Teachers College Record
The Voice of Scholarship in Education

Network Search: A New Way of Seeing the Education Knowledge Domain

by Daniel McFarland & Eric Klopfer — 2010

Background:

The educational knowledge domain may be understood as a system composed of multiple, co-evolving networks that reflect the form and content of a cultural field. This paper describes the educational knowledge domain as having a community structure (form) based in relations of production (authoring) and consumption (referencing), and a cognitive structure (content) based in relations of ideas and concepts.

Purpose:

We propose developing an online interactive system whereby the vast array of available knowledge artifacts can be mined for information reflective of these networks, and which can be visualized, measured, and explored over time. We argue that the creation of such a system will benefit the field of education manifold: it will greatly enhance individual learning of online materials; create more efficient searches; open access to knowledge creation and consumption to a wider public; enable greater practitioner involvement; facilitate direct study of the educational knowledge domain; and identify various means towards accelerated innovation.

Design:

Building on the ideas of online communities, network visualizations, e-commerce, and advanced search engines, Scholar Practitioner Information Networks for Education (SPINE) not only facilitates access to education information resources, but also allows the community to view multiple sources of information in a relational context. At the heart of SPINE are multiple sources of information (journal articles, case studies, reviews, etc.), reviews of information, and ratings of reviewers. Connecting these sources are advanced analyses of production and consumption, allowing unique insights into established fields.

Conclusion:

We illustrate that such a system is being developed in piecemeal fashion within other disciplines and propose a means by which they can be synthesized into a multi-network, interactive system for the field of education.

INTRODUCTION

Like many academic fields, the literature and resource base in the field of education is growing rapidly. This growth is facilitated by a large number of researchers creating academic papers for journals and by information technologies that provide ready access to conference papers, opinions pieces, and other contributions made by various constituencies in the educational field. Each type of knowledge artifact has its place and purpose, and is typically accessed by various sub-communities (Shriffin & Borner, 2004). Practitioners and researchers currently access different materials, but as demand grows on both sides to be more informed about each other's work, those boundaries will begin to break down, at least at their peripheries (Lagemann, 1989, 2002; Mitchell & Haro, 1999).

Current efforts to catalog educational research of all of these types and to make them accessible have taken the time-tested approach of building up large databases, and then providing users with the opportunity to conduct various keyword searchers that return lists of results. Efforts like ERIC (http://www.eric.ed.gov/) attempt to bring together educational literature published in a variety of forms, while journal databases provide a similar, but more narrow, means of cataloging and relating existing research (e.g., JSTOR). These efforts revolve around a central database and searchable resources.

The combination of seemingly exponential growth of materials comprising these databases, along with the information technologies that make more and more of these artifacts accessible, has created a situation in which the resources at the disposal of researchers and professionals have become so vast that their utility may indeed become diminished rather than

enhanced. Good search criteria can generate a long list of knowledge artifacts relevant to a topic, but the lists do not facilitate the viewer's understanding of a topic, its organization, and how it is developed—i.e., how types of knowledge span artifacts and are interrelated, and how social relations under-gird this body of developing knowledge.

It has come to a point where the education community needs to have at its disposal better tools for visualizing and measuring the connections between artifacts, their contents, and the constituencies that generate, evaluate, and consume them. Such a tool could potentially reorient members' explorations and conceptualizations of ideas, and provide a deeper understanding of the field in which such ideas are embedded. A new researcher to the field of education should be able to understand central pieces of research and practice, the communities based around them, and their current level of activity. Professionals should be able to both contribute to the research community through their experiences on the ground and understand enough of the current research to affect their own practices. Sub-disciplines within the field of education should be able to see how they interrelate with each other, as well as how their work informs educational practice. In addition, researchers utilizing different research methodologies should be able to see how their own work relates to research conducted on the same topic but with different methods. The current search and list technologies do not facilitate inquiries like this, and in many ways, they are counterproductive because they perform narrow searches that return a specific result instead of a view from above that captures a broader picture.

Other academic and non-academic domains face problems of a similar nature, and new technologies have arisen to help ameliorate these issues. From knowledge artifacts (e.g., articles, books, patents, reviews, curricula, etc.), we can visually represent different organizational dimensions, or network arrangements, of the educational knowledge domain. Knowledge artifacts are cultural objects (Griswold, 1987) intimately linked to producers, consumers, and evaluators. Through these roles and activities, participants engage in interactions and processes that are central to the field's generation and organization. Moreover, the contents of knowledge artifacts have their own organization, or cultural structure, which informs our understanding of the field. We can take these readily accessible features of knowledge artifacts and represent them as multiple networks reflecting the *form* and *content* of the educational knowledge domain (Simmel, 1971; Mannheim, 1936): such as networks of interrelated knowledge contents, networks of collaboration (production), and networks of reference and critique (consumption). Hence, we assert that the form and content of the educational field can be empirically represented as a multi-dimensional, networked system (Simmel, 1971; Barabasi, 2003; Carolan & Natriello, 2005).

Many benefits can be derived from representing the educational knowledge domain as a system of multiple, interrelated networks. First, network visualizations exploit human vision and spatial cognition, thereby expanding mental capacities of viewers. Network maps are like concept maps and facilitate individual learning more than lists of unrelated concepts (Novak & Gowin, 1984; Novak, 1998). Upon using concept maps in science curricula, researchers find that students retain more information, learn more about scientific processes, and better understand the system of interrelated ideas and processes than lists of unconnected facts could convey. Previous research also shows that people more effectively learn about complex systems, such as that formed by the community of education research, through systemic views (Simon, 1969; Richards & Barnett, 1993; Chen, 2004; Powell, White, Koput, & Owen-Smith, 2005). We believe that interactive network visualizations of the educational field will increase comprehension of the vast array of findings, actors, and concepts over and above the way rank-ordered lists do.

In addition to increasing knowledge retention, and facilitating comprehension, network representations enable efficient searches of the information space (Shriffin & Borner, 2004). Network linkages reduce visual search time because they show how artifacts are related, and thereby reduce error and "wild goose chases." Network depictions reveal hidden relations normally overlooked in list-representations. In this manner, network representations exploit users' visual capacities and facilitate a more intelligent search of the domain.[1]

Third, networks afford a variety of benefits for all sorts of analysis—quantitative and qualitative—by explicitly representing a set of elements and the relations among them When chosen well, network representations make it easier to discover mistaken formulations of the nature of the education field. By affording users the capacity to generate *multiple* network representations, people are not required to conform to a single understanding of how the educational knowledge domain works. Instead, the choice of views will commit users to substantially different understandings. However, their real virtue is to increase the chance of discovering that presupposed relations of artifacts, terms, and participants are wrong, and therefore enable us to learn something new about the educational field (Tilly, 2006).

Fourth, multiple network representations (and in turn, network data and measures) can provide direct information on the educational knowledge domain, both locally in specific relationships, and more globally across a textured web of relations. The depiction of multiple, interrelated networks will show how different collaboration and consumption patterns influence (and are influenced by) the patterning of educational concepts and arguments. Hence, by interrelating networks we see how relations of production influence patterns of conceptualization and knowledge consumption. Therefore, network data, its representation, and its measurement can illustrate the development of the educational field, generate causal hypotheses, and test which mechanisms generate various forms of change in the knowledge domain (e.g., predicting points of innovation and shifts in

communal structures).

We believe the development of an online system capable of generating multiple local and global network representations (and measures) of the educational field is very realistic and made tangible by recent technological developments and empirical efforts arising in other fields. Moreover, we believe such a system will afford many positive returns to the educational knowledge domain. It will enable more efficient searches of materials; facilitate individual comprehension of the educational field; broaden access and participation in the knowledge domain so as to include more professionals and the wider public; enable direct empirical study of the educational knowledge domain and its core practices; identify regions where greater coordination is needed; and identify those mechanisms accelerating the rate of knowledge innovation.

THE EDUCATIONAL KNOWLEDGE DOMAIN AS A SYSTEM OF MULTIPLE NETWORKS

The educational knowledge domain is centered on knowledge artifacts, such as articles, books, curricula, patents, reviews, audiovisual records, etc.—many of which are already stored online and increasingly so, at an exponential rate. These artifacts are cultural objects, or shared meanings embodied in form (Griswold, 1987) and they enable us to grasp the broader system we refer to as a knowledge domain (or cultural domain) and hold that up for analysis. Knowledge artifacts have certain *contents*, or ideas and narratives (Abell, 2004), which have a decisive influence on the social relations of field participants. In turn, these contents may be understood as developments of a society or social *form* (e.g., social structures) composed of actors engaged in various roles and activities (Nadel, 1957; Simmel, 1971). The educational knowledge domain is centered on knowledge artifacts that concern the topics of an institutional field spanning disciplinary and professional endeavors, and whose formation is contingent on social structural relations shaped by various activities of knowledge production (collaboration and mentoring) and consumption (such as citation and usage). Conversely, the structure of meanings (ideas) and their presented narratives shape the form of social relations in the educational domain (Berger & Luckman, 1966; Crane, 1972). By juxtaposing ideas and aligning them, the shared meanings embodied in artifacts cue social relations reflective of research paradigms, camps, and status hierarchies. This dynamic of *form* and *content* represents a core process of knowledge domain creation and change.[2] A multiplicity of network representations can characterize the form and content of the educational knowledge domain, of which we focus on only a few here (Morris & Yen, 2004):

1. Artifacts entail *knowledge contents* and the patterning of these contents in relation to each other reveals a cognitive structure, concept-space, or empirical rendering of the field's lexicon (and upon abstraction, larger structures reflective of narratives and explanations, see Locke and Golden-Biddle, 1997). These structures are partly defined by co-occurring concepts in the contents of publications, practitioner curricula, and case-reports (Burns, Khan, Handeharizadeh, O'Neill, & Chen, 2003).
2. Artifacts are *created* through the interactions and collaborations of authors. The pattern of these "creative" activities and interactions reveals a *community structure of knowledge production* (dissertation committees and coauthoring patterns) or network topologies reflective of disciplinary boundaries and the way knowledge is created within them (e.g., centralized, isolated, or clustered research endeavors, see Kuhn, 1962/1996; Newman, 2004; Moody, 2004).
3. Artifacts are also used and *consumed*, so they are linked to consumers whose patterned relations reveals a *community structure of knowledge consumption* not only internal to the authoring population (e.g., citation or reference patterns) but also more broadly for a general public (e.g., downloading, reading, and purchase patterns). These networks of consumption patterns interrelate various types of knowledge (or education materials) in ways distinct from, but related to, the ways in which knowledge is produced (Munzner, 2000; Krebs, 2004).

We expect these networks to vary in structure and to change over time in ways that reflect different theories about knowledge domains and their development. In the sociology of science, scholars have proposed several different network topologies as reflecting the social structure of knowledge production (Moody, 2004). One view is that collaborations form a *segmented world* (Watts, 1999; Collins, 2001; Lieberson and Lynn, 2002) that is either multi-factional or polar. In such a world, knowledge contents will reflect distinct perspectives where no theoretical consensus is present. Disconnected groups will be small worlds where local clustering is high (collaborating with collaborators) and citations remain within that cluster. Another view is that of a *star stricken world* (Merton, 1968; Crane 1972; Allison, Long, & Krauze, 1982) where knowledge contents are centered on key concepts and theoretical integration depends on interconnecting these star perspectives. Social interactions among producers will arise in hubs where resources, authoring, and citation are all focused on established, famous individuals. A third portrayal regards academic fields as a *paradigmatic world* (Kuhn, 1962/1996) where there is a core set of ideas reflecting a gestalt perspective—single method or interpretive framework—and theoretical consensus is often achieved In effect, there is a core-periphery structure reflective of hegemony. These paradigms can fluctuate, but many are sustained over long periods of time. A fourth conception has some direct empirical evidence in the discipline of sociology (Moody, 2004). Using collaboration networks, researchers find there is a *cohesive, permeable world* (Friedkin, 1998; Abbott, 2001) where there are continuously varying and interchangeable cognitive frames and little theoretical consensus. Scholars and professionals constantly loop through wide sections of the idea space rediscovering prior work. Similarly, the collaboration networks are cohesive and wide-ranging where "neighborhoods" and paradigmatic boundaries are permeable, especially due to the transferable skills of methodologists (so people specialize, and some cross disciplines).

Hence, multiple network topologies could arise in the educational knowledge domain and reflect different theories of how it functions.[3] Unfortunately, most prior work views a knowledge domain from a single relational dimension of collaboration or citation and does not explore the structure of ideas or the interrelation of the multiple network-dimensions we propose. The ability to view the knowledge domain as composed of multiple networked dimensions in interrelation will afford a more rich and informative understanding of the field. One can readily imagine a star-driven system of ideas (core concepts), a factionalized collaboration network, and a cohesive, permeable structure of consumption coexisting in an interrelated fashion.

As we begin to view these networks over time (by "chunking" them into years or eras), we should see processes of polarization, centralization, divergence, and clustering, and we should see processes co-occur in various ways reflecting different developmental stories. The field's lexicon may be a history of increasing convergence, while collaborations grow factionalized and consumption patterns permeate prior divisions. Why some dimensions change in fashions different from the others becomes an empirical question. It may be that key structural processes—in any network dimension—will incite changes in others. Hence, if public consumption of books and articles is factionalized along ideological lines, then we might expect certain ideas and collaboration teams to react in a factional manner to meet such demand. How these network dimensions interpenetrate over time reveals the core process of field development—the interaction between social forms and cultural contents of the educational knowledge domain.

Recent work has offered perhaps the closest analog to what we believe the educational knowledge domain will resemble. It concerns the biotechnology field and depicts it as composed of multiple network dimensions spanning from 1988 to 2000 (Powell et al., 2005). Each network is based in distinct activities central to the biotechnology field: commerce, research and development, licensing, and finance. The authors argue that this field initially entailed segmented worlds of academic and private industry research that have increasingly co-evolved and grown more interrelated. The trend in biotechnology is toward more profitable, publicly consumed ideas, where the concerns of profit and basic knowledge become increasingly fused. At the same time, collaborative relations now stretch into private industry blurring the distinction between the biology discipline and the private industry profession.

In sum, we propose a multiple network depiction of the educational knowledge domain that can be interactively explored via an online system that builds on prior work in related fields and seeks to extend and improve upon it by applying it to the field of education. Very little research has characterized the community of scientific work as a multi-dimensional, dynamic system of relations (e.g., Kuhn, 1962/1996). No one has empirically illustrated how the structure of ideas and concepts relates to the structures of production and consumption (including Powell et al., 2005). In short, no one has empirically depicted a knowledge domain as a multi-dimensional networked system, let alone offered an interactive, online technology by which anyone can view and contribute to it as such. Such an online system has the potential to re-orient scholar and practitioner explorations and understandings of the educational knowledge domain. In what follows, we explain how this idea is already a reality in piecemeal form within other fields and applications, and then propose a feasible synthesis that we hope will occur in the near future.

Creating networks

As we look to other enterprises that have developed tools to solve some of these problems, we can learn a great deal about where we can go using existing technologies to provide a significantly more effective approach for the education community. The tools that we look to can accomplish the following:

- *Visualization of full text searches* as related networks of resources.
- Understanding of the *structure of the community* through relationships of researchers *producing* new work (authoring) and research being *consumed* by other researchers (references and usage).
- *Collaborative filtering* that provides real time construction of relationships, personalization of results and indirect feedback to the system.
- *Community participation* at multiple levels, including contribution, review of resources, and monitoring of the system for abuse or "gaming."

Additionally we can look to other academic fields that have made efforts to use similar technologies to provide networked views of either the community or its artifacts to inform our own efforts to synthesize these technologies into a reality. Ultimately, we take these lessons to bear on a proposed system we dub "SPINE," or Scholar Practitioner Information Networks for Education.

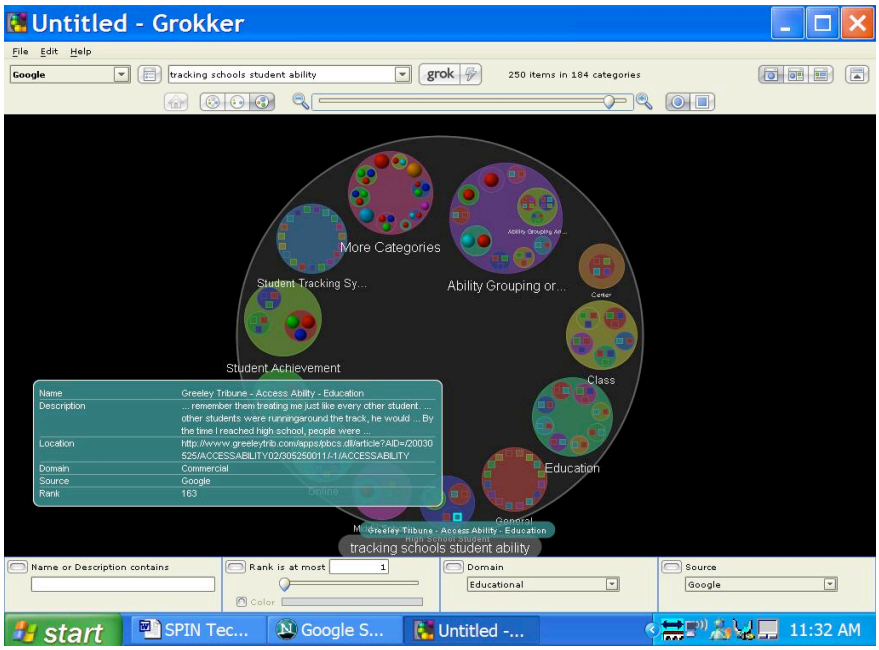
TOOLS FROM OTHER ENTERPRISE

Google (http://www.google.com) is perhaps the most common search engine. It makes searching easier through intelligent algorithms that rate the relevance of results containing the search terms the user has entered. Google represents a quantum

leap over previous generations of search engines that returned search results based on the location of keywords in documents. This early search technology was plagued with problems, many of which have been overcome by Google's page rank algorithm (Brin & Page, 1998). Google's approach is to examine the relationship of pages and calculate ranks using weights for how often those pages are cited broadly across the Internet. This approach has proven to be vastly more useful than the keyword approach employed by earlier search engines and still in use by most databases. Recently, Google has applied their technology to academic literature in the form of Google Scholar (http://scholar.google.com). Using approaches similar to their more general search engine, Google Scholar searches databases of academic literature and returns lists of results for search terms. Searching for the term "tracking schools student ability", Google Scholar identifies a set of results from academic literature that is returned to the user in a familiar list format.

Keyword networks—Mapping cognitive structures of a knowledge domain

When looking for specific items, sorted lists are an efficient way of retrieving and organizing information. However, when exploring broad topics, list formats may not be the easiest way to explore keyword-generated information. There are search tools in existence that can organize these lists into their hierarchical embeddings and allow the user to actively explore the layers until they reach an actual piece of written material. Below is an example of an information retrieval tool called "Grokker" (http://grokker.stanford.edu) that displays the same Google search string results in a hierarchically embedded visual format. Grokker allows the user to select the sources of information they use—from Google, to Stanford's library holdings, to expanded compendiums of online peer-reviewed work from 1987 onward (see Figure 1). In Figure 1, the turquoise square reveals the base reference as noted in the green box.

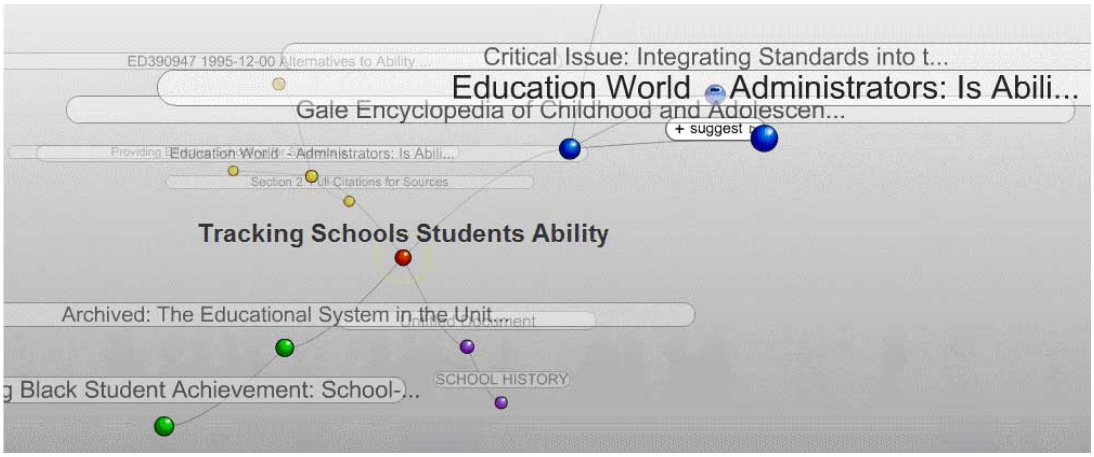


click to enlarge

Figure 1. Grokker visualizing Google search for "tracking schools student ability"

One can take an even more flexible approach by visualizing search results as networks. The procedure begins with a keyword or full-text search that brings up a visualized network displaying the links to related topics or materials, and especially those that have been linked through manual searches in the past. This is a procedure that already appears in databases, dictionaries, and search engines. It allows users to visually explore related concepts and manipulate the network to get to the specific topics that are of interest. A user can start with an abstract topic and drill down to specifics as they explore. With each exploration, a set of links or connections relating various topics to each other builds up.

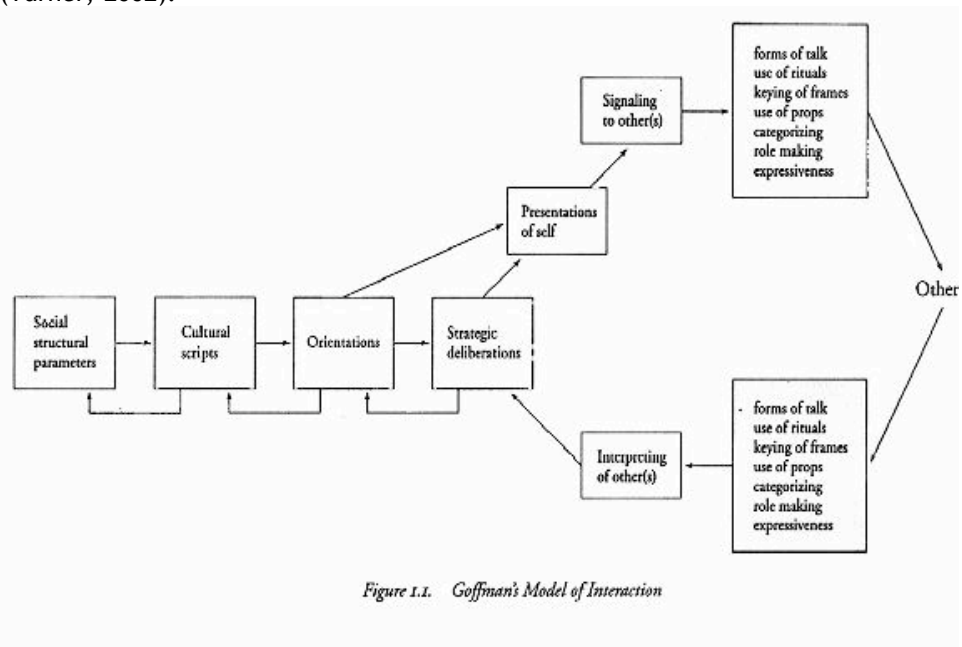
A good example of this practice can be seen in many of the commercial search and visualization packages such as Blinkx.com. For example, a user could enter the database and search the string "tracking schools students ability" (Figure 2). When this is searched, it appears in a network graph along with related topics and materials, such as Black student achievement and publication materials concerning it. By clicking on "black student achievement," the user can then find out what is related to that topic. Sometimes these explorations drill down to more specific or parallel topics, other times they are end nodes that contain actual documents.



click to enlarge

Figure 2—Blinkx.com search engine. The user here searched on "Tracking Schools Student Ability" and saw the related topics

If we take the representation of knowledge contents seriously, then the goal is to show how concepts interrelate in some type of argument or narrative explanation (see Tilly, 2006). Co-occurring keywords are suboptimal in this sense. If we are to better capture meanings of knowledge domains, then we really need to represent how concepts are used with relation to one another in an argument or narrative explanation. Most concept maps and formalistic representations of arguments are manually generated and concern individual texts (Novak & Gowin, 1984; Novak, 1998). Hence, one can turn to theoretical textbooks and find manual depictions of arguments, such as the one shown in Figure 3 depicting the general argument of Erving Goffman (Turner, 2002).

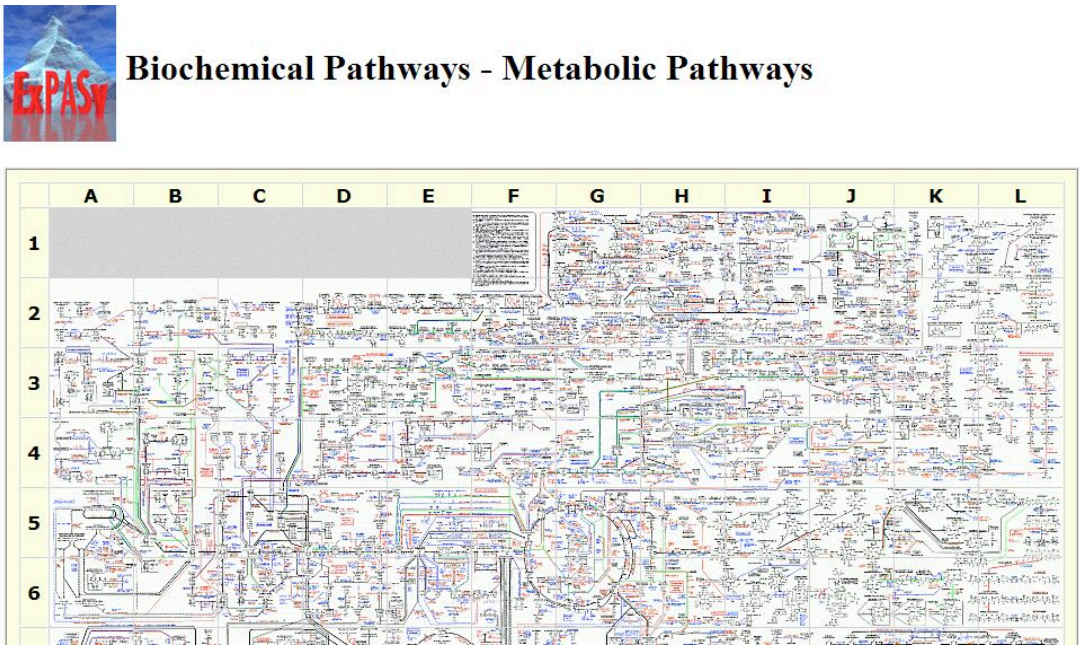


click to enlarge

Figure 3. A manually drawn conceptual map for Goffman's theory of social interaction

This depiction is selected precisely because it concerns a general theoretical formulation—something that is far more common in the educational knowledge domain than in the hard sciences where the lexicon and methodology is far more consistently used. In an ideal world, these argument representations would come naturally and easily so that they could be compiled and interrelated, thereby forming larger maps of what is known and how concepts work together as a system.

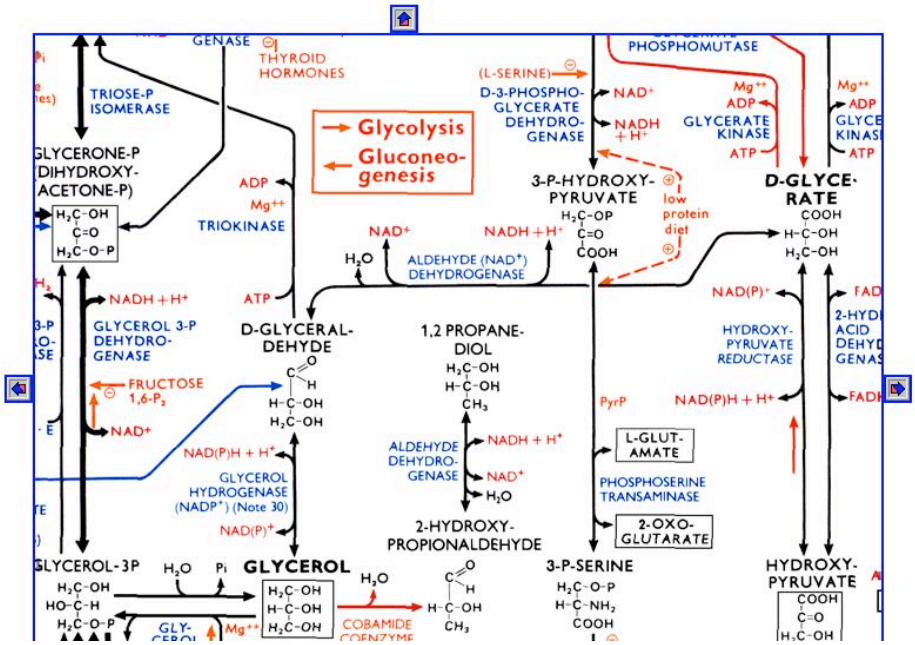
In biology, there has been an effort to do just that. Biologists have taken a special interest in developing larger causal maps that summarize their analyses of proteins. Perhaps the most famous depiction is that of metabolic pathways (http://www.expasy.org/cgi-bin/show_thumbnails.pl). See Figure 4.[4]



click to enlarge

Figure 4. Biochemical Pathways— overview illustration of causal relations

Users can click on the image for more detailed views (see Figure 5). Specific links or pathways can be clicked on to identify research evidence in the form of published materials. By exploring relations in Figure 5, one can dig further and find concept-definitions and references to relevant research articles on MEDLINE (e.g., an online article database for medicine that contains over 11 million documents).



click to enlarge

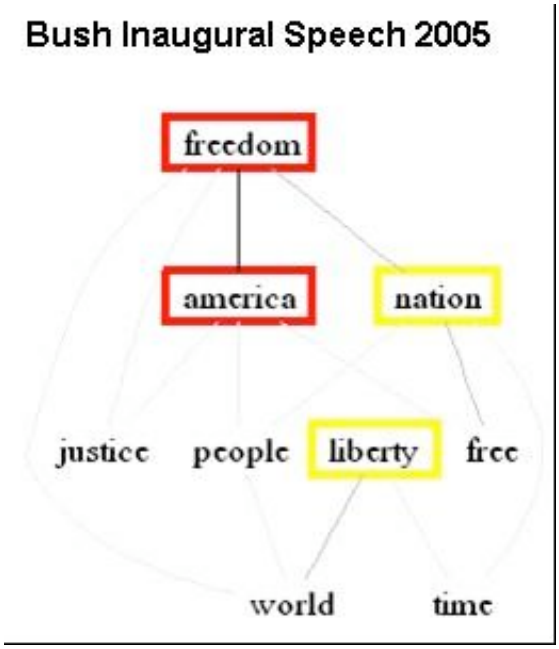
Figure 5. Biochemical Pathways— A blown up image of section G1 in Figure 4

The fundamental goal in these causal diagrams is to develop a visual summary of what is known and how the pieces of what is known are interrelated. There may be unparalleled difficulties in developing such images for the field of education as it has multi-disciplinary foundations and lacks a consistent lexicon. The structures of arguments will vary across philosophers, anthropologists, and economists of education. Moreover, there will be points of contention on each topic, issues of research quality, and so forth.[5] But these challenges are also present in other disciplines and fields. One can find similar issues for

biology where philosophers of biology have different concerns and arguments to molecular biologists, to animal ethnologists, etc. Sub-fields have some degree of consensus (e.g., biochemists), and review panels can be formed to adjudicate points of contentions, and such contention can even be represented to end users who have the ability to actively explore multiple representations. Regardless, an effort to present such synoptical network-views will at least allow communities to see and discuss their current working consensus. Therefore, it seems plausible that the construction of such networks—wherever possible—will be immensely useful to the development of the field of education, and will rely on the support of the education research community in order to flourish.[6]

Another option is to seek out new methodologies that can offer representations of arguments that are comparable to the manual representations discussed above, but afford automated procedures. Here, at least two methods show potential. The first is to employ nascent methods of semantic network analysis (Carley, 1997; Corman, Kuhn, McPhee, & Dooley, 2002). This approach can be used to provide more useful networks than keywords alone, and can also provide information about the ordering, and in turn, directionality in keyword relations. The procedure usually entails the removal of verbs and developing a thesaurus of nouns. Using this reduced set of concepts, a network of their interrelation is depicted. In these networks, terms are related to each other via dependent clauses, sentence structure, and order of presentation. While the representations are coarse, they offer one form of automated argument representation.

One specific method of semantic network analysis is called Centering Resonance Analysis, or CRA (Corman et al., 2002). Whereas other methods equate word frequency with word importance, CRA is based on linguistic theory concerning how people create coherence in their communication. CRA uses natural language processing to create a network model of text. Word influence is calculated based on the structural position of the word within the CRA Network. In Figure 6, we present one such rendering from President Bush's inaugural speech in 2005.



click to enlarge

Figure 6. CRA of President Bush inaugural speech

A second technique is to use computer programs to parse texts and identify how keywords (or citations) are being used. Such an effort has been applied to the Kansas Event Data System (KEDS) which uses automated coding of English-language news reports (consisting of over 500,000 events coded from many thousands of news reports) to generate political event data focusing on the Middle East, Balkans, and West Africa. These data are used in statistical early warning models to predict political change as well as render events into political networks of alignment and conflict (see http://vlado.fmf.uni-lj.si/pub/networks/data/KEDS/KEDS.htm, and Batagelj, Mrvar, & Zaversnik, 2002). The verbs surrounding any particular keyword (in the KEDS case, it is political actors) are then parsed into classes of meaning—such as alignment (e.g., sending arms between nations), threat (e.g., withdrawing embassy officials), etc. Network depictions of this data have only just begun, but illustrate how political actors and issues emerge over time, form polarized debates, and explode into conflicts.

It is a small inferential leap to imagine the application of such methods to online repositories of knowledge artifacts so as to identify how key concepts (or citations) are being used as foils, reinforcing ideas, or as new offshoots. With time, this nascent

method will likely improve and come closer to our manual argument representations, thereby making an automated representation of knowledge contents in the field more feasible. Our point in presenting such preliminary methods is to show that the structure of knowledge contents—as argument structures—is important but difficult, and that it should be rendered in a way comparable to community social structures.

The structure of the education community—Part I: Production networks

In contrast with the cognitive structure of ideas, the social structure of the educational field exists in the pattern of social interactions that are structured by various core activities of field participants. Two core activities are knowledge *production* and *consumption*. Implicit in most every description of knowledge domains are *networks of knowledge production* (Kuhn, 1962/1996). Under-girding what we know in the field of education is a network of knowledge production based in the creative interactions authors have when generating knowledge artifacts. Using tools that can visualize these kinds of networks, when users search a topic they can view how knowledge producers, or people, are organized with regard to the topic. Why would the user want to know this? The capacity to see or understand the relationships by which a body of knowledge is produced enables members of the field to manage their relations so as to guide the production process in more efficient and innovative directions (e.g., to span currently unrelated parallel efforts by different sets of authors). Such relations of production, or collaborations, can be generated automatically from each artifact's listing of authors.[7]

Collaboration networks are perhaps the most consistent with Kuhn's view that social interactions and communities of scholars produce scientific knowledge (Kuhn, 1962/1996; Newman 2004; Moody, 2004). In these networks, nodes are people and arcs are collaborations. Work by Newman on co-authorship in a large private institution shows distinct clusters of research efforts that are linked by particular "boundary-spanning" authors (see Figure 7).[8] If one searched the equivalent of the RNA structure in the educational literature, then one might acquire only the yellow nodes in the network. However, if a system were to prompt the user to select the relational distance from authors they wish to represent (from a default step of 1 relation away, to say 2 or 3 steps removed), then images can be made larger so as to include more distant collaborations.

The problem with the co-authorship perspective is that some knowledge creation does not require formal collaboration while others demand it—e.g., philosophy of education versus educational technology. As such, the collaboration network might seem sparse and non-existent for some groups while for others, they are dense and thriving. This is not an error, but rather reflects the role that formal collaboration plays in different educational sub-fields. One means of more accurately representing relations of knowledge creation is to show mentoring ties (dissertation committees) as related in Dissertation Abstracts listed online (1990 to present). In order to acquire a clearer image of the field, other network representations are needed.[9]

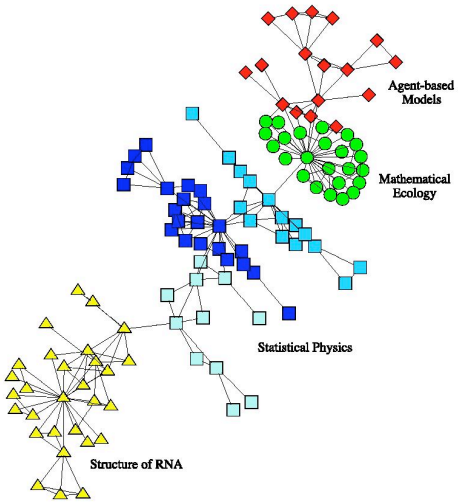


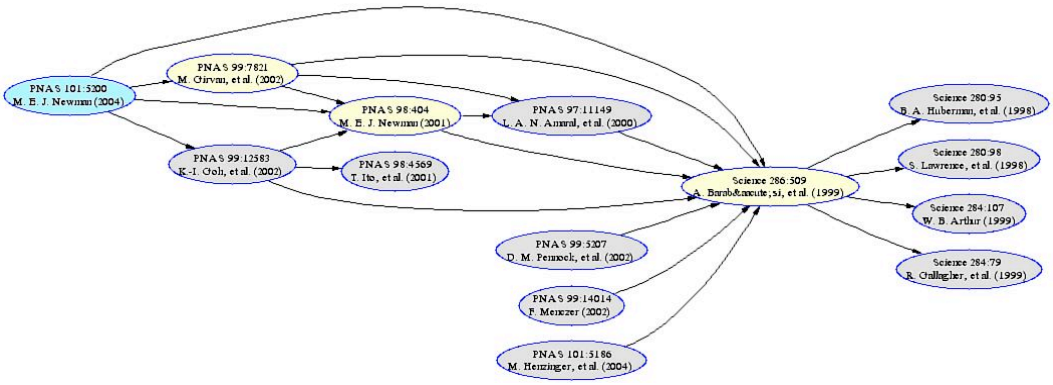
Fig. 1. An example of a small coauthorship network depicting collaborations among scientists at a private research institution. Nodes in the network represent scientists, and a line between two of them indicates they coauthored a paper during the period of study. This particular network appears to divide into a number of subcommunities, as indicated by the shapes of the nodes, and these subcommunities correspond roughly to topics of research, as discussed by Girvan and Newman (37).

click to enlarge

Figure 7. Collaboration networks in private research institution (Newman, 2004)

The structure of the education community—Part II: Consumption networks

Consumption networks are the product of authors and professionals consuming the work of their colleagues. They are another means of identifying the informal groups that shape knowledge. In contrast with collaborations and mentoring relations, references show authors' interests more clearly (e.g., knowledge traditions, etc.) and the specialty structures of common works. In citation networks, nodes are knowledge artifacts (e.g., articles/curricula/or reviews), and relations reflect the references being made across them. The Proceedings of the National Academy of Sciences has instituted network visualizations of this type using citations from its papers, and data contained in a common journal database. When browsing articles, one can easily obtain a directional network map of the papers cited by any central reference (Figure 8). The depth of this map (i.e., the top 10, 20 or 30 most-cited articles) can be chosen by the user so that a wider or narrower ranged network may be viewed. By selecting a particular node, the viewer can explore any cited document.



click to enlarge

Figure 8. Proceedings of the National Academy of Sciences citation maps showing papers cited by Newman, 2004. The blue citation is the central reference and the yellow citations are cited most often.

Work by Chen has focused on the structure of citations over time in articles concerning string theory (Chen, 2004, 2003; Chen and Morris, 2003). He finds various developmental turning points in that line of research (see Figure 9).

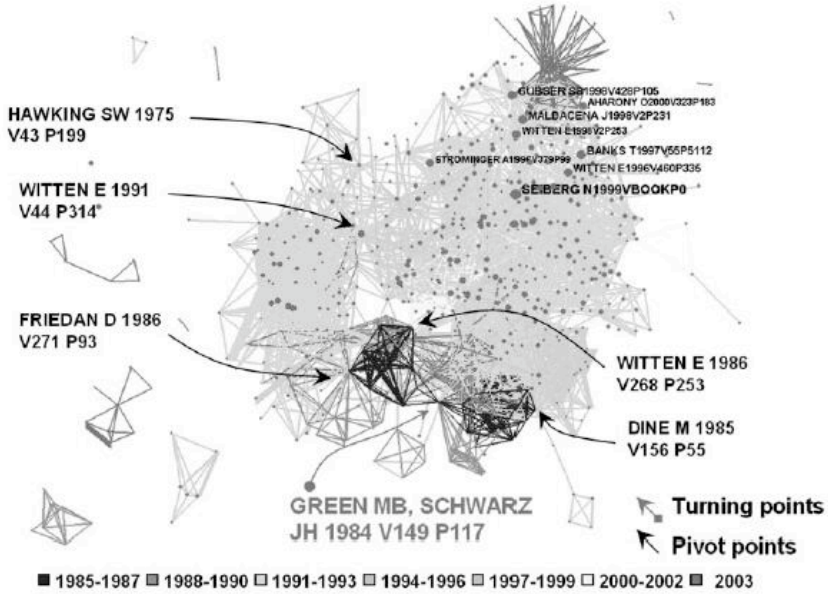


Fig. 3. A 624-node merged network without global pruning. See a color version at: www.pages.drexel.edu/~cc345/citespace/Figure3.png.

click to enlarge

Figure 9. Co-citation networks for string theory in physics (Chen, 2004).

The problem readily seen here is that citation patterns can create overly dense networks so some sort of pruning is required. Several simplifying heuristics are commonly used to render networks more interpretable.[10] Earlier, we described how the Proceedings of the National Academy of Science allows users to depict the top 10, 20, or 30 most-cited articles centered around any particular article. Such a technique is akin to page ranking, and is aimed at capturing the structure of prominent nodes (where prominence can be citation rank, professional regard, etc.). Chen adopts a similar technique when he removes all articles that fail to have a high co-citation level with other articles in the graph. As a result of such pruning, Chen's graph offers a more streamlined presentation and becomes far more interpretable (see Figure 10).[11]

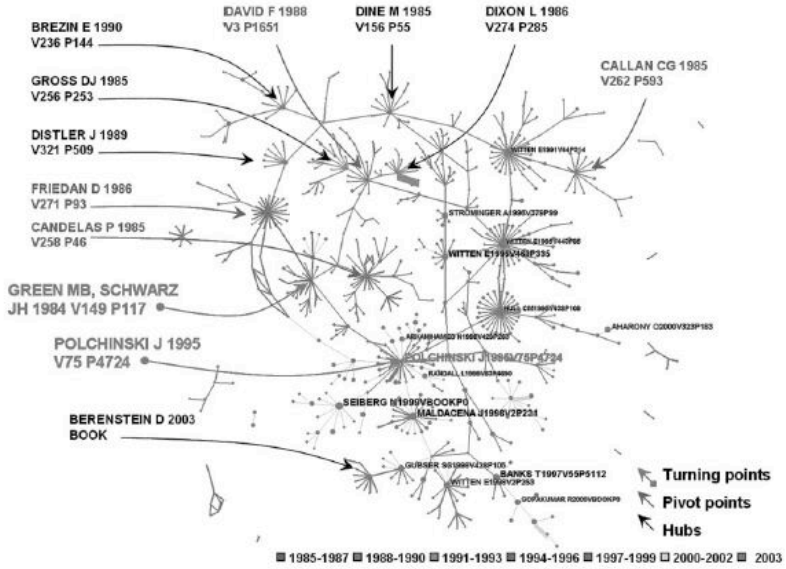


Fig. 4. A 624-node merged network with global pruning by using Pathfinder ($q = N - 1$, $r = \infty$). See a color version at www.pages.drexel.edu/~cc345/citespace/Figure4.png.

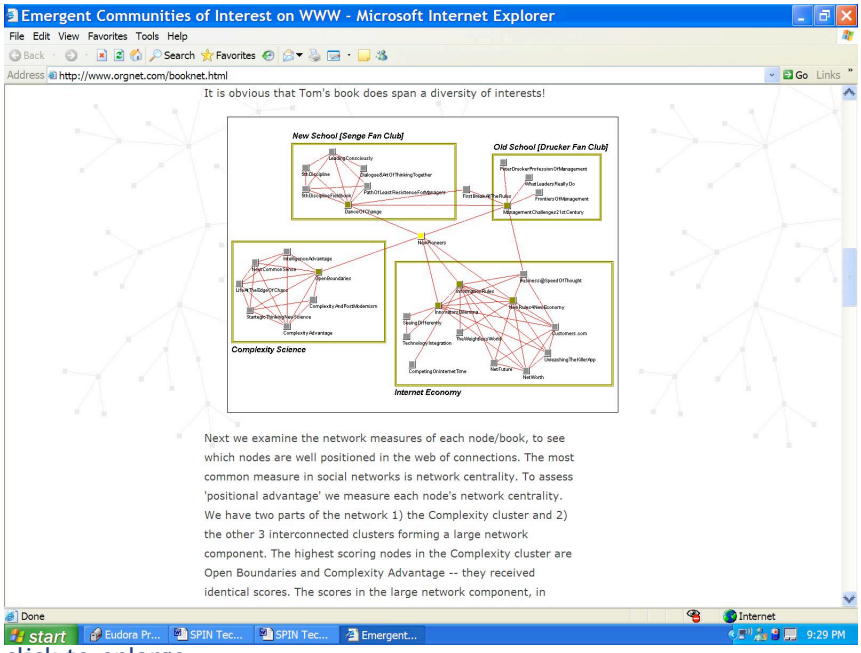
click to enlarge

Figure 10. Pruned co-citation network for string theory in physics (Chen, 2004, p. 6).

Collaborative filtering revealing consumption patterns

Citation patterns are mostly shaped by knowledge-creators, so it does not capture how the wider public engages and consumes these materials. Networks that represent the consumption or page-usage patterns of research databases can reveal additional useful information about the educational field. Many websites are employing forms of collaborative filtering for this purpose (Linden, Smith, & York, 2003). A common example of such filtering can be found at sites like Amazon.com. Amazon.com catalogs the page selections of users, analyzes the general patterns of browsing and consumption (e.g., books), and compares the users browsing activity to normative patterns in order to identify and suggest items they might find interesting ("*Customers who bought this book also bought these books*"). These links are not predefined, but emerge from the usage patterns of people browsing the site and selecting useful resources.

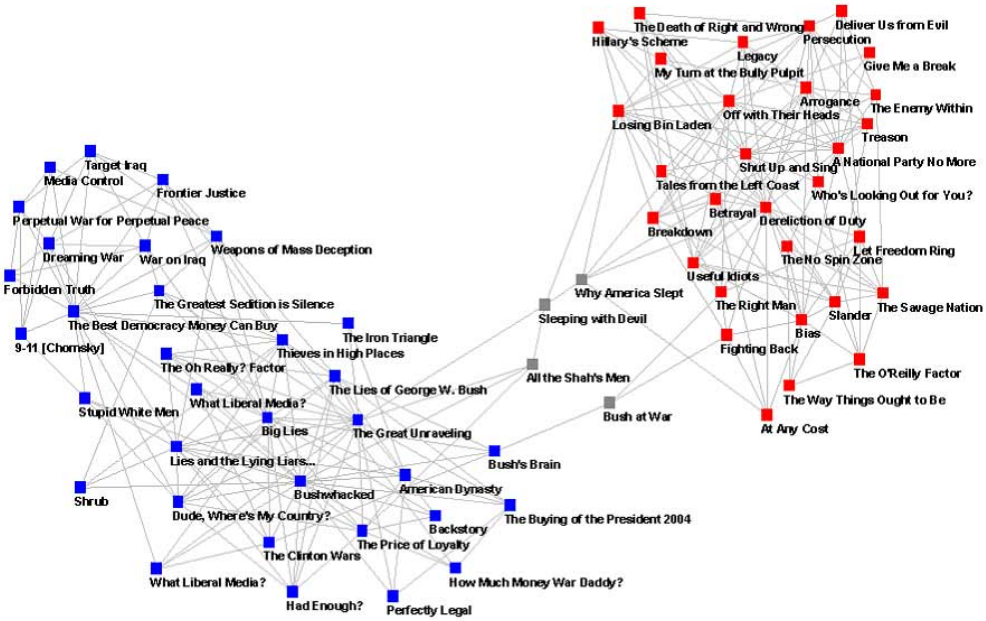
Valdis Krebs has analyzed some of the information Amazon affords by listing the top six books bought by individuals who also bought the book currently being browsed. In one example, he studies the network surrounding the book, *The New Pioneers* (see Figure 11). In the network graph, each node represents a book and the lines indicate which books were purchased together. Krebs finds the buying pattern of books self-organizes into clusters that he names for their content (http://www.orgnet.com/booknet.html).



click to enlarge

Figure 11. Book purchasing patterns around New Pioneers on Amazon.com

In another example, Krebs uses the top 100 political books on Amazon and maps out their interrelations (Figure 12, see http://www.orgnet.com). By doing this, he reveals how consumption patterns of political texts are guided by opposing political orientations (see Eakin, 2004 and Carolan & Natriello, 2005). The blue nodes reflect liberal tastes, the red conservative, and the gray are books bought by both sides.



click to enlarge

Figure 12. Analysis of book purchasing patterns related to the top political books purchased on Amazon.com

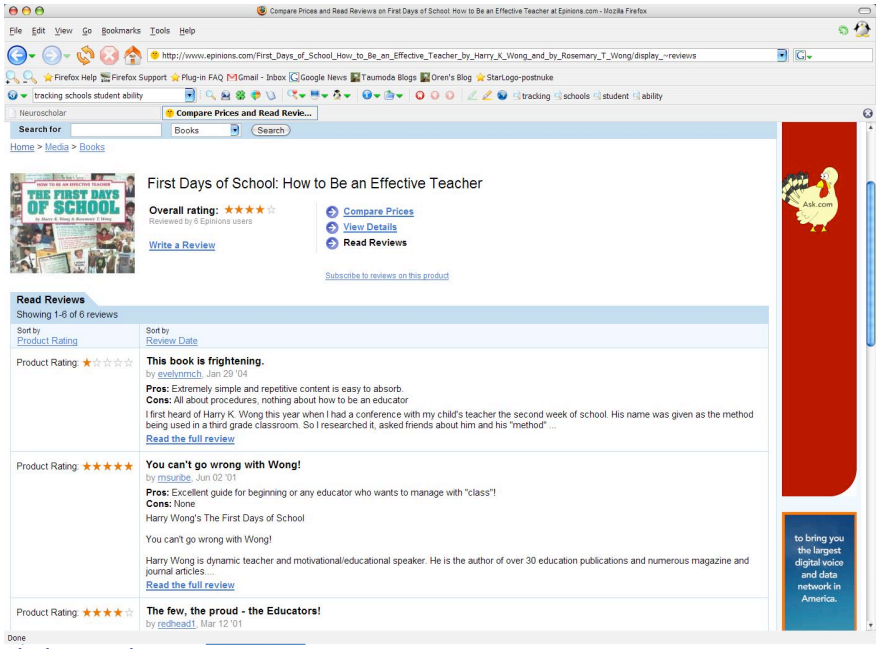
The point is that these visualizations can be derived from information automatically recorded over time to reveal how users consume the information afforded and move across various web pages. These consumption or usage-patterns give us yet another perspective on the field of education. Here we can see how educational materials are being consumed by a wider public, and possibly understand when or why some work has greater cache. Is it due to ideology or current public crises? Who consumes what—are professionals purchasing certain books over others? Why? Furthermore, by studying these networks over time and in interrelation, we can assess how public consumption patterns can have feedback effects on collaborations and the structuring of ideas.

When fully automated, consumption networks of this sort may have certain drawbacks. First, the network form may be contingent on the search algorithm utilized. Users move through topic sets according to the lists generated and the connections displayed in the network graphs. Once a link is shown, then people will be more likely to user that route again—creating a self-fulfilling network form. In addition, the resultant graph will depend on how narrow the search (i.e., focused on a single book, or top 100) and how high the cutoff level is (i.e., Amazon lists six books most co-purchased).

Last, much of the browsing being done entails lots of "noise" or wasted time and "wild goose chases" so usage patterns may be "fuzzy" in comparison with more serious usage like purchasing patterns. Noise in usage-patterns can be reduced, but it requires asking users to review the links being given (e.g., which are more or less useful). In this manner, a snowball sample of "useful" searches forms and generates network links from higher quality browsing experiences. An example of this augmented feedback model can be found at Amazon.com, where users are given book recommendations not only on the basis of their browsing (as described above), but also on the basis of their active evaluation of suggested texts. Users are prompted to "improve your recommendations" and they rate items listed so they get a better profile of the user's tastes.

Community participation, reputation, and critique

Knowledge domains are characterized by normative judgments of the knowledge contents being expressed. In academia, this evaluative and aligning action occurs in the way citations are used. Some work is referenced as a foil—what is being argued against. Other work is referenced as a precursor, or aligned partner in intellectual debates. Technology will soon automatically display authors' usage of references to reveal this normative, evaluative dimension of academic knowledge domains.[12] However, when it does, the wider public's opinion may go unnoticed, or worse, when noticed it may still be regarded as "fad-driven" or "ideologically-driven" and therefore dismissed offhand. The route to acquiring valuable public critiques is already being developed by various websites. These websites accompany descriptive resources with reviews or evaluations of materials. Users write commentaries and reflections on them so opinions are known. In the aggregate, these reviews help reveal points of consensus and contention concerning various materials. These reviews are commonplace at retailers such as Amazon.com where the reviews accompany the product descriptions, but they also exist in separate websites like Epinions.com which has created a community that provides reviews of merchandise including movies, household items, consumer electronics and books (see Figure 13).



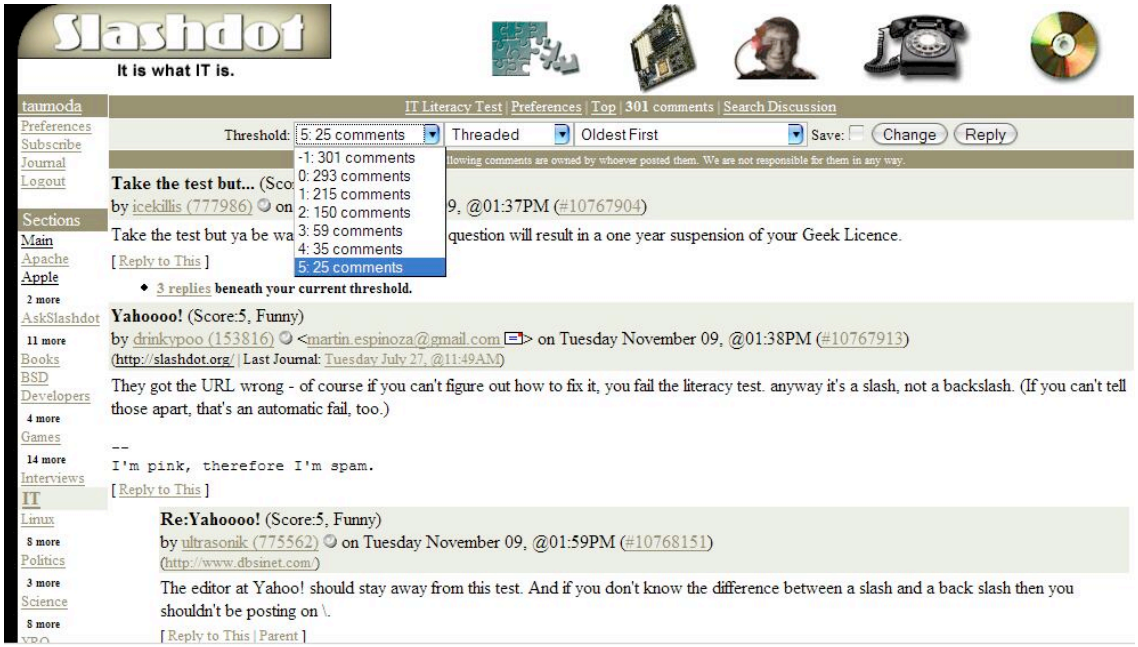
click to enlarge

Figure 13. Reviews of a book on teaching at epnions.com

Some sites contain another layer of user-evaluations—or ratings of reviews (and, in turn, reviewers). Upon reading a review, users are given the option of evaluating its usefulness, or how informative and helpful the review was. Reviews that are consistently ranked as helpful can be sifted to the top for representation while those consistently ranked as unhelpful are sifted to the bottom. In addition, certain reviewers will repeatedly write good reviews, and sites seek to reward (or punish in opposite cases) these submitters with a reputation score reflecting some measure of expertise (i.e., number of reviews performed, average rank of reviews, standard deviation of rank). Such a system allows for broader community input, and it validates

reviews and reviewers so as to ensure relatively stable, consensual standards of evaluation.

In many online communities, the members generate the content themselves. The volume of information that some communities generate, even on a daily basis, is so immense that it requires a degree of ranking, filtering or organization. Rather than having an editor solely responsible for fielding and creating questions, answers, and reviews, this can be done in a decentralized way. Communities such as that of Slashdot.org, a staple in the technological sector (see Figure 14 below), have accomplished this. Slashdot is a community discussion board that is based on news items, often accompanied by a question. Each question or topic is accompanied by a discussion that is open to the community. Unlike many discussion boards that display the entire discussion thread to everyone who views it, Slashdot employs a rating system that allows readers to filter the postings and to view only the resources designated to be the most valuable by the community. The ratings are contributed by active members of the community who are given ratings points to distribute.



click to enlarge

Figure 14—The Slashdot website showing a news story with hundreds of responses. It is easy to filter and get only the most valuable responses as designated by the community.

Any particular news item can receive hundreds of responses. Community members then assign a value to these responses, but not everyone has equal say in the voting. Frequent visitors to the site build up "Karma" points that allow them to vote on the responses. If they use that Karma wisely (as ranked by the community itself) they can continue to build up more Karma points. Other successful websites like Ebay also allow users to build up ratings. On Ebay, buyers and sellers develop reputations based on their past transactions with each other (see Figure 15 below). Amazon.com has a similar process where book reviews are offered and users rate the helpfulness of reviews.

Figure 15. The Ebay website showing a seller and his reputation. This is in the form of positive feedback which can be reviewed by individual comments.

These examples show community participation in a primarily decentralized fashion where practitioners and laypersons can play an important role. This participation provides a means for broader member representation in the community, whether that be as contributors of primary resources, reviews of those resources, or ranking of the value of other's contributions or reviews. As members in the community build up a reputation, they have a greater influence in the community and the knowledge networks being represented as reflective of current "consensus."[13]

SCHOLAR PRACTITIONER INFORMATION NETWORKS FOR EDUCATION

Building on the experiences and resources presented above from academia and other enterprises, we have synthesized a design for an Internet technology, "SPINE" (Scholar Practitioner Information Networks for Education), with the goal of providing members of the educational community with the ability to explore and reflect upon itself as a networked system (a similar idea has arisen for business knowledge at the Massachusetts Institute of Technology, see Malone, Crowston, & Herman, 2003).

SPINE's development would require the involvement of the wider education community, and would draw on the participation of researchers, practitioners, and wider public involved in the educational field. The purpose of SPINE is to serve the diverse needs of the education community by providing it with forms of knowledge useful to multiple audiences from different perspectives, allow direct study of the educational knowledge domain, and identify areas of overlap and disparity. As such, SPINE will allow users to actively explore the networked or systemic nature of educational knowledge through a variety of lenses (see Pea, 1999, for similar effort).

Scholar Practitioner Information Networks for Education would be interactive and customizable, such that users can pick forms of knowledge, their sources of input, and even be limited to those forms of knowledge that they deem of high quality (e.g. a high reputation ranking). The user would be able to move inward on a network toward greater detail and outward in abstraction to view the larger structure. Moreover, SPINE would be a means by which multiple sources of knowledge can be integrated, organized, evaluated, and directly related to other forms of knowledge.

In what follows below, we briefly describe the types of knowledge inputs or materials that represent the data of SPINE. Then we draw upon the network visualization examples and principles above to describe how SPINE can be actively explored in a variety of ways that ultimately lead users to see multiple network representations or multiple systemic perspectives of the educational field. In this fashion, we propose SPINE as a networked system of educational knowledge that will reorient explorations of the educational knowledge domain, facilitate broader participation (via reviews), and become a database for the field's self-reflection.

Sources of SPINE

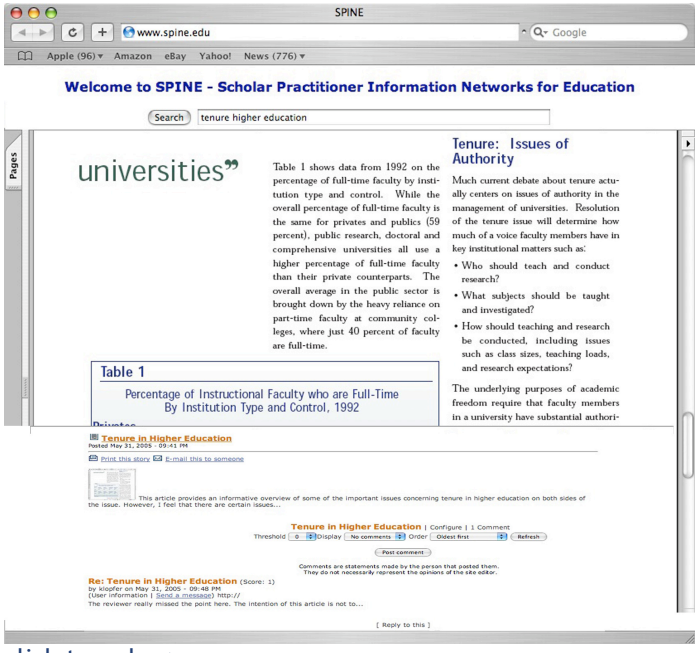
SPINE is conceived to index and interrelate what the educational community knows. "What we know" would be derived from multiple sources. Drawing upon the systems and networks referenced above, we have conceived of SPINE with three layers of knowledge inputs in mind: base written materials, reviews, and ratings of reviews/reviewers.

Academic scholars are primarily concerned with research, and legitimate research (at least to the community as a whole) comes in the form of peer-reviewed books, articles, and conference proceedings. These materials are currently found in commercial databases, and they may appear in full or as abstracts, just as traditional journal databases provide (see ERIC, the Education Resources Information Center, a massive government-run database of published and released research on education). These materials may in fact be supplied by partnerships with commercial or government supported databases, alleviating the need to recreate these materials. Users of SPINE would not necessarily be aware of the multiple data sources working behind the scenes (as happens with Google Scholar). Inputs may also include experiential cases that would be contributed by practitioners and experts in the field. SPINE would store cases, or logged experiences where educators and scholars discuss educational problems and lessons learned from the field (e.g., the case library at Harvard University). A library of experiential cases will make it possible to see how practitioner experiences relate to scientific research. Submissions of case-based experiences will be encouraged by members of the community. Here, practitioners can present case-based evidence that becomes integrated with information from other researchers and existing resources within SPINE. This is one of the basic tenets of SPINE providing for multiple forms of valued input from varied members of the education community. While practitioners are increasingly being called upon to implement research-based practices, and academics are being asked to increase the relevancy of their work to practice, there is little opportunity to see the input from these communities in relation to each other. By using these multiple forms of input, academics can more readily access knowledge from practitioners, and see their own field of work in relation to practice. Similarly practitioners can provide valuable contributions to academics, and make sense of their own work in a vast field of academic knowledge. They can also do so without wading through an overwhelming amount of information, and instead be guided by the consumption networks of their expert peers.

With these multiple forms of input, filtering in and filtering out content will be critical. Evaluations will be pervasive for all of the base materials of SPINE, though those evaluations may take different forms for the various inputs. For academic materials, the evaluations might be viewed as reviews of published materials that summarize the main points and critique the work. For the case materials, the community will be encouraged to comment on the validity and utility of the written case-experiences. By having reviews and a subsequent level of rankings of reviews and reviewers, we would integrate a design feature in SPINE that would encourage the development of experts and informal standards of practitioner (and research) knowledge. Reputations would be associated with the multiple forms of input into SPINE, through source materials and ratings. In such a fashion, we hope that highly regarded experts and opinions will be compared to prevailing research to identify when, where, and how they align or contradict each other. In addition, user evaluations of academic scholarship would convey to academics other perspectives on their work. In SPINE, reputations would be built in ways similar to the many communities cited above, and be used to provide credentials for submitting materials, moderating discussions, or changing the strength of links in the visual maps we will describe in the following section. As the community grows, those who have earned higher reputations may ultimately have greater sway over search results in SPINE.

This three-layered input of knowledge—from base materials of peer-reviewed research and experiential cases, to critiques of

these documents, to rankings of critiques—is the “stuff” of education knowledge. It is the means of understanding and evaluation. By rendering these inputs for an online system (as shown in prototype in Figure 16) that inter-relates them in a variety of ways, we would make the system of knowledge and its variable quality the focus of representation and reflection in online searches. As Figure 16 shows, a simple search has produced a set of results (for the organization of those results, see the following section). When drilling down to a particular piece of content (in this case a journal publication), the content is viewed in an upper frame, while the lower frame has information on the ranking of this paper, and a threaded and ranked discussion of the paper. Readers may not only view the paper, but sort the rankings by data and value to the community. SPINE members may also contribute comments, and as their reputation increases they will also be allowed to rate the comments of other members. That reputation is earned by reading relevant articles, making valued postings, and having agreement with other users on any rankings that they have already done.

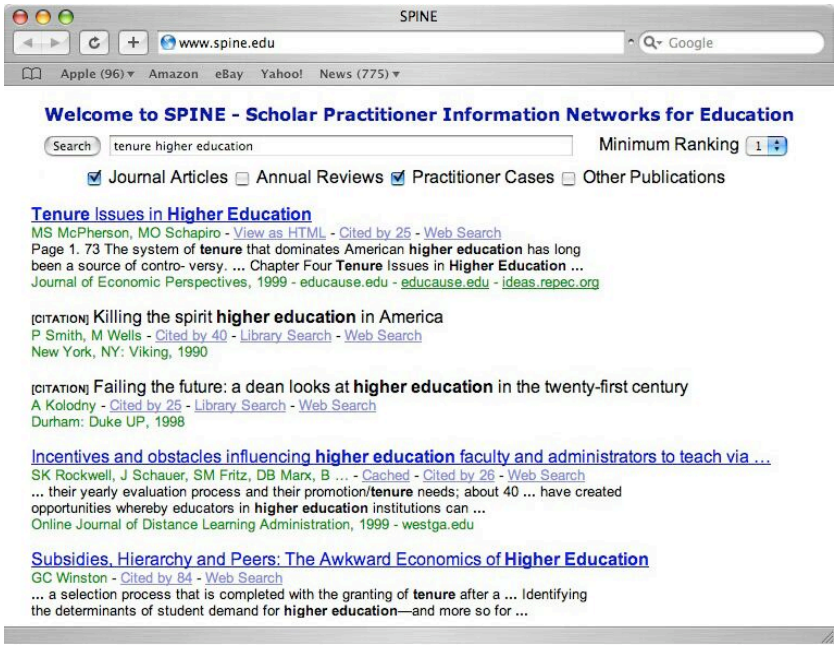


click to enlarge

Figure 16. A mockup of SPINE showing a PDF of an article on top that was identified, along with comments, ratings and feedback on the bottom.

Exploring the networks of SPINE

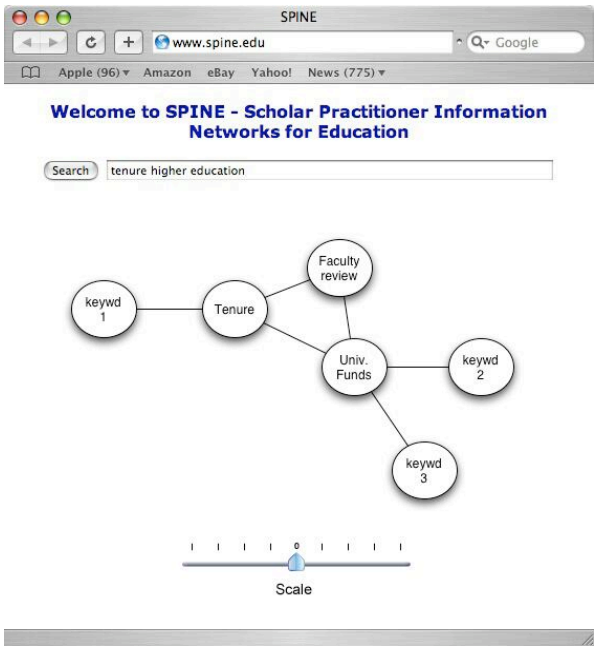
Users of SPINE begin by selecting various databases of interest and then inputting a normal search string (like Grokker—user can pick which databases to search [e.g., JSTOR, ERIC, Stanford library, or combinations]). From this search results a traditional listing of links (Figure 17). Using a dynamically updated view, the user can narrow the listed links by excluding or including certain kinds of knowledge (e.g., experimental designs, non-experimental designs [survey data], narrative studies, or practitioner cases). They can further narrow the range by selecting only those topics that have been reviewed and considered of high quality (e.g., the rating of reviews are good, and the reviews of the materials are positive). What results is a traditional keyword association, with relevant links listed by ranked relevance. Even this somewhat more familiar view provides several advantages over current search techniques. First, the ability to dynamically filter information allows the researcher to explore the available resources much more fluidly, allowing them to appropriately adjust the criteria of their search. Second, it simultaneously explores multiple forms of knowledge, and lists them in relevance to each other. And finally, it allows the searcher to only view those resources that other people have also found of value. Together, this allows for greater customization of the search and enhanced reliability of information.



click to enlarge

Figure 17. A mockup of SPINE showing the traditional list view along with the criteria to change the sources of the search and a threshold for the ranking of the materials presented.

The list view provides easier and more reliable access to the information contained in SPINE. It also draws upon the networks of relationships that we have established to be important in this academic domain. But it does not present that to the researcher in an explicit way. SPINE will allow the user to quickly jump from the list view (Figure 17) to one of the other networked views, such as the content relationship network (Figure 18). This view allows them to see the same content as the list view but clumped together in spheres of relationships, with linkages shown within this content area and between this content area and others. The scale at which the network is viewed can be dynamically changed via a slider to see the relationship between closely related individual knowledge items, or more distantly related clusters.



click to enlarge

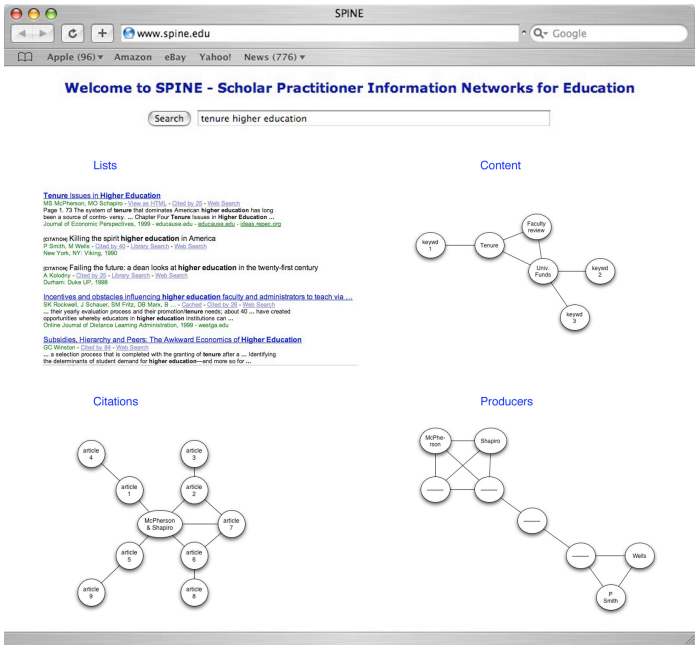
Figure 18. A SPINE visualization of a content network resulting from a search. The user can probe down into individual items (resulting in a view like Figure 16), or zoom out to see the relationship between the initial resources and others that might be more distantly related.

Similar network representations for the other networks are also presented as options to the researcher, and she can quickly move between these views. The networks may also be presented simultaneously on screen (via split screen representation—see

Figure 19 for a prototype view) to facilitate insights into the relationships between the networks and to identify gaps in knowledge. To recap, here are the four networks and their formulation (see Table 1):

Table 1. Summary of SPINE's network representations and their formulation

	Keywords and Phrases Network	*Production Network of Co-Authors*	*Consumption Network by Researchers through Citation*	*Consumption Network by Readers through Collaborative Filters*
Basic relation	Pattern of co-occurring keywords (categories) derived from listed materials	Co-authorship links derived from listed materials	Citation links derived from references in materials	Pattern of webpage usage/searching derived from listed materials
What is displayed	Displays most common co-occurrences, and/or those so many steps removed, and/or those deemed useful (by reviews).	Displays knowledge as interrelated according to the degree of overlap (# of authors in common) and path distance from topic (# of steps).	Displays knowledge as interrelated according to the degree of overlap (# of cites in common) and path distance from topic (# of steps).	Displays most common paths, and/or those deemed most useful (by reviews), and/or so many steps removed from original topic.
User selections	User selects # of steps out from core topic (e.g., 1-3), # of most common co-occurrences (e.g., 1-10) and/or those deemed of high enough quality (e.g., a rating of "excellent").	User selects how many co-authors are a tie, and how many steps from the core topic relations will be displayed (e.g., 1-3).	User selects how many citations are a tie, and how many steps from the core topic relations will be displayed (e.g., 1-3).	User selects # of steps out from core topic (e.g., 1-3), # of most common paths (e.g., ranked 1-10) and/or those deemed of high enough quality (e.g., a rating of "excellent").
Means of Data Collection	Automatic but supplemented by user feedback	Automatic	Automatic	Automatic, but supplemented by user feedback



click to enlarge

Figure 19. A mockup of SPINE showing the traditional list view along with the primary network views resulting from the user's search which may be presented to the user simultaneously allowing them to choose the most appropriate representation, and view how they relate to one another.

IMPLICATIONS OF SPINE

A technology like SPINE can provide scholars, practitioners, and the broader public multiple network perspectives of the educational field that they can explore. The most likely contribution of this type of system will be in providing a new orientation, or a new means of exploring and studying the knowledge base present in the field of education. SPINE has the potential to alter the way we initially approach and consider the educational knowledge domain, thereby altering the way we do research.

By altering community members' initial conceptualizations of the educational knowledge domain, this technology also has the potential to bridge some of the most persistent dualisms in education research (e.g., dualisms of quantitative-qualitative and theory-practice, see Lagemann, 2002).[14] By representing concepts, producers, and consumers of educational knowledge as interrelated, we will learn whether there are distinct sets of ideas, authors, and publics who follow different research paradigms, or whether said incomensurabilities between methodological traditions are actually few and far between. SPINE's proposed network representations will utilize multiple types of knowledge that comprise the education literature, including small-scale qualitative studies produced by practitioners that can be put on an equal footing with quantitative forms of knowledge, and then interrelated and represented in larger webs. Network representations, therefore, have the potential to show us when and where actual intellectual divisions exist, and on what dimensions. If quantitative and qualitative researchers use the same concepts and have commensurable finds, but they have distinct communities of production and consumption, then it suggests that the supposed divide is more the result of factionalized social relations than distinct sets of ideas. If divides only exist with regard to production, then it suggests that many of our espoused dualisms do not reflect the way we use concepts and consume ideas but reflect whom we work with in producing ideas. The implication then is that this divide can be treated via pointed bridging efforts and collaborations.

Merely depicting networks of relations will reveal to scholars and practitioners where linkages exist that they did not know of before. In such a fashion, it will at least force different research paradigms to recognize each other due to their shared focus. As such, a technology like SPINE could show the field of education as simultaneously a community of collaborating participants, as a community of consumers, and as a network of interrelated ideas. When compared, these network perspectives afford the opportunity for critical reflections (from one standpoint viewing another) on how the field of education is developing and how we can direct its development in a more fruitful manner.

SPINE also has the potential of lessening the divide between theory and practice and bringing practitioner and stakeholder concerns into greater salience. Traditionally, it has been difficult to render practitioner knowledge and expertise in a systematic way that can be brought into greater relation to scholarly work (Hiebert et al. 2002). The proposed technology gives practitioners and members of the public a more prominent role in the way we search, conceive of, and define the educational knowledge domain. First, SPINE will be a public search engine and knowledge repository. Hence, it will expand access in a way akin to Google. Second, SPINE will be a repository for practitioner materials (curricula) and evaluations that members of the public generate. This broadens participation in the production and evaluation of educational knowledge. Last, SPINE will include consumption networks, thereby making the public's usage of educational knowledge a core component of the technology. When it comes to consumption, it is clear that the broader public far outnumbers the select group of scholars using educational knowledge, and will thereby bring public, professional, and scholarly conceptions of educational knowledge into greater relation. We believe all of these proposed features will give practitioners and educational stakeholders (parents and students) motives to use and contribute to SPINE.

Another potential payoff of SPINE is that it will afford data by which we can study the causes of innovations and the pace at which a knowledge domain develops. There have been theories of what constitutes a new idea or innovation, as well as theories as to what kinds of relational systems generate collective synergy (e.g., Carnegie Mellon in 1950–1965, where the beginning of organizational science arguably occurred and led to today's graduate schools of business). Some argue that when two key but disparate concepts are merged, that a new topic space and niche for funding/innovation arises (innovation as conceptual "bricolage"). Others argue that innovation arises from taking the methods of one arena and applying it to another—so seeing the same thing a different way (conceptual embedding or reframing). Via an online system, we could acquire data and generate predictive models identifying the social mechanisms driving various forms of intellectual innovation and stagnation, and so on, all in an effort to facilitate greater efficiency and innovation in the field of education.

CONCLUSION

We propose a new Internet technology, SPINE (Scholar Practitioner Information Networks for Education), that is based on existing technologies and theories, and which will provide researchers, practitioners, and education stakeholders the ability to explore and reflect upon itself as a networked system. SPINE's base materials consist of peer reviewed work, curricula, and practitioner cases. In addition, reviews will be affixed to these materials, and ratings will be affixed to reviews.

As the field of education moves forward, both scholars and practitioners must be able to develop insights into the field through a systemic view (Carolan & Natriello, 2005). Developing a networked system perspective is a challenge, and it is all too easy to instead develop a narrow view favoring only a limited subset of available evidence and expertise when making decisions or

conducting research. This is particularly evident as one looks at the relative role of scholars and practitioners and how their contributions become part of the field. It is critical to have inputs from and accessible outputs to both scholars and practitioners and find ways to bridge the gaps between their very complementary contributions. SPINE is a feasible technology that can provide a bridge across this gap that can better equip all members of the field of education.

Notes

1. It is important to note that there is a difference between search efficiency and research efficiency. More comprehensive, intelligent searches of knowledge artifacts facilitate academic work by identifying the context of ideas and social relations in which an idea is embedded. It does not make reading and comprehending the actual texts more efficient. Moreover, such a system should not challenge the role of foundational texts, but rather make their centrality in the field all the more apparent.
2. Work in the sociology of knowledge has overlooked this claim related by Simmel (1971) and Karl Mannheim (1936). Instead, the focus is overwhelmingly on social forms, or the relations of production and collaboration (Collins, 1998). Entire philosophical traditions are characterized as the result of networks of interpersonal interaction and fail to take into account characteristics of the knowledge itself and the ways in which it is consumed and judged.
3. In technical terms, we propose rendering artifacts into affiliation networks (artifact x author, etc.), which through matrix multiplication can reveal a variety of comparable structures. We intend to go beyond prior work and introduce asymmetry into affiliation networks (See Breiger, 1974). Breiger makes an argument about primary and secondary affiliations—simply A = primary affiliation where cells 1,0 (say core concept, first author, or a primary citation), and B = all other concepts, authors, or citations. One takes Q=ABt (where t=transform) or inversely R=BAt depending on what kind of network representation you want to generate. If C=A+B, then we can generate S=CCt or T=CtC. Valued directional networks would then be a simple result of element-wise multiplication, where X=Q#S or Z=R#T (#=element-wise multiplication). In addition, citation placement and frequencies within articles may offer a means to identifying core concepts and cites. Once directionality is had, we can explore a far wider array of network forms reflective of political processes and hierarchical understandings of the field (see blockmodels in Wasserman and Faust, 1994).
4. Another relatively famous figure has been developed to show cancer pathways (http://www.nature.com/nrc/journal/v2/n5/weinberg_poster/). There, different kinds of mechanisms or links are rendered in different colors.
5. The challenges in generating argument networks for the field of education are perhaps precisely those for knowledge writ large, or the knowledge being searched on the Internet. In effect—if we can accomplish this for the field of education, then we can accomplish it for most other knowledge domains.
6. One of the best examples of viewing community resources in multiple networked ways can be found in the neurosciences using the Neuroscholar website (Burns et al., 2003). Neuroscholar contains a visualized network of related topics and end-nodes that represent the resources themselves (e.g., publications, curricula, and written cases). Neuroscholar has the advantage of being able to call up primary resources inline when a user drills down to that level of the network. However, the neurosciences have a distinct advantage over the field of education in organizing topics and their interrelations. Neuroscientists generally agree on physical and biological keywords and their limited kinds of relations (e.g., physical-chemical), whereas the system of classification in education is quite under-developed.
7. Because some sub-disciplines will not use collaboration very heavily, we also intend to rely upon dissertation committee listings that can be readily acquired from dissertation abstracts (starting 1990). In this fashion, “relations of production” will include collaboration and mentoring. As with other networks, these relations can be weighted to favor first authors and illustrate some directionality in affiliation networks.
8. URL-http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=387296andaction=streamandblobtype=pdf
9. Careful reading and study of entire research traditions will of course offer another important perspective. Work by Donald Levine (1995) and Randall Collins (1998) depict entire theoretical traditions of sociology and philosophy as networks of friendship, opposition, teacher-student relations, and critical reception. The drawback of such work is that minor figures are never fully taken into account. They are histories that emphasize the stars of intellectual traditions, not the community. In addition, this work fails to relate the nature of ideas and their formal properties as an analytically separate network that can influence that of social relations.
10. Knowledge domains can entail millions of documents and many millions of citations. Thus, global visualizations are often infeasible because they are too big or too dense to interpret reliably (“multicolored spaghetti”). Knowledge domains and their networks are high-dimensional phenomena that often cannot be squashed down into two-dimensional representations without incurring a great deal of error (in addition to there being too many nodes and arcs to follow). Scholars have used at least three routines to reduce this complexity: (1) *limiting scope* (or distance) limits, (2) *pruning* unpopular artifacts, and (3) *clustering*. Limiting the scope of a graphical representation is perhaps the easiest—here, the user makes a single artifact the central focus and then looks at the network of relations say 1–2 steps outward from it. Scope-limits are often augmented by popularity-pruning (Chen, 2004), where users select to only show the relations to the most popular or most referenced artifacts (often a ranking is used so the top 10 or 30 most cited pieces, or the 10% most cited pieces). Both distance limits and degree pruning throw away some data in order to create visualizations of more local, central relationships in a knowledge domain. Hence, it is not surprising that scholars have sought other means to simplify and illustrate the structure of knowledge

communities, such as clustering (Newman 2004; Newman & Girvan, 2004; Morris & Yen, 2004). Agglomerative clustering methods typically identify sets of terms, authors, cites, or artifacts that have greater interconnection within a group than between them (an optimization of this is performed), or which identifies clusters of articles or authors with the most similar citation patterns (thereby clustering on similarity scores). In theory, these clusters represent research fronts (as a set of artifact), reference groups (as a set of co-cited materials—thought paradigm), or collaboration teams (groups of authors working together, see Morris and Yen, 2004, 5292). However, in large systems, clusters do not always have clear cutoffs, so the user can decide how coarse or fine-grained they want to characterize the field (see Newman 2004a figure). Some techniques are even useful for identifying where the greatest segmentation arises (e.g., Fourier analysis). Our point is that different options are available to simplify the complexity of knowledge so that useful, accurate visual representations can be had.
11. Even though pruning treats the density of ties, it does not solve the problem with citation data itself. Just as some persons and fields entail more collaboration than others, some authors cite more than others. Because of this, there may be unexpected outcomes that future research should identify and describe. One route we are exploring is to discern just how citations are used (e.g., whether as foil or as support, etc.), and to create a parsing code that will capture such variation and render it automatically for an online system.
12. Here, the KEDS study seems most relevant. There, parsing code was used to identify aligning words reflective of politics. One can readily extend that idea to aligning words before, after, or near certain citations, thereby reflecting the “type” of citation usage. In this manner, a more normative rendering of knowledge consumption could be had for citations.
13. We emphasize the communal input in this section, but are readily aware that most communities still have editors who prune unrelated or offensive postings, and start new threads of discussion. SPINE would clearly retain such a feature.
14. See Simmel (1971, p. ix-lxv) for a similar discussion of resolving dualisms in sociological theory.

References

Abbott, A. (2001). *Chaos of disciplines*. Chicago, IL: University of Chicago Press.

Abell, P. (2004). Narrative explanation: An alternative to variable-centered explanation. *Annual Review of Sociology, 30*, 287–310.

Allison, P. D., Long, S. J., & Krauze, T. K. (1982). Cumulative advantage and inequality in science. *American Sociological Review, 47*, 615–25.

Barabasi, A. (2003). *Linked: How everything is connected to everything else and what it means for business, science, and everyday life*. Cambridge, MA: Plume.

Batagelj, V., Mrvar, A., & Zaversnik, M. (2002). Network analysis of texts. Preprint series, Vol. 40 (2002), 833. Paper presented at the Third Language Technologies Conference, October 14–15, 2002, Ljubljana, Slovenia.

Berger, P. L., & Luckman, T. (1966). *The social construction of reality: a treatise in the sociology of knowledge*. New York, NY: Anchor Books.

Breiger, R. L. (1974). The duality of persons and groups. *Social Forces, 53*, 181–190.

Brin, S., & Page, L. (1998). *The anatomy of a large-scale hypertextual web search engine*. Paper presented at the Seventh International World Wide Web Conference, Brisbane, Australia.

Burns, G., Khan, A., Handeharizadeh, S., O’Neill, M., & Chen, Y. (2003). Tools and approaches for the construction of knowledge models from the neuroscientific literature. *Neuroinformatics, 1*, 81–110.

Carley, K. M. (1997). Network text analysis: The network position of concepts. C. W. Roberts (Ed.), *Text analysis for the social sciences: Methods for drawing statistical inferences from texts and transcripts* (pp. 79–100). Mahwah, NJ: Lawrence Erlbaum Associates.

Carolan, B. V., & Natriello, G. (2005). Data-mining journals and books: Using the science of networks to uncover the structure of the educational research community. *Educational Researcher,* April, 25–33.

Chen, C. (2003). *Mapping scientific frontiers: The quest for knowledge visualization*. New York: Springer Verlag.

Chen, C. (2004). Searching for intellectual turning points: Progressive knowledge domain visualization. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*.

Chen, C., & Morris, S. (2003). *Visualizing evolving networks: Minimum spanning trees versus pathfinder networks*. Paper presented at the meeting of the IEEE Symposium of Information Visualization, October 19–21, Seattle.

Collins, R. (1998). *The Sociology of philosophies: A global theory of intellectual change*. Cambridge: Harvard University Press.

Collins, R. (2001). Why the social sciences won't become high-consensus, rapid discovery sciences (pp. 62–84). S. Cole (Ed.). *What's Wrong With Sociology*. New Brunswick, NJ: Transaction Press.

Corman, S.R., Kuhn, T., McPhee , R., & Dooley, K. (2002). Studying complex discursive systems: Centering resonance analysis of communication. *Human Communication Research, 28*(2), 157–206.

Crane, D. (1972). *Invisible colleges: Diffusion of knowledge in scientific communities*. Chicago, IL: University of Chicago Press.

Dewey, J. (1916/1997). *Democracy and education: An introduction to the philosophy of education*. New York: Free Press.

Eakin, E. (2004). Study finds a nation of polarized readers. *New York Times*, March 13, 2004.

Friedkin, N. E. (1998). *A structural theory of social influence*. Cambridge, England: Cambridge.

Griswold, W. (1987). A methodological framework for the sociology of culture. *Sociological Methodology, 17*, 1–35.

Hiebert, J., Gallimore, R. & Stigler, J.W. (2002). A knowledge base for the teaching profession: What would it look like and how can we get one? *Educational Researcher*, *31*(5), 3–15.

Krebs, V. (2004). *Inflow Software* v. 3.0 (and http://www.orgnet.com/).

Kuhn, T. S. (1962/1996). *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Lagemann, E. C. (1989). The plural worlds of educational research. *History of Education Quarterly, 29*(2),185–214.

Lagemann, E. C. (2002). *An elusive science: The troubling history of education research*. Chicago: University of Chicago Press.

Levine, D.N. (1995). *Visions of the Sociological Tradition*. Chicago, IL: University of Chicago Press.

Lieberson, S., & Lynn, F. B. (2002). Barking up the wrong branch: Scientific alternatives to the current model of sociological science. *Annual Review of Sociology,* 28, 1–19. New Brunswick, NJ: Transaction Press.

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative riltering. *IEEE Internet Computing, 7*(1),76–80.

Locke, K., & Golden-Biddle, K. (1997). Constructing opportunities for contribution: Structuring intertextual coherence and "problematizing" in organizational studies. *Academy of Management Journal, 40*(5), 1023–1062.

Malone, T. W., Crowston, K., & Herman, G. A. (2003). *Organizing business knowledge: The MIT process handbook*. Cambridge, MA: MIT Press.

Mannheim, K. (1936). *Ideology and utopia: An introduction to the sociology of knowledge*. (Louis Wirth and Edward Shils, Trans.). New York: Harcourt, Brace & World.

Maxwell, J. (2004). Causal explanation, qualitative research, and scientific inquiry in education. *Educational Researcher, 33*(2), 3–11.

Mitchell, T., & Haro, A. (1999). Poles apart: Reconciling the dichotomies in education research. E. C. Lagemann & L. S. Shulman (Eds.). *Issues in Education Research* (pp. 42-62). San Francisco: Jossey-Bass.

Merton, R. K. (1968). The Matthew effect in science. Science, *159*, 56–63.

Moody, J. (2004). The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999.

American Sociological Review, 69, 213–238.

Morris, S. A. & Yen, G. G. (2004). *Crossmaps: Visualization of overlapping relationships in collections of journal papers*. Proceedings of the National Academies of Science, *101*(1), 5291–5296.

Munzner, T. (2000). Interactive visualization of large graphs and networks. Unpublished doctoral dissertation, Stanford University.

Nadel, S. F. (1957). *The theory of social structure*. London: Cohen and West.

Newman, M. E. J. (2004). Coauthorship networks and patterns of scientific collaboration. Proceedings of the National Academy of Sciences, USA. 2004 April 6; 101(Supplemental 1): 5200–5205.

Newman, M.E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. Physical Review E 69, 026113.

Novak, J.D. (1998). *Learning, Creating, and Using Knowledge: Concept Maps as Facilitative Tools in Schools and Corporations*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

Novak, J.D. & Gowin, D.B. (1984). *Learning How to Learn*. Cambridge, United Kingdom: Cambridge University Press.

Pea, R. (1999). New Media Communications Forums for Improving Education Research and Practice. In *Issues in Education Research: Problems and Possibilities*. Ellen Condliffe Lagemann and Lee S. Shulman, Eds. San Francisco: Jossey Bass.

Powell, W, White, D.R., Koput, K.W. & Owen-Smith, J. (2005). Network Dynamics and Field Evolution: The Growth of Inter-organizational Collaboration in the Life Sciences. *American Journal of Sociology* (January).

Richards, W. D. & Barnett, G. A. (eds). (1993). *Progress in Communication Sciences, Volume XII*. Norwood, NJ: Ablex Publishing.

Shriffin, R.M. & Borner, K. (2004). Mapping Knowledge Domains, *PNAS 101,* supplemental 1, 5183-5185.

Simmel, G. (1971). *Georg Simmel: On Individuality and Social Forms*. Donald N. Levine ed. Chicago: University of Chicago Press.

Simon, H.A. (1969). *Sciences of the Artificial*. Cambridge, MA: MIT Press.

Tilly, C. (2006). *Why? What Happens When People Give Reasons…And Why*. Princeton, NJ: Princeton University Press.

Turner, J. H. (2002). *Face to face: Toward a sociological theory of interpersonal behavior*. Stanford, CA: Stanford University Press.

Wasserman, S. & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

Watts, D. J. (1999). Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology,* 105, 493–527.

Cite This Article as: *Teachers College Record* Volume 112 Number 2, 2010, p. -
http://www.tcrecord.org ID Number: 12844, Date Accessed: 5/1/2008 5:02:08 PM