. M. (1991). *Instructional Design: Implications from Cognitive Science*. Englewood Cliffs, New Jersey: Prentice Hall.
- White, H. and McCain, K. (1998). Visualizing a Discipline: An Author Co-citation Analysis of Information Science, 1972-1995, *Journal of the American Society for Information Science*, 49(4), 327-356.
- Whitehouse, D. (2004). *Scientists Seek 'Map of Science.'* BBC News. Wednesday, 7 April, 2004. Available at:
<http://news.bbc.co.uk/1/hi/sci/tech/3608385.stm>
- Winn, W. and Holiday, W. (1982). Design Principles for Diagrams and Charts. In: D. H. Jonassen (ED). *The Technology of Text: Principles for Structuring*,

Designing, and Displaying Text. (pp. 277-299). Englewood Cliffs, NJ:
Educational Technology Publications.

Winn, W. (1994). Contributions of Perceptual and Cognitive Processes to the
Comprehension of Graphics. In: W. Schnotz and R. W. Kulhavy (Eds.)
Comprehension of Graphics (pp. 3-27). Amsterdam: North-Holland.