Available online at www.sciencedirect.com

ScienceDirect

Int. J. Human-Computer Studies 65 (2007) 659–673

International Journal of Human-Computer Studies

www.elsevier.com/locate/ijhcs

Beyond ontologies: Toward situated representations of scientific knowledge

William Pike*, Mark Gahegan

GeoVISTA Center, Department of Geography, 302 Walker Building, University Park, PA 16802, USA

Abstract

In information systems that support knowledge-discovery applications such as scientific exploration, reliance on highly structured ontologies as data-organization aids can be limiting. With current computational aids to science work, the human knowledge that creates meaning out of analyses is often only recorded when work reaches publication—or worse, left unrecorded altogether—for lack of an ontological model for scientific concepts that can capture knowledge as it is created and used. We argue for an approach to representing scientific concepts that reflects (1) the situated processes of science work, (2) the social construction of knowledge, and (3) the emergence and evolution of understanding over time. In this model, knowledge is the result of collaboration, negotiation, and manipulation by teams of researchers. Capturing the situations in which knowledge is created and used helps these collaborators discover areas of agreement and discord, while allowing individual inquirers to maintain different perspectives on the same information. The capture of provenance information allows historical trails of reasoning to be reconstructed, allowing end users to evaluate the utility and trustworthiness of knowledge representations. We present a proof-of-concept system, called Codex, based on this situated knowledge model. Codex supports visualization of knowledge structures through concept mapping, and enables inference across those structures. The proof-of-concept is deployed in the domain of geoscience to support distributed teams of learners and researchers.

Keywords: Knowledge representation; Situated cognition; Concept maps; Collaboration; Web applications

1. Introduction

Ontologies, as they are typically implemented in information systems, are often hierarchical and authoritative. These ontologies are useful formalizations in circumstances where formalization is called for, such as mapping terms between domains. But real-world cognition is often more fluid, flexible, and context-dependent than strict formalizations suit. Can ontologies properly capture the nuance of human knowledge? Does any single ontology reflect what is truly relevant to a particular application or domain, or is the ontology more a reflection of its creators' worldview than of neutral or common belief? Here, we propose that knowledge representations for computational environments should reflect the *situated* nature of human understanding. We also introduce a proof-of-concept application that enables distributed collaborators to share knowledge and data in a manner that is cognizant of the contexts and perspectives surrounding the creation and use of those resources.

We focus on scientific cognition as an example of human reasoning processes and on the domain of earth science as an application testbed. Like other domains, geoscience produces volumes of information, from hyperspectral satellite sensors to millions of words in journal articles. The scientific community's ability to generate new information—ever more detailed observations, about more diverse phenomena—tests its ability to turn these measurements into useful knowledge. The problem is not that there is no wisdom contained in the digital artifacts of modern science; the problem is the need to make efficient and effective use of increasingly complex descriptions of the world. Even within a single discipline, the variety of information sources and analytical methods brought to

---

*Corresponding author. Present address: Information Analytics Group, Pacific Northwest National Laboratory, MSIN K7-28, P.O. Box 999, Richland, WA 99352, USA. Tel.: +1 509 375 2689; fax: +1 509 375 3835.

E-mail addresses: bill.pike@pnl.gov (W. Pike), mng1@psu.edu (M. Gahegan).

1071-5819/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ijhcs.2007.03.002

bear on a problem can complicate assessing commensurability between researchers' approaches. The clearest picture of a problem might only be painted when diverse points of view are integrated into an explanation broader than any one alone could provide. The communication of conceptual models between collaborators is crucial to accomplishing this integration, especially in earth and environmental applications (Heemskerk et al., 2003).

The information science literature is rife with efforts to represent human "concepts" computationally, but concepts are often treated as category labels useful for integrating heterogeneous data sources. Computational data *contains* knowledge, to be sure, and it is used to create and apply knowledge, but that knowledge is not yet represented richly. As a result, information integration tasks are often data-centric; semantics are important to the extent that they support data interoperability, but the human knowledge and practices that guided the collection or use of that data remain implicit somewhere in its syntax or schema. A land cover map, for instance, says something about the place it depicts, although what it says to an individual researcher is either locked in the data, locked in the researcher's head, or described elsewhere in natural language text. In any case, it is not easily accessible to others who want to know how or why to use this information (say, to devise a new theory), or how it has been applied in existing explanations. For domains where meaning depends partly on the subjective perspectives of inquirers, a restricted view of what constitutes a concept does not do justice to the complexity of human knowledge structures. This complexity can be seen as a nexus of interconnected attributes (Fig. 1), all of which a knowledge representation should model to be more faithful to the nature of human understanding.

Our work approaches the problem of capturing, storing, and communicating scientific knowledge by treating science as a process of inquiry. Knowledge is constructed and applied during this process as observations are collected and manipulated, hypotheses generated and tested, and results transmitted and built upon. Here, concepts rather than datasets are the primitive elements of scientific inquiry, and these concepts are often in flux. This approach emphasizes interoperability of ideas, not simply data; it recognizes that the knowledge these ideas embody is by turns a shared and contested conceptualization, the result of collaboration, negotiation, and manipulation by teams of researchers. Whereas modern ontology is very much concerned with Aristotelian classification of terms, we contend that much knowledge is best modeled through representations of inquiry and interpretation. If systems can help represent concepts as evolving, cooperatively constructed, and experientially grounded resources, more effective integration of disparate observations into coherent explanations will result.

2. Merging top-down and bottom-up representations

There are two broad approaches to the problem of knowledge representation. The ontological approach is characterized by projects such as Cyc (Guha and Lenat, 1991), a top-down, authoritative encyclopedia. Ontological tools such as this focus mainly on enabling sharable underlying representations of knowledge and less on interfaces and supporting infrastructure to let collaborators construct this knowledge together. The alternative approach emphasizes the bottom-up, discursive nature of knowledge. This approach acknowledges the perspectives of collaborating inquirers (rather than an imposed

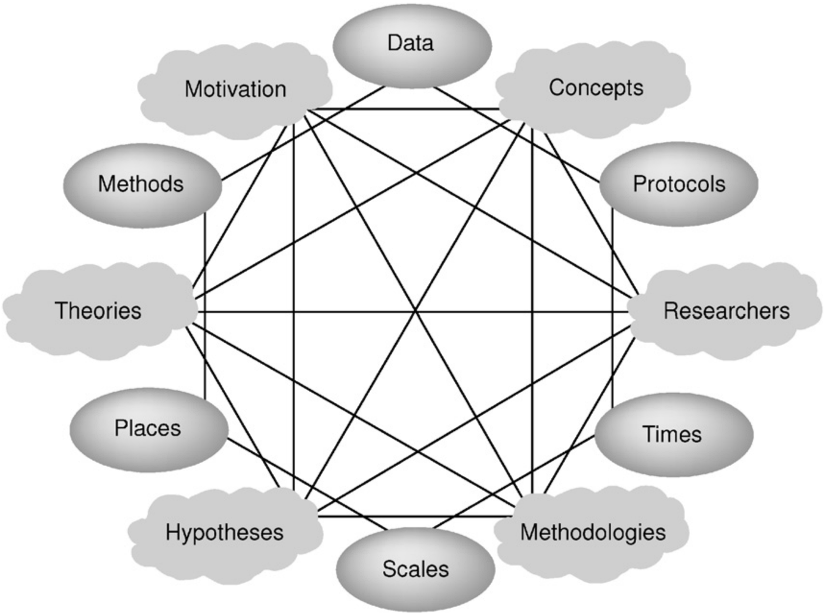


Fig. 1. Nexus of constructs concerning the development and application of scientific knowledge. Some (ellipses) are often made explicit in scientific reports or metadata, while others (clouds) are not; the latter, however, are crucial to understanding, communicating, and reusing scientific knowledge.

ontology) in defining concepts relevant to a community. The cooperative approach is evident in computer-mediated communication methods such as the Delphi method (Turoff and Hiltz, 1996), where the aim is to identify common understanding (or areas of disagreement) over time. Cooperative tools focus on effective interfaces to collaborative work, but often lack explicit semantics.

2.1. *Asserting knowledge from the top-down*

To the end user, strictly ontological approaches to knowledge sharing (e.g., Bozsak et al., 2002; Sugumaran and Storey, 2002) often appear as received wisdom: neutral concept structures isolated from the collaborative experiences of their creation and the fluid processes of their change. Indeed, there have been attempts at "national knowledge infrastructures" to be deployed across diverse domains and applications (Cao et al., 2002). Some ontological frameworks have been extended to represent the process of constructing natural science knowledge through experimental procedures described in existing text (Noy and Hafner, 2000).

There have also been some commendable recent efforts to examine the problem of ontology versioning and change as it relates to maintaining logical consistency (notably Klein et al., 2002), but a change in the way we express a concept (through its intension, extension, or relations) over time reflects something deeper than relabeling; the change indicates a shift in ourselves that necessitates modification to the interpretive stance others must take to understand our knowledge (Buzaglo, 2002).

It is possible that domains could agree on semantics by committee, but this approach is complicated by the diversity of opinion that should be represented rather than removed (Sowa (2000) writes that "independently developed, but convergent theories that stand the test of time are a more reliable basis for standards than the consensus of a committee"). Kazic (2000) suggests a middle ground, where domains create ontologies for only the most abstract, simple concepts. The simplicity of this approach is also its shortcoming: those ideas "most likely to engender controversy are left where they belong—as the private opinions of people, databases, or algorithms" (Kazic, 2000). But it is these controversial ideas, opinions, hypotheses, and theories that are often important to forming, evaluating, and modifying scientific explanations. Ideally, our ability to represent semantics computationally should not be reduced to the lowest common denominator upon which we can all agree. Disagreement may identify topics ripe for breakthrough.

There is now growing recognition in the knowledge representation field that its tools should reflect the situated work practices of their users (Schultze and Boland, 2000) and accommodate the dialogical, interactive nature of exploration (Nake and Grabowski, 2001; Dustdar, 2004). Marcos and Marcos (2001) argue that ontologies in information science are often treated as unassailable schema for "external" knowledge rather than as representations of shared knowledge with their own context and schema. Magnani (2001) suggests that *situatedness* is precisely what makes abduction a useful model for computer-based hypothesis creation—even under conditions of hypothesis failure, it produces useful information. And in light of the deductive nature of ontology languages like OWL, some have suggested a turn toward techniques that enable communities to test, refine, and implement emergent, rather than top-down, solutions (de Moor et al., 2002).

2.2. *Building knowledge from the bottom-up*

The bottom-up, cooperative approach to knowledge construction is characterized by the tools and methods of Computer-Supported Cooperative Work (CSCW). CSCW applications for scientific collaboration often take the form of electronic notebooks, organized into hierarchies of chapters and pages (e.g., Lysakowski and Doyle, 1998; Myers et al., 2001), in which researchers can enter and search for free-form records. The descriptive nature of CSCW relaxes many of the normative constraints of formal knowledge representations.

Recently, the CSCW community has begun to embrace ontologies as the basis for tools to support scholarly discourse. These ontologies describe the kind of entities that a CSCW system is capable of expressing (van Bruggen et al., 2003) and are typically considered generic containers for scientific work (e.g., Suthers, 1999), not as choices of perspective that are themselves contestable. Some CSCW methods, like ScholOnto (Buckingham Shum et al., 2000), develop discourse ontologies to express connections between researchers and published topics; ClaiMaker (Li et al., 2002) provides a structured vocabulary for articulating scientific claims and a mechanism to annotate documents so as to encode relations between text documents and the concepts expressed therein. However, the public record of science tells only part of the story, and not always faithfully; it does not reveal all of the analysis procedures, decisions, wrong turns, and intermediate results that underlie published work. Publications are a high-level mechanism for knowledge transfer within a large community, but much of the discourse relevant to science is inaccessible outside of the small groups in which it occurs. Practitioners in other places or times can have difficulty in reconstructing the discursive process that lead to a particular finding. Compendium (Buckingham Shum, 2006) is designed to capture these discourses—both personal and collective—in the contexts in which they occur through visual dialogue mapping. Our approach, articulated in the following sections, extends such methods by explicitly accommodating the recording of analysis strategies, integration of multiple complementary analyses, and long-term, automated tracking of knowledge construction. We return to related work in Section 6.4 to reflect on the relationships between our approach and the growing body of research on the computational support of analytical discourse.

2.3. The best of both

Between the ontological and the cooperative views of knowledge sharing lies an opportunity for combining formal representations of knowledge with the ability for communities to elicit and refine them over time. While ontological representations (at various levels of generality, from task to domain (Guarino, 1997)) have been described as no less than a "silver bullet" (Fensel, 2001) for information integration—and have been rapidly accepted in the information science community—there has been relatively little reflection on how conceptual structures emerge from practice and how they can reflect the evolving nature of that practice. In Engelbart's (1962) famous vision for human–computer cooperation, users not only manipulate complex concept and symbol structures, but systems also capture how these structures are flexibly linked across multiple roles. Unfortunately, in contemporary practice it is too often the experts in ontology who determine how to represent a domain, not communities of practitioners. The top-down imposition of an ontologist's domain model masks understanding of how practitioners themselves construct meaning in ill-structured scientific problems where there may not even be initial agreement about nature of the domain itself.

Formal ontology is widely practiced in domains such as medicine and business, but even where a community can (and ought to) agree on the structure of some aspects of reality—such as the mereology of the human body or the rules of financial transaction—there are inevitable differences between an individual's worldview and the concepts expressed in the ontology. When it comes to how and under what circumstances an individual inquirer applies those ontologies to a problem—instantiating concepts or asserting relationships as one makes a medical diagnosis or decides to buy or sell—the ability to manage flexibility, fluidity, and differential application is essential. In a domain like geoscience, characterized by elusive concepts such as hazard risk, explanations often derive from highly personal observations about the world (Frodeman, 2003). Ontologies intended to apply across the entirety of such domains may emerge from shared experiences, but can be difficult to agree on a priori without a mechanism for capturing and integrating those experiences. Methods from the cooperative work community can help address the bottom-up nature of knowledge construction but generally lack semantic richness; the expressions of meaning they produce are not routinely grounded in knowledge representations that allow concepts to be efficiently shared, searched, and reused in other problems or by other tools. The present study brings cooperative construction and emergence to ontologies, and richer semantics to cooperative tools. In our case, the discourse central to achieving collective understanding is implicit in the re-use of shared data, concepts, and analysis procedures. As an alternative to using predefined concepts or relations to describe a domain or the discourse within it, we aim to allow researchers to describe their domain understanding in their own terms, while encouraging them to find, apply, and modify knowledge resources created by others as a means of creating shared understanding.

3. The construction of knowledge

Knowledge involves the ability to apply information—consciously or otherwise—to solve a problem. Knowledge is not just the accumulation of facts, but also the experiences and reasoning facilities that guide the use of those facts. Central to this study's approach is an effort to explicitly preserve representations of utility—the "how" and "why" that create knowledge out of information.

3.1. The ingredients of knowledge

We define three components that are required to represent knowledge in the context of its use: concepts, metadata, and situations. The *concept* expresses the existence of an abstract category and encompasses everything in its extension. A given concept may have different names in different circumstances while preserving the same underlying meaning (its intension).

There are a number of basic views on the form of mentally held concepts. Classical Aristotelian theory holds that concept membership is defined through a set of singly necessary and jointly sufficient conditions, although consistent definitions are scarce in practice and limit the capacity for conceptual change (Wittgenstein, 1953). In dealing with these shortcomings, the probabilistic approach to conceptual structure suggests that some concepts are better examples of their category than others. This approach finds support in empirical classification studies (e.g., Rosch, 1975; Rosch and Mervis, 1975) that indicate humans classify by assessing an observation's proximity to prototypes. A "conceptual atomist" approach defines concepts through relationships between mental symbols and the objects they represent, a position in which concepts themselves are devoid of intensional definitions (Fodor, 1998).

A key problem for the present work is to implement a model for concepts that can be communicated efficiently across the human–computer interface. Although concepts may be difficult to define internally (that is, their intension may be vague), it is possible to describe them in terms of their relationships with other concepts, and by how they are applied. In this vein we can model concepts in terms of their perceptual-functional affordances (Tversky, 2005)—in our cases, these affordances are the roles a concept plays or the capabilities it enables. Representing concepts' functional roles in a larger knowledge structure is important to depicting "how" and "why" in scientific reasoning.

In our model, each concept is wrapped in metadata that consists of the attributes that can be recorded regardless of how or why a concept is used: who created it, using what tools, at what time and place, and so on. In the process of inquiry, concepts are selected based on relevant criteria or roles and linked together into larger structures. These acts of conceptual manipulation have been described as *situation* (Solomon et al., 1999), the bringing together of background information and current observations and analyses toward some goal. Situation is important to knowledge representation because it explicitly reproduces the enactment that is part of selecting and reasoning with a set of concepts (Barsalou, 2002). Lemke (1997) calls situation an "ecology," a term that evokes the dynamic interaction between concepts and thinkers in the process of knowledge construction.

Situation, then, encompasses the coordinated activity that is directed toward some goal. A given concept—we might think of it as a node in a conceptual network—can be reused in different circumstances, but there will be some information we want it to carry with it regardless of circumstance (this we have called metadata), and some that will be unique to the role it plays in a particular case (this we have called situation). To denote the particular choice of concepts, metadata, and situations that a particular thinker (or community of thinkers) uses to describe a process, problem, or phenomenon, we can use the term *perspective*. In Fig. 2, situations A and B might represent different perspectives on a problem taken by two thinkers. Each might use concepts from the same body of shared understanding but see them as being directed toward different explanations. For thinker A to appreciate B's perspective, it is necessary to reproduce for A both the entities that are relevant (the concepts and contexts) as well as their surrounding situation (the directed aim of B's reasoning). This work creates an infrastructure for achieving that reproduction and a visual mechanism for depicting concepts and situations.

4. Implementing a situated knowledge model

The convergence of the notions of situation, community, and scientific process can motivate more complete models for knowledge-centered computing than are currently available. In this section we describe the basic architecture for a system to support collaborative and situated knowledge representations. To represent our thought structures, we should not start by adopting a language like OWL, asking only "How can we represent our ideas using this structure?" Instead, we should ask, "What are the fundamental characteristics of knowledge-based work, and how should the design of computational tools follow from them?"

We have developed a Web-based application called Codex as proof-of-concept platform for capturing and sharing situated knowledge representations. As part of a suite of knowledge sharing tools, Codex has been described elsewhere in MacEachren et al. (2006). (The name Codex follows from the names given to manuscript notebooks such as da Vinci's, which are archetypal examples of observation and theory interwoven into a record of knowledge construction). Codex uses a portal model to organize distributed resources—data, concepts, and collarators—under a common interface.

Codex is principally a knowledge-sharing medium. It operates at a level of abstraction different from that of CSCW tools intended to support data-sharing; while Codex allows data files to be stored and linked together, data are described foremost by the human concepts they signify. Codex also builds on online scientific workbenches (e.g., Stevens et al., 2003) that emphasize data integration for automated analysis; Codex treats problem-solving as an issue of human consideration and interpretation. Typical scientific problem-solving environments might integrate several analysis tools to support hypothesis generation (Sanchez and Langley, 2003), but fail to capture the knowledge embedded in the manipulation of resources.

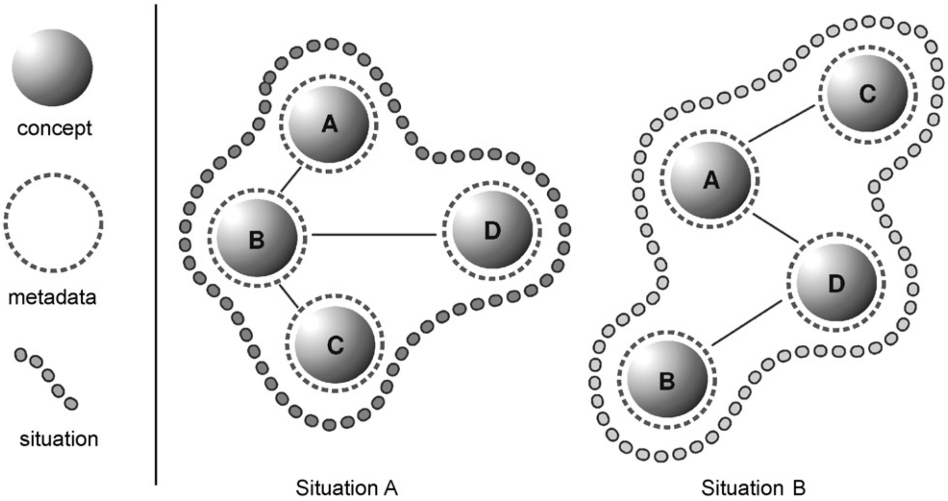


Fig. 2. Metadata in our model describe the circumstances surrounding the creation or use of an individual resource or concept (a node in a conceptual network); situations describe the circumstances of larger knowledge structures arising from the different ways these nodes can be connected.

Codex is at once a CSCW tool that enables rich semantic descriptions, and a semantic markup platform that relaxes the constraints of common ontological approaches. Codex is built around the concept of workspace. Workspaces can be both private and communal. Each Codex user has a personal workspace to store his or her ideas, data, and hypotheses. Researchers can move resources to shared workspaces where they can be accessed, applied, or modified by collaborators. The cooperative, Web-based nature of Codex means that one user's insight can be made immediately accessible to colleagues a world away.

The researcher logging in to Codex is first presented with a nexus-like view onto the workspace (Fig. 3). This starting point groups resources together under a set of default categories, providing quick access to the basic units of an investigation. From the workspace home page the researcher can rapidly upload a file, look in on collaborators, or describe a new analysis. Six types of resource are supported on this page, although these entry points can be supplanted with user-defined categories.

- *People*: The individuals and groups who create or apply resources accessed through the Portal. Each person maintains a profile that can communicate elements of his or her background and expertise.
- *Concepts*: Descriptions of abstract ideas, such as "flood" or "earthquake".
- *Files*: Binary data that express something about a concept. Files could include spreadsheets, text documents, images, audio clips, maps, or other data formats (quantitative or qualitative) that connect observations or measurements to the cognitive structures represented by concepts.
- *Tools*: The methods used to analyze data and to construct instantiations of concepts from data. Tools could include statistical tests, visualization methods, predictive models, interviewing instruments, and so on.
- *Places*: Geography is fundamental to integrative research, and places help researchers define the locations and scales under study. Place also helps to account for differences in epistemology between researchers.
- *Tasks*: People, concepts, files, tools, and places are linked together through tasks that might describe a workflow process, an experimental procedure, or a problem-solving approach.

A nexus view into these categories underscores the preeminence of interrelations between resource types in Codex; knowledge structures express the interdependence of data, analyses, hypotheses, and results. For instance, Codex might reveal a thread showing how a *file* was produced by a particular *person* as a step in a *task* aimed at describing the relevance of a *concept* to a *place*. Codex emphasizes that the act of making connections is central to knowledge construction.

Codex is built on a two-layer client–server design (Fig. 4). At Codex's core is a set of server-side applications that manage the basic functionality for maintaining a shared knowledge base. By enforcing a separation between core functionality and client interfaces, the set of services that manages situated, perspective-based representations of scientific knowledge can be used by multiple applications at once. These applications can all be thin clients, leaving the server to do the heavy lifting of storage and inference while the clients provide interfaces or extensions to the core functions. This architecture also lets users interact with the same knowledge base through a variety of interfaces,

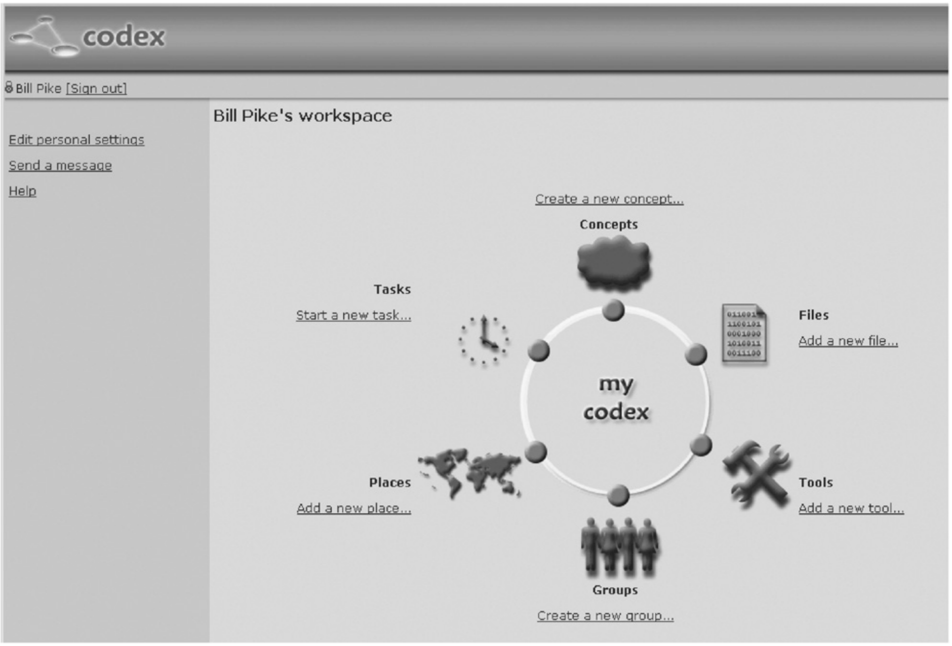


Fig. 3. The home page for a Codex user's workspace.

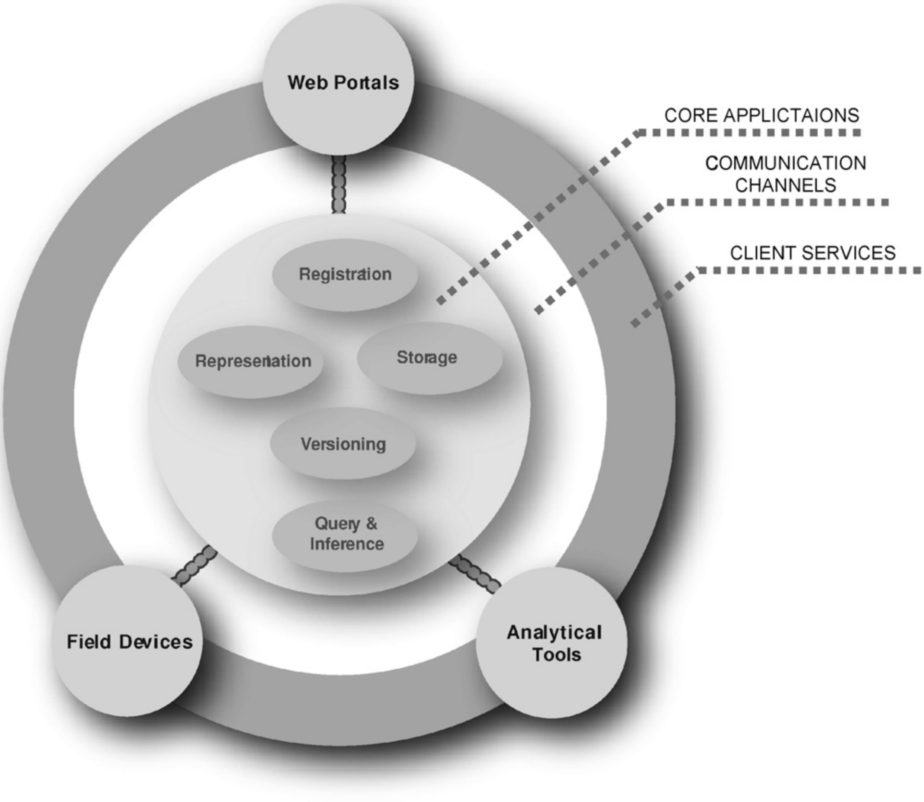


Fig. 4. Codex architecture.

supporting the insight that can come from examining information in different formats. Third party clients can also be developed over time that layer domain-specific views over the same underlying applications (for instance, a Codex customized for geoscientists). Currently, the Web portal has received primary development attention, although some work on enabling access to Codex functionality through mobile devices has also been conducted.

There are five components to the Codex server:

- *Representation*: All resources in Codex are described in the OWL-DL markup language, to make them interoperable with other knowledge representations on the Web and to allow Codex users to easily borrow from and extend imported representations. However, the OWL statements that are composed by the system in response to user activity are not full ontologies; they are simply snippets that reflect whatever understanding the user is currently expressing. Users never come face-to-face with OWL, as the GUI masks implementational details.
- *Storage*: Concept and workspace files are stored as OWL files to make them accessible to Semantic Web crawlers that index knowledge resources. Since the knowledge structures built with Codex are intended to be shared and reused, it makes sense to ease third party search and retrieval. Representing workspaces in OWL (that is, not only the content of a workspace, but the definition of the Codex workspace itself) means that researchers can use any third-party, OWL-compliant tool to access and manipulate their personal resources. The storage module is also responsible for maintaining links to external resources, such as referenced data files.
- *Registration*: Each resource belongs to one or more workspaces, which represent sandboxes in which individuals and groups work. The registration module tracks the assignment of resources to these workspaces and the promotion of resources from one workspace to another (e.g., from a private space to a shared space as a resource gains approval by, or is ready to be exposed to, a wider community). The registration component also validates user permissions, determining what resources should be exposed to a user or a search engine (keeping hidden, for instance, those tentative knowledge structures that investigators do not yet wish to share).
- *Query and inference*: Storing resources is one thing; finding them is another. The query module allows users to find resources based on their relationships to concepts, people, or observations. For instance, a user interested in a particular idea can query for tasks in which it was used to find out what roles it has played.
- *Versioning*: Codex resources change through use, so it is not possible to keep just one copy of a resource for all collaborators to share; this is a primary distinction between a traditional ontology and a situated knowledge

representation. Every user will make slight modifications to align concepts with his or her perspective. For each modification, Codex transparently creates a new version of the resource that contains a reference to its immediate predecessor (or predecessors, if it was created by merging properties from several resources). Audit trails thus emerge that show how a resource came into its current form, and users can follow these ancestral paths to evaluate each others' ideas. The versioning module also provides the facility to revert to a previous version of a resource while preserving a record that changes had been made but reversed. These tentative paths that were later abandoned can still be informative.

The modular architecture of Codex leaves open the possibility of integration with other components of a knowledge representation infrastructure. For example, Codex could serve as the knowledge management node on a broader network, handling the capture and communication of explanations that emerge from manipulating information in a larger online workbench.

5. Modeling knowledge in codex

The Concept is the universal set in Codex. The six resource categories introduced above represent five specializations of Concept: file, group, place, task, and tool. The sixth, concepts, contains direct members of the class Concept (such as "chair"). The reason for this top-level category is that it allows certain rules to be instituted for the format of conceptual information and simplifies reification of resource collections as instances of another resource type (for instance, a file, place, and tool can be gathered and reified as a task). Each of the specializations extends the default Concept with unique characteristics; the file subset, for instance, adds properties for file location, size, type, and so on; the place subset accommodates attributes like place name and geographic coordinates.

The use of Concept as a universal quantifier also places Codex's knowledge model in explicit opposition to contemporary style. Conventional wisdom holds that the ontology is the top-level category, the container for all knowledge represented computationally. This "ontology-first" model is highly restrictive; it presupposes a structure where there may be none (indeed, the act of manipulating resources in a tool like Codex might be for the very purpose of devising this structure, so imposing one at the start would be futile). Given the emphasis of the present work on structuring knowledge from the bottom up through pragmatic use, Codex takes a "Concept-first" approach.

Codex maintains one-to-many relationships between Concepts and situations. A situation is an arbitrary group of resources and the relationships that connect them; a given resource can be used in any number of different situations. Codex does not explicitly represent situations through a special set of tags. Representation of situational meaning is instead query-driven, based on the semantic and contextual relationships already stored within resources. Codex supports two varieties of situation, *user-defined* and *inferential*. A user-defined situation is formed on the basis of resource selection and/or definition by an individual. For instance, in the course of defining an "Earthquake risk" concept, the researcher might:

(1) Define two new concepts, "Earthquake risk" and "Distance decay."
(2) Find an existing concept, "Geographic area" and create a new instance of it, "Fault zone."
(3) Relate "Earthquake risk" to "Fault zone" through a distance decay property.

There is now a situation that contains a small set of concepts and relations. Should another user query for "Fault zone," Codex can show that in one situation, a fault zone is a geographic area prone to earthquake risk.

Inferential situations result from detecting relationships between the contextual elements of the resources in a given set. What makes them special is that they do not require resources to have any predefined semantics; relationships between resources are inferred on the basis of co-occurrent context attributes (Langley et al., 2002). Codex allows users to search for inferential situations over sets bounded by (1) the resources contained in a given workspace or (2) the result set of any query over a larger set. Inferential situations can be useful for spurring hypothesis generation by presenting candidate knowledge structures to the user. These structures are not ones that he or she created, but that represent other ways in which the resources in one's workspace could be connected.

Exposing inferential situations in Codex is a variation on collaborative filtering, a self-organizing approach to detecting ad hoc relationships (Ansari et al., 2000). Collaborative filtering is typically used to produce recommendations in e-commerce environments—for instance, based on the books for which a customer has expressed a preference, an online bookseller will recommend other books by examining what others who share those preferences have read. The same connections are possible in Codex, where an inferential situation can be built around any subset of resources that share a time, place, creator, and so on. The collaborative nature of Codex means that one can look across a user community to find relevant resources. In the simple case, Codex could build a situation around a user's query for resources that his or her collaborators created within one day of the time a target Concept was created, creating a situation of temporal association, or could recommend one researcher's "Seismology" resource to someone who already uses that person's "Earthquake risk" concept.

To understand how situations are presented to the Codex user as *perspectives* on an information space, consider the children's toy where a colored plastic lens is passed over a complex background; suddenly a pattern

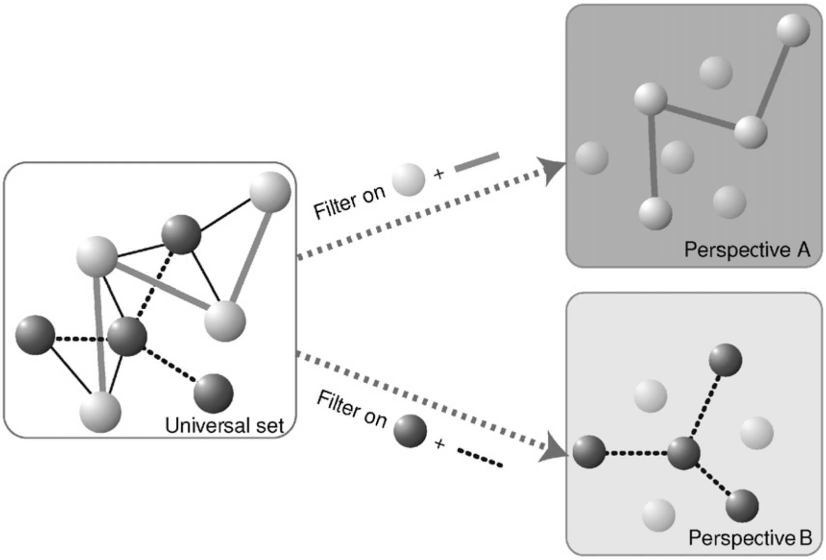


Fig. 5. Perspectives filter an information space according to particular situations. Perspectives A and B preferentially select different types of resources and relations; the ability to view perspectives can show how someone else made sense of a given set of resources.

appears, often a word or image in answer to a riddle. The lens absorbs certain wavelengths of light while permitting others to pass through. The result is that some of the complexity of the printed background is obscured, allowing only the salient elements—those that are compatible with the composition of the lens—to be seen. A Codex perspective works on the same model (Fig. 5). The perspective filters out some information, revealing only certain "wavelengths" of meaning that conform to the resource types present in a given situation. The remainder of a concept space is masked.

To examine a set of resources from different perspectives, the Codex user foviates on a resource and queries for the situations in which that resource is found. Codex can combine situations to either restrict or expand the selection of resources salient to a perspective.

- In the *union* of situations, the researcher finds the bounds of a problem space, given by the complete set of resources that a community deems relevant to it.
- The uniqueness of a particular perspective is found in the relative *complement* between situations. That is, the uniqueness of a researcher's perspective can be described as the set of resources that are in the situation through which he or she describes a problem, but that are not found in anyone else's. Taking the complement of a perspective can also reveal areas of disagreement or uncertainty, where concepts in one user's perspective fail to correspond to those in others'.
- The resources and relations in an *intersection*, the consilient set, constitute a new situation that represents the points of agreement within a community. Codex uses the consilient set as the basis for expressions of community or domain belief that might qualify to be used elsewhere as top-down knowledge structures (e.g., ontologies).

Although perspectives can be compared and integrated in Codex, Codex does not mandate the use of a "neutral" ontology as might be the case with other tools. A neutral ontology amounts to a mapping vocabulary that regulates interoperability between terms. In Codex, a common vocabulary can be *discovered* if one exists, but it is not an a priori requirement to describe knowledge.

5.1. Creating and using situations through concept mapping

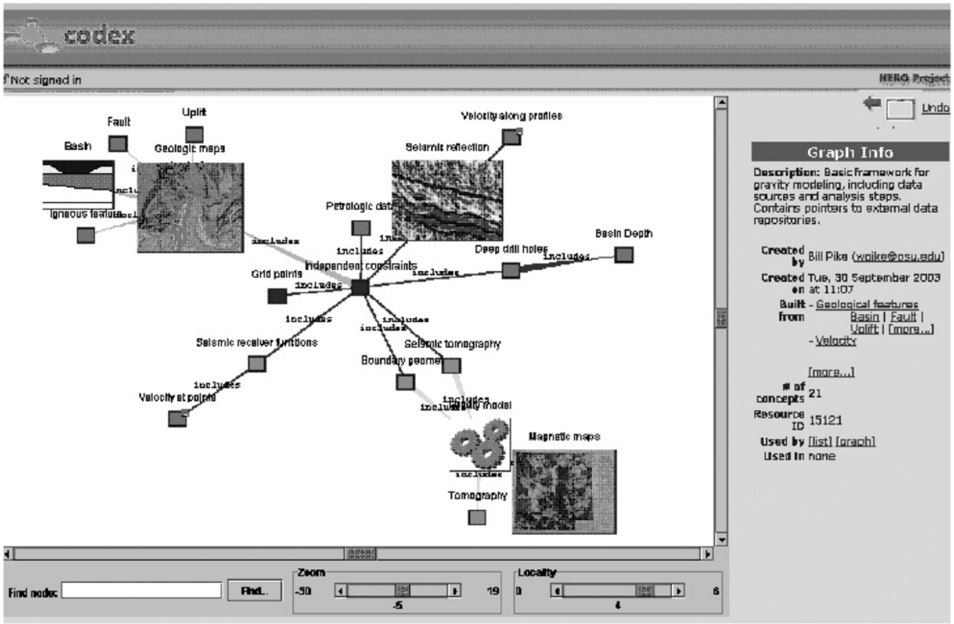
Codex's primary user interface is a concept mapping utility developed around an open-source dynamic graph browser.[1] Fig. 6 shows a sample Codex concept map for a set of information resources related to the notion of seismology (the examples in this section and the next are drawn from users in the GEON project,[2] a large cyberinfrastructure project for the geosciences). In the client, OWL classes or individuals are depicted as nodes; relationships between nodes (their properties) are edges. The semiotic nature of concept representations in Codex is seen here in the form of iconic representations for selected nodes. For instance, the concept "Seismic reflection" (which this user has chosen to signify through a particular data set) stands for a particular reflection profile, which is pictured. The graph depicts one situation in which the resources it contains are found.

As a user draws a graph, he or she is creating a situation for its set of resources. When the user adds a copy of an existing resource (whether as a node or an edge), Codex now contains multiple situations for that resource: the situation in which it was originally created, and the new situation in which it is being applied. Each of these situations constitutes a different perspective on the same

---

[1] TouchGraph: http://www.touchgraph.com/.

[2] http://www.geongrid.org.



Graph Info

Description: Basic framework for gravity modeling, including data sources and analysis steps. Contains pointers to external data repositories.

Created by Bill Pike (wpike@osu.edu)
Created on Tue, 30 September 2003 at 11:07
Built from - Geological features
Basin | Fault | Uplift | [more...]
- Velocity
[more...]
# of concepts 21
Resource ID 15121
Used by [list] [graph]
Used in none

Fig. 6. Codex concept map client.

resource (either the perspectives of different researchers, if the resource has been used by multiple people, or the various perspectives that an individual scientist adopts when thinking about the resource in different situations). In facilitating reuse of existing resources, Codex enables users to extend their understanding by borrowing ideas from collaborators. This mechanism is a variation on the network elaboration technique (Eckert, 1998), which has been shown to be especially useful in pedagogical settings: those learning a new domain can start with a simple structure provided by an instructor, text, or colleague and gradually extend it with new information as their learning progresses.

A simple example shows how different perspectives are displayed in the Codex concept map client. Suppose a geoscientist has created a concept map describing the domain of seismology (Fig. 7a). In this example, the graph represents a taxonomy and contains only concepts, so we would say that this situation is directed toward describing the *structure* of a domain. Now that the user has described this structure, he or she is interested in finding other situations for one of the concepts it contains, ''seismic velocity.'' (It is possible to find situations for a set of resources of any size, but for simplicity we will use a single node). The user selects the seismic velocity node and queries for a ''task'' perspective. Codex searches for situations in which a task has been described that include this concept; Fig. 7b reveals such a task. Now the user has gained a new perspective on the concept of seismic velocity, seeing it situated in a ''how-to'' network. Next, the user might want to know which collaborators have also used this seismic velocity concept; executing this query produces Fig. 7c. In this case, two of the ''users'' are in fact groups that have applied the concept in shared workspaces, so the researchers shown use the resource indirectly though their membership in the groups. Finally, the geoscientist wants to know what data files contain information on seismic velocity. Running this query produces the perspective on seismology shown in Fig. 7d; here, nodes represent instances of files that other researchers have included in the extension of seismic velocity. While the first two situations are user-defined (they only exist if a Codex user has built them), the latter two are inferential—they can be built automatically. This depiction of a single concept in multiple situations helps achieve the goal of knowledge representations that support interpretation and inquiry, not just classification.

Codex users can search for resources and situations to reuse both syntactically and semantically. A keyword search can query across workspaces for resources that have been marked as shareable, and a researcher can drag a copy of any of these resources into his or her own workspace where they can be instantiated or modified, always preserving their relationship to their progenitor. A semantic search might ask for other resources that are connected to a concept of interest by a given property; because resources are represented as sets of statements, one can query by those statements' subjects, predicates, objects, or some combination. The inquirer can traverse the networks of resources in which a search result is situated, gaining perspective on how that item was applied and knowledge structures emerged around it.

By enabling users to quickly shift perspectives on a problem, the visual display of information in Codex can change to correspond with a user's cognitive focus. Rather than show all possible relations between seismic velocity and any other resource, the perspective filters facilitate interpretation by displaying only the nodes that are salient

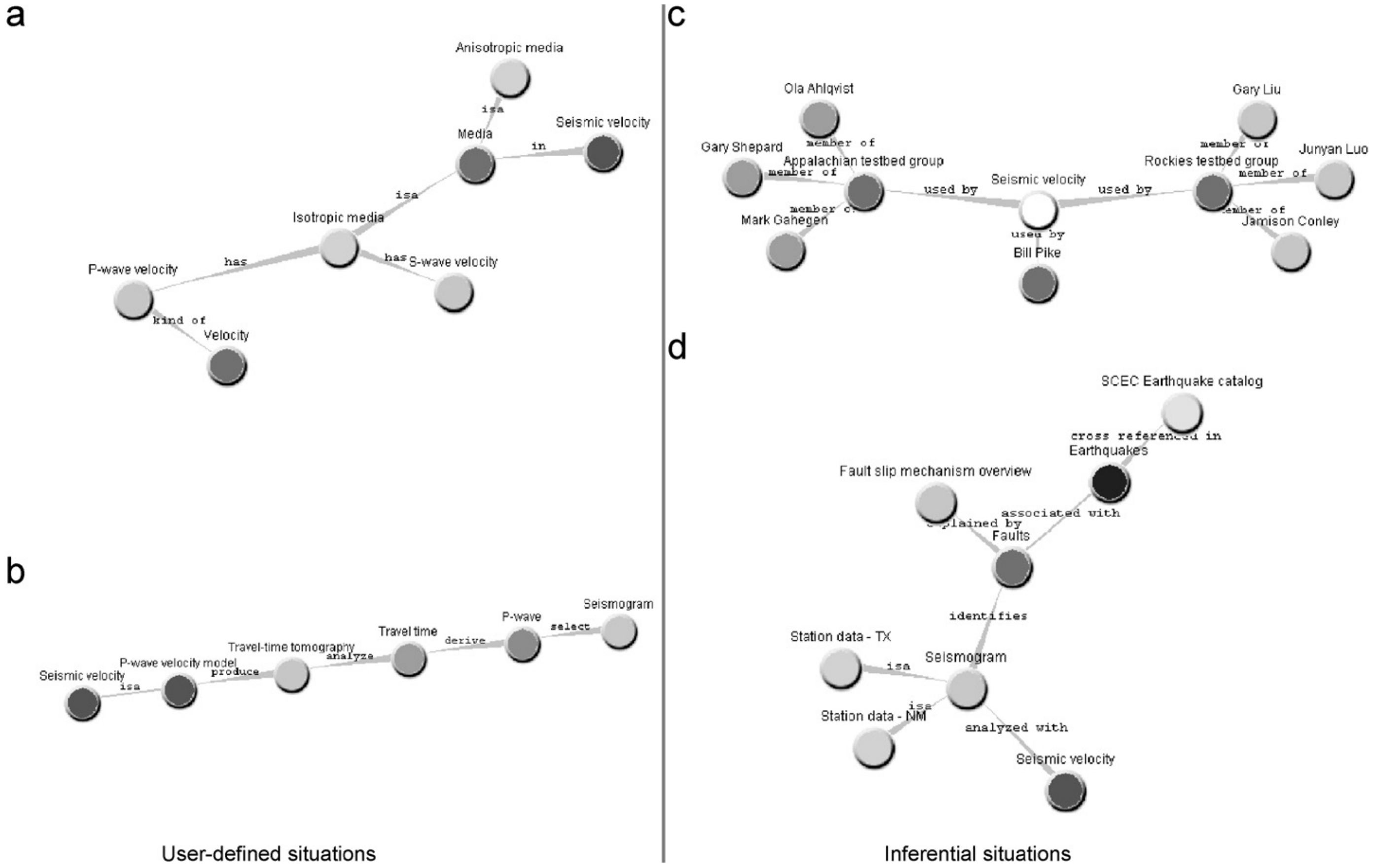


Fig. 7. Four perspectives on a "seismic velocity" concept. (a) Intensional concept structure. (b) A task that describes how seismic velocity can be measured. (c) A social network built around users of the concept. (d) Data resources that have been used to describe seismic velocity.

to the chosen situation. Changing the visualization to suit the user's goals has been shown to increase the efficiency and effectiveness of interpretation (Neuwirth et al., 1998); in light of the fluid, adaptive nature of problem-solving, enabling this sort of change can quickly communicate multiple approaches on a problem to a thinker coming to grips with its complexity (Chung et al., 2003). The Codex concept map client and its mechanism for navigating across perspectives is a step toward increasing the transfer of understanding among collaborators.

5.2. *Combining perspectives and versioning to track knowledge evolution*

When versioning histories and registration of concepts with different perspectives are combined, it becomes possible to present a single view onto the evolution of a resource. Fig. 8 shows such a view as retrieved by Codex, summarizing the intensional and extensional changes in the geoscience "Depositional environment" concept. The original concept, at upper left, was created by one sedimentologist and contained a single extensional element. As successive researchers adopted the concept, the connections they made between this concept and those in their personal ontologies were added to the concept's intension. For extension (nodes with document icon), practitioners are permitted to use different data to signify the same concept, or the same data to signify different concepts; registration between data and ontology is made flexible through a versioning mechanism that allows users to modify concepts over time without affecting the other users of those concepts. As a result, resources can fluidly migrate between personal and shared conceptualizations. A common resource might also branch into alternate versions that are used by different communities, while retaining references to their shared ancestry. Whereas a mechanism this flexible will likely give rise to many versions of concepts, we see this as an accurate reflection of the social process of knowledge construction and refinement that we seek to facilitate. How such a concept versioning mechanism should be maintained, visualized, and administered is an ongoing research question.

6. Issues in adopting situated knowledge-sharing tools

6.1. *Bootstrapping*

A tool like Codex will only be effective if it demonstrably increases the efficiency or quality of its users' work. But bootstrapping a knowledge-sharing system with enough cases to enable useful queries and reliable inference from the start is problematic. In our case, for geoscience users we

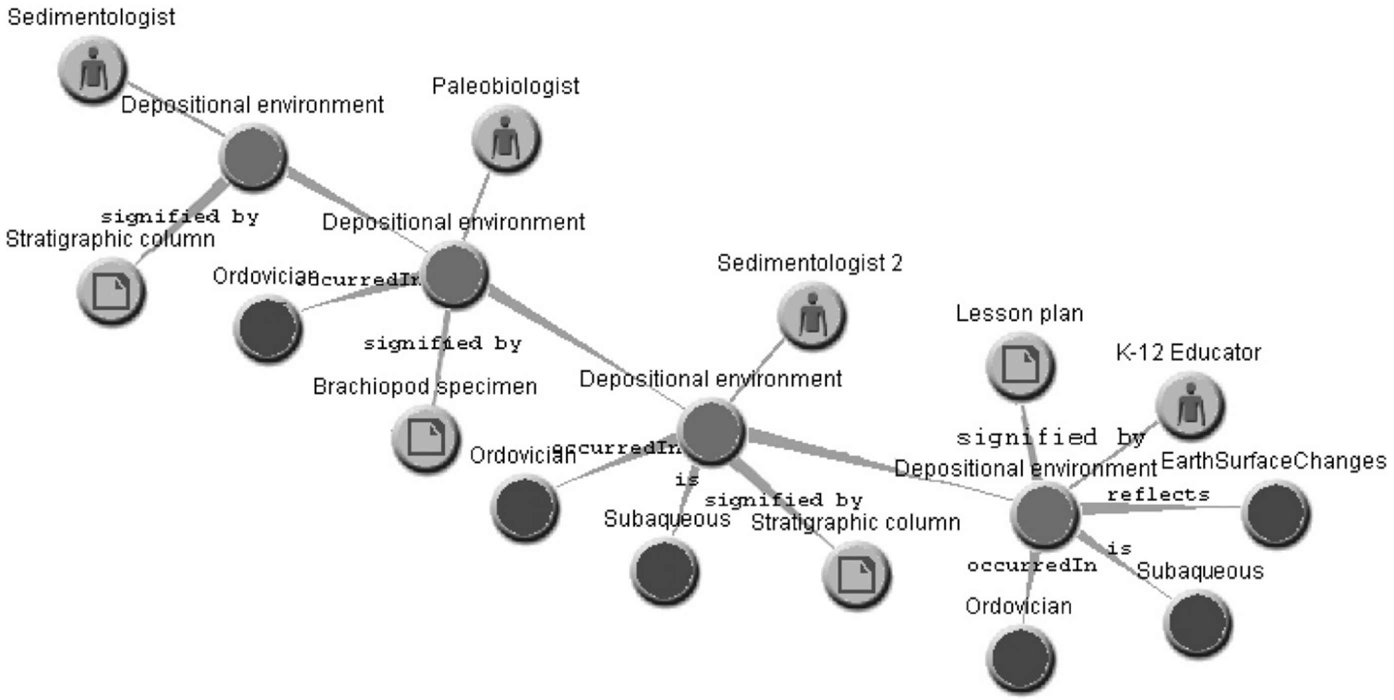


Fig. 8. Evolution of "Depositional environment" concept through use by different researchers over time, progressing from upper left to lower right.

have preloaded a selection of earth science concepts (drawn from sources such as NASA's SWEET[3]) that can be the basis for borrowing, modifying, and instantiating concepts in their own workspaces. But the benefits of knowledge sharing only truly accrue through use, so more "personal" bootstrapping is desirable. Because it is unreasonable to expect users to pause every few minutes to update the system on their current thinking, the knowledge-collection costs imposed by tools like Codex can be reduced if concept capture is embedded in users' everyday workflow. Recent research into detecting knowledge communities using Web links (Henzinger and Lawrence, 2004), extracting entities and relationships from a researcher's existing documents, or tracking the electronic resources that come into one's workspace (Dumais et al., 2003) may provide usable fodder for a knowledge-sharing environment.

6.2. Cultural barriers

The professional culture of many scientific domains creates impediments for a knowledge sharing system such as that proposed here. When accolades, publications, promotions, and other measures of status are often attained by carefully protecting one's own intellectual property, a fundamental change in science could be required to create a culture of open sharing. Codex can be a victim of this culture if users choose not to share, or share selectively, but it can also be a force for change. The reward of instant access to rich bodies of reasoning and reusable knowledge artifacts can motivate users to contribute. But to preserve a sense of ownership, we are beginning to explore the integration of licenses from Creative Commons[4] into knowledge resources; such licenses can give resource authors some control over how their creations are later applied. The ability to share knowledge structures in Codex corresponds with wider efforts to create new publication models (e.g., Martens et al., 2003, DSpace[5]). The repeatability of studies can be enhanced when the analysis process is described as a workflow that can be imported into a researcher's workspace, modified, and re-applied.

6.3. Evaluation

It can be difficult to measure the efficacy of knowledge-sharing tools because we cannot know what insight users would have gained without them—but we can look at uptake in user communities to see how adopters come to view these tools as necessary technologies. Among one Codex user group, for instance, our system replaced an earlier "notebook"-style data sharing tool (the genesis of Codex as a response to alternative knowledge sharing techniques is described further in (Pike et al., 2005)). When we examined the logs of the notebook tool, we noticed that the predominant behavior of users was to upload files (because they thought they might be useful to their collaborators on a project). Very few ever downloaded those files, so the contributed data would appear to have made little difference to collaborators' knowledge con-

---

[3] http://sweet.jpl.nasa.gov.

[4] http://www.creativecommons.org.

[5] http://www.dspace.org.

struction. Anecdotal evidence indicated that the reason for this one-way flow of information was that the data files, even with some descriptive metadata, were of limited utility; what collaborators needed to know was how the data were used, or what conclusions were drawn from them. With Codex deployed, we noted greater parity in the ratio between the average user's original contributions to and selections from the knowledge space they shared with collaborators. This finding is in line with Gray and Meister's (2004) theory of "knowledge sourcing": instead of measuring knowledge-sharing success from the supply-side by looking at the amount of information stored, we should look at demand-side measures of the extent to which researchers intentionally access and reuse each other's knowledge. Thus it is not the size of the knowledge base that matters, but the ability of users to find out what they need to know.

A group of students from several universities around the US used Codex to share descriptions of human–environment interactions observed in their local regions; the graph structure of knowledge maps in Codex has proved appealing to them since it can easily show complex connections that might be otherwise difficult to articulate. The students actively borrowed relevant concepts produced by others to use in their own depictions, building understanding collaboratively. Students reported that the visual reasoning facilities provided in Codex encouraged them to look for intersections between their own developing understanding and that of others. From a pedagogical perspective, the ability to recover resource histories has been particularly useful. Students working to synthesize knowledge across teams could explore the process of a resource's construction (as in Fig. 8), learning how others developed knowledge over time. This insight into the knowledge creation process offered the students insight into the "how" and "why" that we have earlier argued is central to scientific collaboration but that is otherwise difficult to capture, let alone communicate.

Codex was developed to support asynchronous collaborative exploration, but we found that students began using it synchronously, sometimes even seated together manipulating a shared conceptual graph. The original intent of the tool's deployment in this student group was to encourage them to record the concepts and relations that emerged from their exploration of human–environment interaction data (including maps, interviews, and documents). By providing a mechanism for the long-term archiving of their insight, it was hoped that changes over time would be more readily noticed, and that future researchers would be able to recover both the data and the reasoning that were the basis for a finding. While this mission stayed true, students began to use the tool in support of very short-term interaction, almost as a sketchpad to help communicate their ideas to each other. However, the ability to express the connections between their data (by importing it into the map, as is shown with some nodes in Fig. 6) and their abstract concepts made Codex more useful to them than a whiteboard or similar interactive device.

6.4. Related work

Codex represents a growing recognition in the Computer-Supported Cooperative Work and scientific cyberinfrastructure communities that computational aids to science work must incorporate richer representations of the reasoning process. As more examples of systems that support scientific introspection and communication appear in the literature, our hope is that scientific communities will increasingly see these systems as necessary to continued scholarly advancement. For its part, Codex encourages increased reflection on the structure of knowledge by explicitly representing the alternate perspectives that researchers can take on the structure of their reasoning.

There have been other notable recent efforts to build visual interfaces that help collaborating investigators fuse perspectives (Brennan et al., 2006), although Codex provides the additional capability for analysts to work with multiple perspectives at once and to explore machine-generated perspectives based on semantic overlap in the underlying OWL representation. The Biological Sciences Collaboratory (Chin and Lansing, 2004) provides a data-centric view into scientific analysis and interpretation, and includes graphical depictions of provenance, workflow, and versioning, although these are focused on supporting formal experimental procedures instead of unstructured, exploratory investigation. Again, Codex is distinguished by its ability to synthesize situations based on multiple individual conceptual structures and by its facility for expressing the relationships between abstract concepts and concrete data. Personalized knowledge maps (Novak et al., 2003) have been proposed as one mechanism for linking structure observed in data to structure inferred by the human analyst. Such maps provide an alternative formulation for depicting the association between data and knowledge, although do not offer historical tracking or explicit knowledge sharing among users.

ClaiMaker (Li et al., 2002) sits at a different point in the knowledge construction process; it is designed to express and contest arguments that exist in published material. Our intent with Codex is to capture knowledge in situ, that is, during the course of its construction. To this end, Codex models the domain knowledge that emerges from the relationships between people, data, tools, and concepts. Moreover, in capturing this knowledge we provide little organizing structure or predefined taxonomy, instead allowing this structure to emerge from cooperative analyses and researchers actively searching shared workspaces for concepts to modify and re-use. Like Codex, ClaiMaker allows different users to take different perspectives on the relationships between resources. We believe that the additional capability to search across users' workspaces to build inferential situations (Fig. 7) and to rapidly visualize alternative perspectives on a domain represent advancements necessary to increase the adoption of new knowledge sharing tools in the sciences. The preservation of use histories that allow any user to recover the reasoning trail that produced a finding is a further contribution of the Codex approach.

7. Conclusions

We argue for the use of knowledge representation techniques capable of reflecting the situated nature of human cognition. By representing knowledge in the contexts of its creation and use, it is possible for these situated representations to integrate both ontological (top-down) and discursive (bottom-up) approaches to knowledge elicitation and structure. By capturing knowledge as it is constructed and evolves, this work draws from the benefits of computer-supported cooperative work. By allowing groups to treat certain concept structures as representative of their community (and to share them as such), this work acknowledges the utility of top-down ontologies and the structure they can lend to a problem. By allowing users to collaboratively create and apply knowledge representations, the Codex system helps bridge the ideals of both the CSCW and ontological domains.

The benefits yielded by situated knowledge models and their implementation in collaborative systems should be felt across the range of inquiry-based activities. At the outset, pedagogy and basic scientific research are obvious application contexts—students and researchers are actively engaged in knowledge exposition and communication. In these applications, a tool like Codex formalizes, extends, and speeds an extant knowledge transfer process. Still, there are possibilities for a Codex-like approach in other application domains. Indeed, in any field where information is reused, built upon, or acted upon, the representation of the situated reasoning that went into a resource can help guide appropriate use. Planning and policymaking tasks, for instance, can require searching for consilience in group perspectives over time. In strategy-setting endeavors, understanding the coherence of individual perspectives and detecting common threads can lead to policy candidates likely to have widespread approval. (Although Codex, like policy, need not be democratic; particularly distinctive perspectives on a problem can be detected, appreciated, and adopted or rejected by others.) A knowledge model that preserves audit trails of resource manipulation and concept growth can increase the transparency of a research enterprise. Furthermore, the trustworthiness of any information resource can derive from an examination of how it came to be and how it has been used—thus, trust need not be a consistent measure for all circumstances, but can vary according to the needs of an inquirer or the conditions of a situation.

Tools that attempt to capture and integrate individual perspectives on knowledge require scientists to confront the issue of knowledge-sharing and to consider its benefits—not least a greater ability to integrate observations and hypotheses across space and over time. But without institutional support (from academic departments, funding agencies, research laboratories, journal publishers, and so on), a cultural shift may be difficult. There are early signs, though, that the need to achieve community synthesis is changing the way science is performed; the US National Science Foundation, for instance, is beginning to encourage greater knowledge sharing through community ownership of resources produced by its funding. Moreover, the techniques suggested in this paper will have benefit even if adopted at the individual, not collaborative, level. Indeed, it is possible that the benefits individual researchers might realize from using such tools will encourage the cultural shifts necessary to make their collaborative use possible.

Acknowledgments

The material presented in this paper is based upon work supported by National Science Foundation grants EAR/ITR-0225673, BCS-9978052, and BCS/ITR-0219025.

References

Ansari, A., Essegaier, S., Kohli, R., 2000. Internet recommendation systems. Journal of Marketing Research 37, 363–375.

Barsalou, L., 2002. Being there conceptually: simulating categories in preparation for situated action. In: Stein, N., Bauer, P., Rabinowitz, M. (Eds.), Representation, Memory, and Development: Essays in Honor of Jean Mandler. Lawrence Erlbaum, Mahwah, NJ, pp. 1–16.

Bozsak, E., Ehrig, M., Handschuh, S., Hotho, A., Maedche, A., Motik, B., Oberle, D., Schmitz, C., Staab, S., Stojanovic, L., Stojanovic, N., Studer, R., Stumme, G., Sure, Y., Tane, J., Volz, R., Zacharias, V., 2002. KAON—towards a large scale Semantic Web. E-Commerce and Web Technologies, Proceedings 2455, 304–313.

Brennan, S., Mueller, K., Zelinsky, G., Ramakrishnan, I., Warren, D., Kaufman, A., 2006. Toward a multi-analyst, collaborative framework for visual analytics. In: IEEE Symposium on Visual Analytics Science and Technology 2006. Baltimore, MD. IEEE, New York.

Buckingham Shum, S., 2006. Sensemaking on the Pragmatic Web: A hypermedia discourse perspective. In: First International Conference on the Pragmatic Web.

Buckingham Shum, S., Motta, E., Domingue, J., 2000. ScholOnto: an ontology-based digital library server for research documents and discourse. International Journal on Digital Libraries 3 (3), 237–248.

Buzaglo, M., 2002. The Logic of Concept Expansion. Cambridge University Press, New York.

Cao, C.G., Feng, Q.Z., Gao, Y., Gu, F., Si, J.X., Sui, Y.F., Tian, W., Wang, H.T., Wang, L.L., Zeng, Q.T., Zhang, C.X., Zheng, Y.F., Zhou, X.B., 2002. Progress in the development of national knowledge infrastructure. Journal of Computer Science and Technology 17 (5), 523–534.

Chin, G., Lansing, C., 2004. Capturing and supporting contexts for scientific data sharing via the Biological Sciences Collaboratory. In: 2004 ACM Conference on Computer Supported Cooperative Work, Chicago, IL. ACM, New York.

Chung, P.W.H., Cheung, L., Stader, J., Jarvis, P., Moore, J., Macintosh, A., 2003. Knowledge-based process management—an approach to handling adaptive workflow. Knowledge-Based Systems 16 (3), 149–160.

de Moor, A., Keeler, M., Richmond, G., 2002. Towards a pragmatic web. In: Priss, U., Corbett, D., Angelova, G., International Conference on Computational Science, Proceedings, vol. 2393. Springer, Berlin, pp. 235–249.

Dumais, S., Cutrell, E., Cadiz, J., Jancke, G., Sarin, R., Robbins, D., 2003. Stuff I've Seen: a system for personal information retrieval and reuse. In: SIGIR 2003, Toronto.

Dustdar, S., 2004. Caramba—a process-aware collaboration system supporting ad hoc and collaborative processes in virtual teams. Distributed and Parallel Databases 15 (1), 45–66.

Eckert, A., 1998. The "Network Elaboration Technique": A computer-assisted instrument for knowledge assessment. Diagnostica 44 (4), 220–224.

Engelbart, D., 1962. Augmenting Human Intellect: A Conceptual Framework. Stanford Research Institute, Menlo Park, CA.

Fensel, D., 2001. Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce. Springer, New York.

Fodor, J., 1998. Concepts: Where Cognitive Science went Wrong. Oxford University Press, New York.

Frodeman, R., 2003. Geo-Logic: Breaking Ground Between Philosophy and the Earth Sciences. SUNY Press, Albany.

Gray, P.H., Meister, D.B., 2004. Knowledge sourcing effectiveness. Management Science 50 (6), 821–834.

Guarino, N., 1997. Understanding, building, and using ontologies. International Journal of Human–Computer Studies 46, 293–310.

Guha, R.V., Lenat, D.B., 1991. Cyc—a midterm report. Applied Artificial Intelligence 5 (1), 45–86.

Heemskerk, M., Wilson, K., Pavao-Zuckerman, M., 2003. Conceptual models as tools for communication across disciplines. Conservation Ecology 7 (3).

Henzinger, M., Lawrence, S., 2004. Extracting knowledge from the World Wide Web. Proceedings of the National Academy of Sciences 101 (Suppl. 1), 5186–5191.

Kazic, T., 2000. Semiotes: a semantics for sharing. Bioinformatics 16 (12), 1129–1144.

Klein, M., Fensel, D., Kiryakov, A., Ogyanov, D., 2002. Ontology versioning and change detection on the Web. In: Gomez-Perez, A., Richard Benjamins, V. (Eds.), Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web, vol. 2473. Springer, Berlin, pp. 197–212.

Langley, P., Shrager, J., Saito, K., 2002. Computational discovery of communicable scientific knowledge. In: Magnani, L., Nersessian, N., Pizzi, C. (Eds.), Logical and Computational Aspects of Model-Based Reasoning. Kluwer, Amsterdam.

Lemke, J., 1997. Cognition, context, and learning: A social semiotic perspective. In: Kirshner, D., Whitson, J. (Eds.), Situated Cognition: Social, Semiotic, and Psychological Perspectives. Erlbaum, Mahwah, NJ, pp. 37–55.

Li, G., Uren, V., Motta, E., Buckingham Shum, S., Domingue, J., 2002. ClaiMaker: Weaving a semantic web of research papers. In: Horrocks, I., Hendler, J. (Eds.), The Semantic Web—ISWC 2002: First International Semantic Web Conference, vol. 2342. Springer, Berlin, pp. 436–441.

Lysakowski, R., Doyle, L., 1998. Electronic lab notebooks: paving the way of the future of R&D. Records Management Quarterly, 23–28.

MacEachren, A., Pike, W., Yu, C., Brewer, I., Gahegan, M., Weaver, S., Yarnal, B., 2006. Building a geocollaboratory: supporting Human –Environment Regional Observatory (HERO) collaborative science activities. Computers, Environment, and Urban Systems 30, 201–225.

Magnani, L., 2001. Abduction, Reason, and Science: Processes of Discovery and Explanation. Kluwer, New York.

Marcos, E., Marcos, A., 2001. A philosophical approach to the concept of data model: Is a data model, in fact, a model? Information Systems Frontiers 3 (2), 267–274.

Martens, B., Turk, Z., Bjork, B.C., Cooper, G., 2003. Re-engineering the scientific knowledge management process: the SciX project. Automation in Construction 12 (6), 677–687.

Myers, J., Mendoza, E., Hoopes, B., 2001. A collaborative electronic notebook. In: Proceedings of the IASTED International Conference on Internet and Multimedia Systems and Applications, Honolulu.

Nake, F., Grabowski, S., 2001. Human–computer interaction viewes as pseudo-communication. Knowledge-Based Systems 14, 441–447.

Neuwirth, C.M., Morris, J.H., Regli, S.H., Chandhok, R., Wenger, G.C., 1998. Envisioning communication: task-tailorable representations of communication in asynchronous work. In: Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work, Seattle. ACM Press, New York.

Novak, J., Wurst, M., Fleischmann, M., Strauss, W., 2003. Discovering, visualizing and sharing knowledge through personalized knowledge maps. In: Proceedings of the AAAI Spring Symposium on Agent-Mediated Knowledge Management, Lecture Notes in Artificial Intelligence 2926.

Noy, N.F., Hafner, C.D., 2000. Ontological foundations for experimental science knowledge bases. Applied Artificial Intelligence 14 (6), 565–618.

Pike, W., Yarnal, B., MacEachren, A.M., Gahegan, M., Yu, C., 2005. Retooling collaboration: a vision for environmental change research. Environment 47 (2), 8–21.

Rosch, E., 1975. Cognitive representations of semantic categories. Journal of Experimental Psychology 104 (3), 192–233.

Rosch, E., Mervis, C., 1975. Family resemblances: studies in the internal structure of categories. Cognitive Psychology 7, 573–605.

Sanchez, J., Langley, P., 2003. An interactive environment for scientific modeling and discovery. In: Proceedings of the International Conference on Knowledge Capture, Sanibel Island, FL, ACM Press.

Schultze, U., Boland, R.J., 2000. Knowledge management technology and the reproduction of knowledge work practices. Journal of Strategic Information Systems 9, 193–212.

Solomon, K., Medin, D., Lynch, E., 1999. Concepts do more than categories. Cognitive Science 3 (3), 99–104.

Sowa, J., 2000. Ontology, metadata, and semiotics. In: Ganter, B., Mineau, G. (Eds.), Conceptual Structures: Logical, Linguistic, and Computational Issues, vol. 1867. Springer, Berlin, pp. 55–81.

Stevens, R., Robinson, A., Goble, C., 2003. Mygrid: personalised bioinformatics on the information grid. Bioinformatics 19 (Suppl. 1), i302–i304.

Sugumaran, V., Storey, V.C., 2002. Ontologies for conceptual modeling: their creation, use, and management. Data & Knowledge Engineering 42 (3), 251–271.

Suthers, D., 1999. Representational support for collaborative inquiry. In: Proceedings of the 32nd Hawaii International Conference on System Sciences, Maui, HI, IEEE.

Turoff, M., Hiltz, S., 1996. Computer based Delphi processes. In: Adler, M., Ziglio, E.E. (Eds.), Gazing into the Oracle: The Delphi Method and its Application to Social Policy and Public Health. Kingsley, London.

Tversky, B., 2005. Form and function. In: Carlson, L., van der Zee, E. (Eds.), Functional Features in Language and Space. Oxford University Press, New York.

van Bruggen, J., Boshuizen, H., Kirschner, P., 2003. A cognitive framework for cooperative problem soving with argument visualization. In: Kirschner, P., Buckingham Shum, S., Carr, C. (Eds.), Visualizing Argumentation. Springer, London, pp. 25–47.

Wittgenstein, L., 1953. Philosophical Investigations. Macmillan, New York.