

## Re-engineering the GIS&T Body of Knowledge

Sean C. Ahearn<sup>a\*</sup>, Ilknur Icke<sup>a,b</sup>, Rajashree Datta<sup>a</sup>, Michael N. DeMers<sup>c</sup>,  
Brandon Plewe<sup>d</sup> and André Skupin<sup>e</sup>

<sup>a</sup>CARSI, Department of Geography, Hunter College, New York, NY, USA; <sup>b</sup>Department of Computer Science, University of Vermont, Burlington, VT, USA; <sup>c</sup>Department of Geography, New Mexico State University, Las Cruces, New Mexico, USA; <sup>d</sup>Department of Geography, Brigham Young University, Provo, Utah, USA; <sup>e</sup>Department of Geography, San Diego State University, San Diego, CA, USA

(Received 9 September 2012; final version received 1 May 2013)

A computational framework is presented for re-engineering the Geographic Information Science and Technology Body of Knowledge (GIS&T BoK). At its core is an ontology that is meant to simplify and extend the original BoK hierarchical structure to better capture relationships existing among concepts. Our approach builds on several key ideas. First is the notion of a knowledge corpus, an aggregate of both the internal cognitive forms of knowledge held by domain actors and the content of external artifacts that are produced and consumed by domain activities. Second is the notion of a reference system within which such artifacts are located and relationships among artifacts can be expressed. Third is the idea that by structuring the GIS&T concepts through the use of semantic web standards for formal ontologies and envisaging it as a reference system for GIS&T artifacts, activities, and actors, a fundamentally different approach to the redesign, content generation, and maintenance of the GIS&T BoK is enabled. This new approach affords replacing the top-down strategies used to generate the original GIS&T BoK, with a bottom-up strategy that combines analytical and participatory components. On the analytical side, computational and visual techniques are used to provide alternative means for accessing BoK content, examining the semantic consistency of current BoK structures, transforming the existing hierarchy into a semantic network, identifying overlaps and gaps in the current BoK, and performing projection of knowledge artifacts onto the BoK to inform its maintenance and update. Participatory approaches to bottom-up restructuring and maintenance of the BoK will support authoring, editing, and validation of concepts using a wiki-like community editing service. The system we describe is deployed as a web service that can be accessed by a range of applications for visualization, analysis, exploration, and contextualization of concepts and their related classes in the new GIS&T Body of Knowledge. The goal is for the new GIS&T BoK2 to evolve into the centerpiece of a cyberinfrastructure ecosystem for the GIS&T domain.

**Keywords:** ontology; cyberinfrastructure; semantic web; reference system; visualization

### 1. Introduction

#### 1.1. Background

Cyberinfrastructure (CI) is transforming what scientists can do by changing how they can do it. There is a major paradigm shift underway in how we think about computational

---

\*Corresponding author. Email: [sahearn@hunter.cuny.edu](mailto:sahearn@hunter.cuny.edu)

problems and how we collaborate with other individuals or groups (Wang 2010). The internet has enabled distributed collaboration on an unprecedented scale (Haklay 2008) and fundamentally changed the way we do what we do. Critical components that act as enablers of this web-based revolution are: high performance computing (Wang *et al.* 2002, Huang *et al.* 2006), distributed data and services (Chervenak 2000), distributed collaboration (Haklay 2008), and the semantic web (Maedche and Staab 2001, Egenhofer 2002, Kuhn 2003). The other major shift in CI concerns our thinking about and interaction with knowledge. The semantic web has enabled us to create linkages between different domains in ways that were previously impossible (Doan *et al.* 2003) and proposals have been made for ontologies specific to geographic information (Couclelis 2010).

Underlying this revolution is the dynamic nature of the knowledge base on which CI is founded and how it interacts with the domain of interest. CyberGIS is, in effect, the conflation of two knowledge domains, CI and geographic information science, that are both relatively young, dynamic, and quickly evolving. Forging a common language in such a young field is essential and the newly engineered Geographic Information Science and Technology Body of Knowledge (GIS&T BoK2) that we propose in this paper can play that role (see DiBiase *et al.* 2006 for the first version). Its goal is to create a realm in which the core of knowledge within the GIS&T domain can evolve more organically than it has in the past, to both include and *make evident* different means of coalescing around a concept. This will enable the maintenance and expansion of the knowledge base of GIS&T in a dynamic, interactive, and collaborative fashion and be a foundational component of CyberGIS by providing a platform for research, collaboration, teaching, and workforce development. In the realm of GIS&T, we have the benefit of an actual document to use as a starting point for this common core, the Geography Information Science and Technology Body of Knowledge (DiBiase *et al.* 2006).

### **1.2. History of the geographic information science and technology Body of Knowledge**

The Model Curricula project was initiated by the University Consortium for Geographic Information Science (UCGIS) Model Curricula Task Force in the late 1990s as an effort to address educational challenges in the field (DiBiase *et al.* 2009). The report produced by the task force called for the development of a *Body of Knowledge*, envisaged as a comprehensive inventory of the GIS&T knowledge domain. The first edition of the GIS&T BoK (DiBiase *et al.* 2009) was a collaborative effort within the UCGIS and was published by the Association of American Geographers in 2006. The BoK was developed by a team of seven editors in consultation with a 54-member Advisory Board, with content coming from over 70 additional scholars (DiBiase *et al.* 2006). The GIS&T BoK includes more than 350 topics organized in 79 units and 10 knowledge areas. Each topic is accompanied by terms of one or more educational objectives (UCGIS 2006).

The first edition of the GIS&T BoK caused a great deal of excitement in the domain community. Although a number of papers discussed the importance of using BoK in educational settings (Unwin *et al.* 2011), frameworks for actual implementation have been limited. As Foote *et al.* (2011, p. 8) point out

few departments have the staff and resources to address the full scope of the BoK. They must make choices about the core concepts and optional topics they will cover in their curricula. Although the BoK suggests developing ‘multiple pathways to diverse outcomes,’ none were developed for the first edition.

Prager and Plewe (2008) introduced a rubric for evaluating competency in BoK concepts, while Painho *et al.* (2011) presented a web-based ontology model for GIS&T BoK, along with a visual data exploration tool as a means for curriculum design. In Prager (2011), an approach to integrating the BoK into the design of GIS educational activities at the K-12, undergraduate, and graduate level is described. Among the concerns raised within the domain community are the relative weight and breadth of coverage of particular concepts, as well as the placement of specific concepts within the hierarchy (Toppen and Reinhardt 2009, Reinhardt 2012). Nevertheless, the GIS&T BoK was a landmark accomplishment in the field and the basis for a way forward (Rip 2008).

### 1.3. Re-engineering the GIS&T Body of Knowledge

Our vision of the new GIS&T BoK is founded on a computational framework that changes the manner in which content is represented, structured, accessed, and maintained. It enables several complementary approaches to maintaining and updating the BoK, including computational and collaborative techniques and support for a series of applications for visualization, exploration, and contextualization of domain concepts.

At the heart of the new representation and structuring of the GIS&T BoK2 is a *concept-based ontology* that is represented using semantic web technologies in the Jena framework (Apache Jena Project 2012). Where this system departs from other ontology-based bodies of knowledge (Riechert 2010) is in the conceptualization of the BoK as the basis of a reference system within which artifacts of a knowledge domain can be located through either assertion or inference. The metaphoric transfer of geographic approaches for creating and using locational reference systems is a key consideration in this regard, carrying with it notions of measurement (i.e., taking stock of the size, shape, structure, and evolution of the knowledge domain), standardization, and projection. One major concern addressed by a reference system, when applied to knowledge domains, is its provision of a framework for projecting the attributes and activities of various actors into a shared space. In effect, it enables applications that operate on the spaces that we intellectually and productively ‘inhabit’ (Figure 1).

Structuring core domain concepts via semantic web standards for formal ontologies and envisaging the GIS&T BoK as a reference system via which to project and analyze

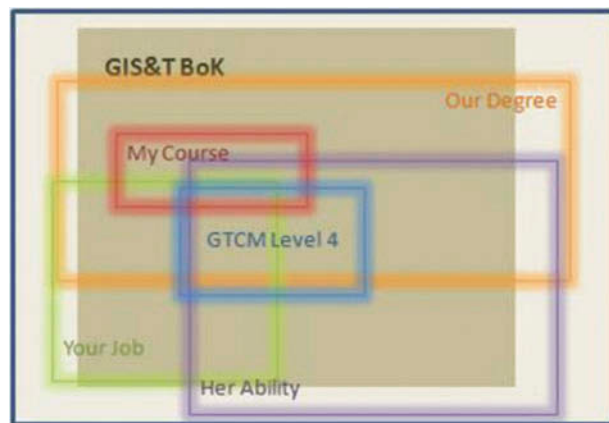


Figure 1. BoK as a reference system.

knowledge artifacts has significant implications for the transformation of existing BoK structures toward a new BoK2. Primary among these is the ability to re-evaluate the content structure of the current BoK, the ability to restructure its hierarchical model into a semantic network model, and the affordance of continuous update and maintenance of the BoK2.

The first edition of the GIS&T BoK (DiBiase *et al.* 2006) was created by domain experts in a *top-down* manner common to similar projects in other disciplines (Sahami *et al.* 2011). A major goal of this project is to shift BoK development toward *bottom-up* strategies by applying computational and visual analysis to the *artifacts* and *activities* of the knowledge domain (see Section 4). Such evaluation and restructuring of the BoK can be supported by a variety of data mining and spatialization techniques, including topic modeling (Blei *et al.* 2003) and self-organizing maps (Skupin 2009, Skupin *et al.* 2013). Another bottom-up strategy aims to actively leverage the collective knowledge and insight of domain *actors* (i.e., members of the GIS&T community) in a participatory, collaborative setting where BoK2 concept development, editing, and validation are supported by a visual wiki environment (see Section 4 below).

#### 1.4. Conceptual model

The conceptual model for a re-engineered GIS&T BoK2 consists of a three-layer system with the BoK Ontology (BoKOnto) at its core, a server layer for administration and management, and a service application layer (Figure 2).

BoKOnto (orange center in Figure 2) is an ontology that uses current RDF and OWL standards (McGuinness, D.L., and Van Harmelen, F., 2004). It describes a flexible, but standardized set of concepts, allowing for multiple situated views of the knowledge corpus by researchers, educators, professionals, novices, or experts. The server layer (green) handles the entire administration and management of BoKOnto. Implemented as a semantic

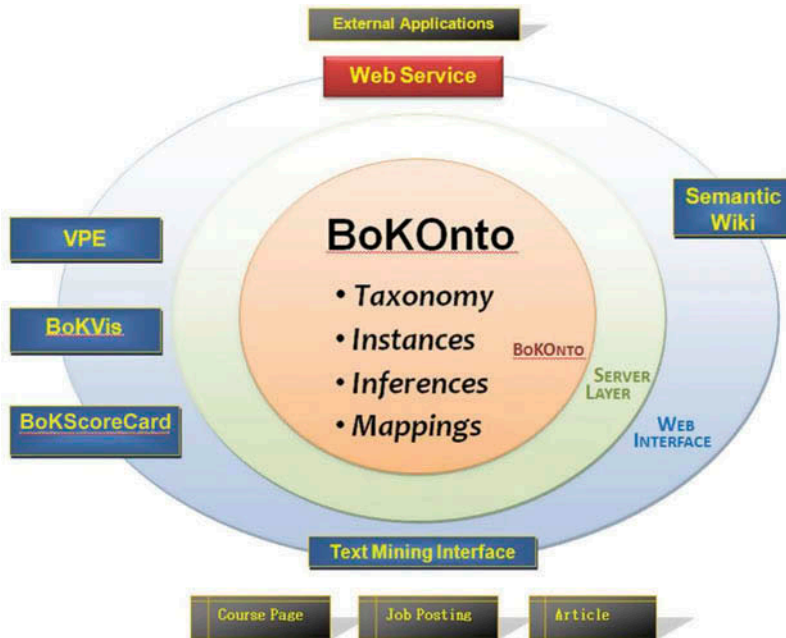


Figure 2. System model for the GIS&T BoK2.

web application, it includes an approval process for populating the GIS&T BoK2 ontology, customized analysis, database versioning, and general database management tasks. The ‘applications’ (blue layer) enable the users and maintainers of the system to interact with it in a range of ways including visualization, update and maintenance, and ontologic mapping. The system relies on a RESTful service to access the BoKOnto core.

## 2. The role of knowledge in a domain ontology

Our reformulation of the GIS&T BoK as a domain ontology is central to its re-engineering. Several key concepts help to give context to our approach:

- (1) *Knowledge domain*: This notion refers to the aggregate of all the actors, activities, and artifacts that are bound by certain recognizable social constructs, including exhibiting significant thematic and epistemic coherence. It encompasses the notion of disciplines, but recognizes the existence of persistent structures of knowledge construction and consumption beyond disciplinary boundaries. GIS&T forms such a knowledge domain.
- (2) *Domain actor*: An entity (e.g., person, organization, institution) actively engaged in producing and consuming knowledge artifacts. The participatory mode of bottom-up BoK creation depends on actors bringing their experiences, abilities, and perspectives – their internal knowledge – to bear. Domain actors are not only among the intended users of a formalized BoK, but can also become objects of analysis, such as when a person’s capabilities are being scored with respect to BoK concepts.
- (3) *Domain activity*: Actors within a domain are engaged in various activities that produce and/or consume knowledge artifacts, within the broad categories of research, education, and professional practice. Examples might include a research project, a university course, a geographic analysis task, or a job. Depending on an actor’s role within the domain, there may be significant variety in the types of activities as well as in the types of artifacts entailed and in the productive/consumptive relationship between actors and artifacts.
- (4) *Knowledge artifact*: Processes of knowledge production and consumption involve physical artifacts (i.e., externalized as compared to internal, cognitive artifacts, which most commonly are just referred to as ‘knowledge’) (Newman 2003), such as research articles, grant proposals, course syllabi, websites, textbooks, and software code.
- (5) *Knowledge corpus*: We shall refer to the aggregate of all internal knowledge (e.g. actors’ relevant skills) and external artifacts contained in the domain as its *knowledge corpus*. The task of constructing a BoK consists of conceptualizing, capturing, and transforming that corpus. Developing the framework and key parameters for accomplishing this within the context of the GIS&T domain is a major goal for the project described in this paper.

## 3. Representation and restructuring of the GIS&T BoK

### 3.1. GIS&T BoK2 ontology

The document design for the original GIS&T BoK provided significant inspiration for the design of the GIS&T BoK2 ontology. In the original, concepts were nested in a

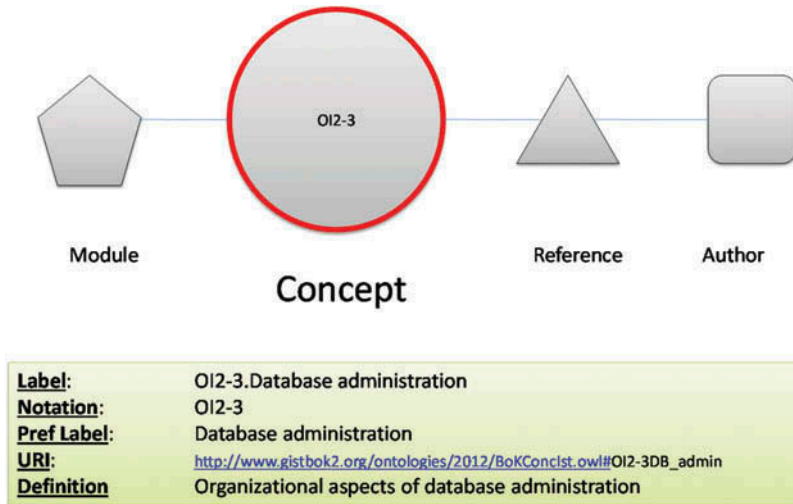


Figure 3. GIS&T BoK2 ontology, with BoK concept instance example.

hierarchy of ‘Knowledge Area’, ‘Unit’, ‘Topic’, and ‘Learning Objective’, overtly imitating the structure of the Association for Computing Machinery (ACM) Computing Curricula (ACM/IEEE-CS 2013). This was followed by a short list of references.

The GIST BoK2 ontology seeks to transform this hierarchical structure to better capture relationships among concepts through part-whole hierarchies and associations, employing semantic web standards for formal ontologies (Miles and Bechhofer 2008). The reasoning here is that while any particular ‘Knowledge Area’ (for example) may be a penultimate (or top-level) designation in our taxonomy today, it may be subsumed by another category in the future while still retaining its original subcategories. In truth, the history of the evolution of academic disciplines suggests that the actual contents of some original disciplines remain intact while their hierarchical relationships are in flux (e.g., Newtonian physics). Each of the core classes of the domain ontology (i.e., BoKOnto) are described below (Figure 3).

- **Concept:** parent class of BoKConcept and BoKSkill, it captures similar properties exhibited by both, such as hierarchical (part/whole) and similarity relationships. Unlike other elements of the ontology, BoKConcept leverages the Simple Knowledge Organization System (SKOS) approach to the representation of controlled vocabularies (Miles and Bechhofer 2008).
- **BoKConcept:** this is the fulcrum around which the entire BoK is constructed, and is a simplification of the ‘Knowledge Area’, ‘Unit’, and ‘Topic’ designations in the original document. However, where the original document designated a fixed three-level hierarchy, the ontology seeks to be more flexible, to better represent the dynamic, complex structure of the knowledge corpus, by combining all of the levels into a single concept class. A *BoKConcept* is thus a distinct unit of knowledge of any breadth or depth: *analysis*, *proximity-based analysis*, *buffer*, and *polygon buffer* are all concepts; linked together by a ‘part-whole’ relationship.
- **BoKSkill:** this class encapsulates and extends the learning objectives listed under each topic in the original document. *BoKSkill* is a task that applies and demonstrates

the knowledge of a concept; for example, ‘create buffers of a particular distance from a set of point, line, or area features.’ A course may teach this skill, a person may be capable of performing this skill, and a project may require this skill in order to be successfully completed.

- **BoKElement**: Parent of BoKRef, BoKAgent and BoKModule, this class captures these three subclasses, two of which are clearly relatable to concepts: namely references and agents. The third subclass is the BoKModule which is less central to the support of a concept and can include any set of knowledge artifacts, such as courses, an individual’s skill set, and curricula.
- **BoKRef**: This class is reserved for knowledge artifacts that are central to understanding a concept. In other words, it can be thought of as providing *primary* sources for the concept. Figure 4 shows the relationship between a reference and a concept. It is important here to note that should the reference have pointed to a concept *narrower* than our concept of interest, then the ontology would have been able to infer that it was a reference for our concept of interest by extension. This is accomplished using a basic property chain to assign reference *by extension* through part-whole relationships.
- **BoKAgent**: The class is the agent associated with the BoKRef, such as its author or institution. The relationship between References and Agents most commonly consists of a journal article and an associated set of authors.
- **BoKModule**: The most basic collection of elements is the module, which could be any set of knowledge artifacts that are defined by some theme, entity, or unifying process. For instance, a course which consists of parts including a basic introduction, a query module on SQL, and a technical introduction could represent a module (Figure 5). It should be noted that any one of these modules could potentially have

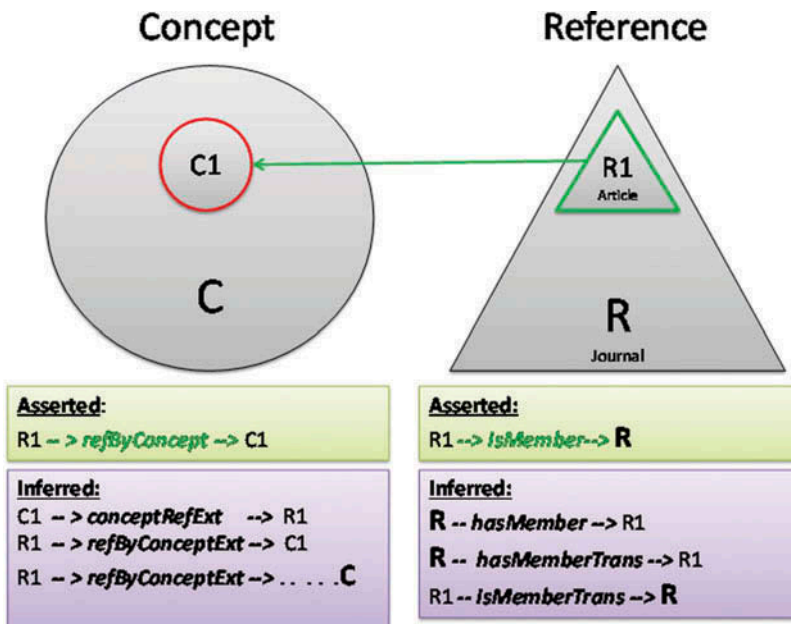


Figure 4. Relationship between concept and reference.

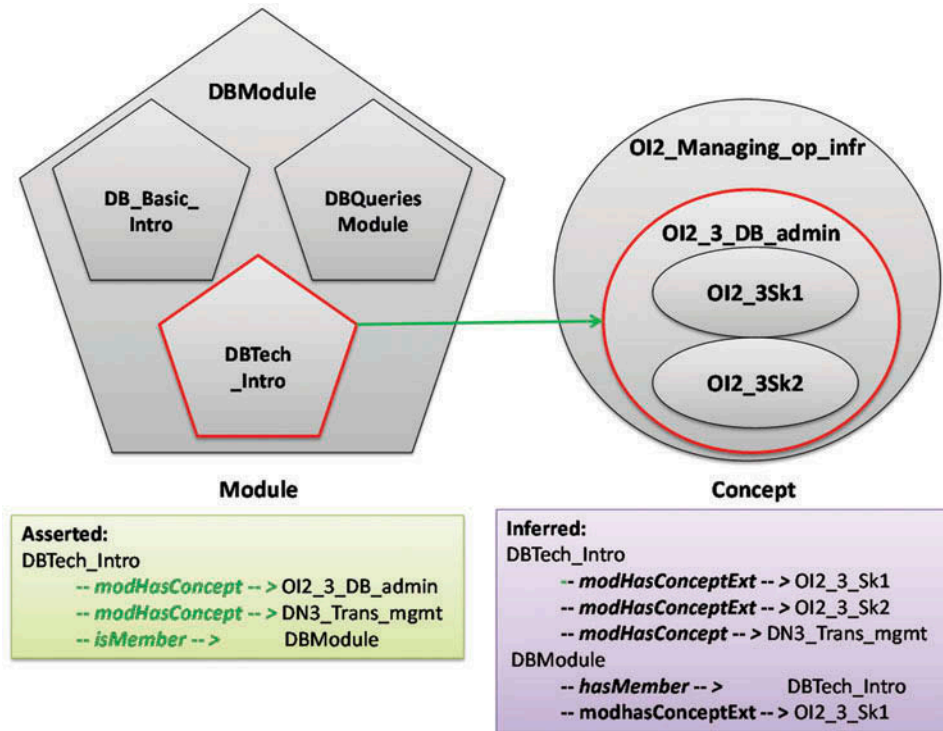


Figure 5. Module relationships and concept inheritance with asserted and inferred relationships.

their own nested series of modules, each with an associated set of concepts. Much like the hierarchical inheritance (and use of a property chain) between references and concepts, the connection between modules and concepts leverages membership to extend a module's attachment to a concept.

The core ontology infrastructure of the GIS&T BoK is implemented using several semantic web technologies, such as Jena software and ontology standards such as SKOS (Miles and Bechhofer 2008). We chose to use SKOS precisely because it is a W3C standard that was specifically designed for classification schemes and subject-heading lists. While currently only a subset of SKOS's object and data properties are leveraged in the BoK, using this common standard could allow us to both *import* and *export* classification schemes for concepts which might be already available in other subject areas. Given the transdisciplinary nature of GIS&T, it is plausible that we might seek to interface with other domains using accepted standards such as Learning Object Metadata (LOM) and Sharable Content Object Reference Model (SCORM) (Ritzhaupt, 2010, Klemke, 2010).

### 3.2. Restructuring the BoK from concept hierarchy to concept network

One major concern regarding the first edition of the GIS&T BoK has been its hierarchical organization. The highest level in this organization is the knowledge area, of which there are ten. Each knowledge area contains at least one unit and each unit contains one or more topics. A hierarchical tree structure is a very constrained form of knowledge organization



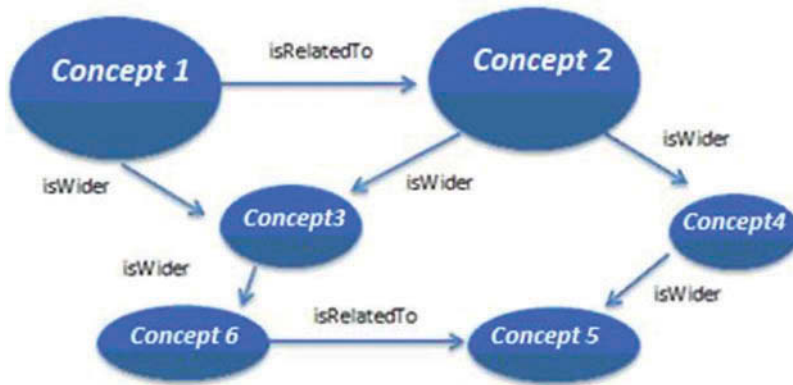


Figure 6. An example of semantic network of concepts.

since it only supports the parent–child relationship. In reality, there are many different types of relationships among concepts such as similarity, possession of multiple parents, or existence of a natural sequence (i.e., prerequisite concepts). Our representation for the GIS&T BoK2 is instead based on a network topology that eliminates certain shortcomings of a simple hierarchical structure (Figure 6).

Our approach for this migration from *hierarchy* to *network* is to use analytical and participatory techniques to find associations across the network for which significant relationships exist (see Section 4 below).

### 3.3. Design framework

Conceptually, the new semantic network for GIS&T BoK2 could be likened to an entity–relationship model used in relational database design. However, relational databases are limited to one kind of relationship that is identified by the foreign key. The semantic web offers more complex relationships, such as inheritance, part-of, associated-with, and many others, including logical relationships and constraints (Hebeler *et al.* 2009). Figure 7 shows the comparison between relational database-based and semantic web-based knowledge modeling.

### 3.4. Application architecture

Figure 8 shows the architecture of the public domain web services environment (i.e., knowledge exploration tool) providing services to a wide range of applications. The live ontology is kept on a text file-based triple store, *TDB* (Apache Jena Project 2012). The triple store is then interfaced with a *SPARQL* endpoint, which allows the ontology to be queried using the *SPARQL* query language. The *SPARQL* endpoint is not directly open to the public, since it would require users to learn the *SPARQL* query language and issue queries directly, but could be at a future time. Instead, we designed a set of web services that run queries internally and return data in an XML format. These web services can be accessed by users via web interface or they can be accessed programmatically.

These services can be configured for a range of applications and we have successfully done so for the BoKVis and BoKScoreCard applications discussed in Section 5.

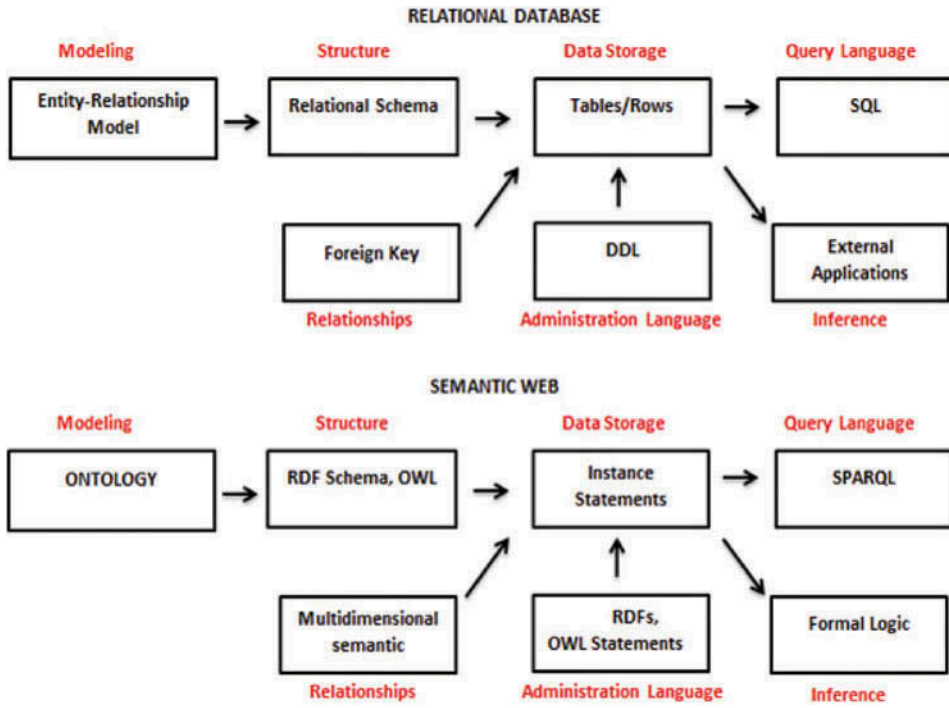


Figure 7. Relational database versus semantic web-based knowledge modeling.

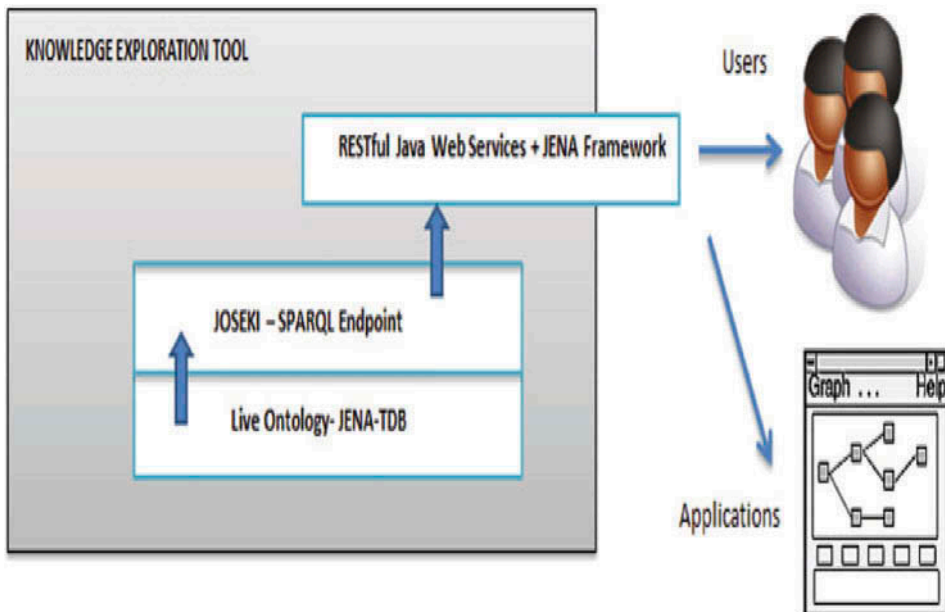


Figure 8. Architecture of the general purpose GIS&T BoK application (public).

#### 4. Generation, maintenance, and update of the BoK2

There are a number of challenges to creating a BoK that will maintain its currency through time. GIS&T is a large, diverse, international enterprise that is still in its adolescence and, thus, changing rapidly. To represent the knowledge corpus underlying this domain, the BoK needs to be maintained dynamically in a bottom-up fashion. To that end, we are employing two approaches; one collaborative, using experts from many subdisciplines, including academe and professional practice, and from many countries; the other relying on computation and visualization.

Collaborative content development, often called *crowdsourcing*, has proven to be a useful and efficient way to collect and share the knowledge corpus held by a large community, as evidenced by popular services such as Wikipedia and OpenStreetMap. Therefore, we have developed a wiki-based tool, called BoKWiki, for community editing of the next GIS&T BoK. The first prototype was built using existing text-based Wiki tools, with extensions to support the ontology-based semantic web. The second prototype, shown in Figure 9, leverages our research in visualizing the BoK to create a ‘visual wiki.’ In this site, users can navigate a network graph of related BoK concepts, view detailed information about each concept and relationships between concepts, propose their own additions and changes to the content of the first edition. Alternative interfaces are also being developed that balance the text-based and visualization-based approaches to navigation. The idea of combining textual knowledge management with information visualization tools has recently been proposed in information management literature (Hirsch *et al.* 2009) and our research suggests it is a powerful approach in a semantic web environment.

A major issue with previous crowdsourcing efforts is the lack of editorial control; however, the diversity and breadth of GIS&T makes it difficult for a small set of editors (as in the first edition) to effectively arbitrate debates about concepts. The BoKWiki

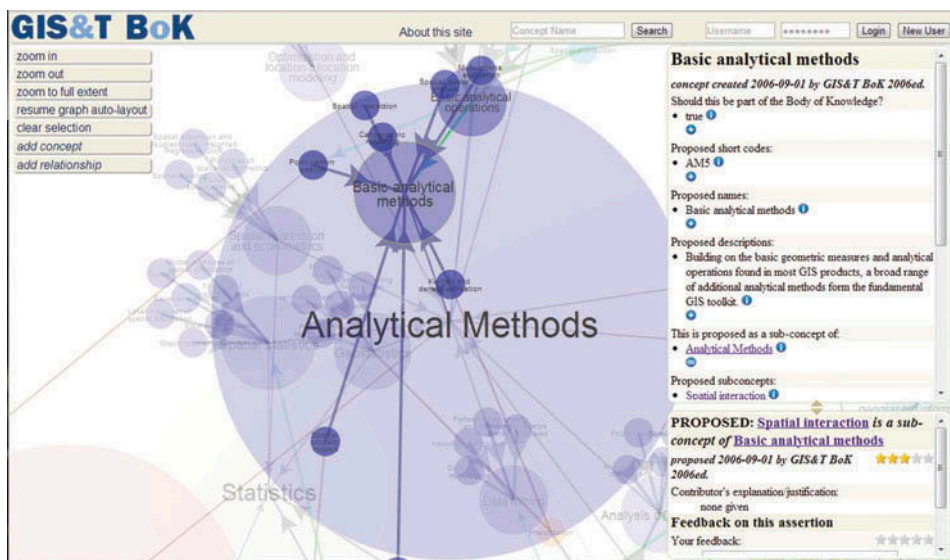


Figure 9. The BoK visual Wiki integrates the community editing features of a wiki with a visualization of the GIS&T knowledge space.

takes a bottom-up approach to this as well by crowdsourcing the editorial process. All contributors are able to rate and give feedback on the previous edition and proposed changes thereto. However, we are currently moving this beyond a simple rating system. Users' contributions and feedback will be weighted according to their reputation in each concept, which is based partially on their past record and partially on the quality of their previous contributions to the BoK. In this system, anyone can propose changes to the BoK, but the community collaboratively chooses which proposals will enter its next edition.

A bottom-up computational approach for populating and maintaining a BoK leverages data-driven techniques, which aim to harvest artifacts contained in the GIS&T knowledge corpus. Computational analysis of current BoK1 content is a first step toward the bottom-up generation of a reference system for the GIS&T domain, for example, by identifying latent semantic dimensions of the domain through topic modeling (Blei *et al.* 2003). Together with a similarity-based inference engine, the BoK1 text content can then be expressed in terms of this reference system and, via dimensionality reduction, projected into a two-dimensional display space. That is how the BoK1 text content can be transformed into a GIS&T base map (Figure 10), which allows the visual examination of semantic coherence of BoK1 hierarchy structures. More importantly, the combination of semantic reference system and dimensionality reduction allows *overlay* of other knowledge artifacts, even if these had not been available or used at the time of model creation (Skupin 2009). This is illustrated in Figure 10 with an overlay of 11 publications seen as the core background of the NSF-funded CyberGIS project (CyberGIS Project 2012). Titles, abstracts, and keywords of these publications are the basis of their overlay onto the BoK1-derived GIS&T base map. Notice that all 11 publications end up being mapped within the boundaries of the 'Analytical Methods' knowledge area! Overlays like these can be the starting point and context for crucial discussions regarding the current state and future evolution of the GIS&T BoK, specifically on how emerging trends should be reflected. For example, while

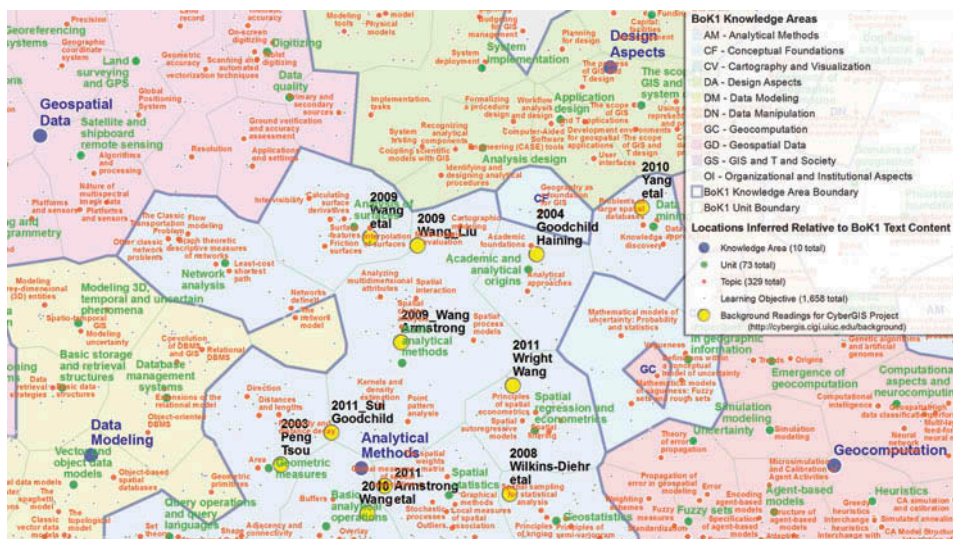


Figure 10. Domain base map derived from a topic model of the GIS&T BoK. Overlaid are 11 foundational papers on CyberGIS.

many of the overlaid publications end up in meaningful surroundings, certain recent trends are clearly having a hard time finding a home in BoK1. Such is the case with Peng and Tsou's 2003 monograph on Internet GIS, distributed geographic information services, and wireless networks, topics that are virtually absent from BoK1.

## 5. BoK applications

The RESTful web service can be accessed by a range of applications. These applications will enable the construction of varying narratives of the knowledge domain, enable its filtering, and provide the ability to compare and contextualize perspectives. The REST services can be configured to access a flexible set of concepts, (see Section 3.4), which enable the creation of multiple 'versions' of the BoK as seen from differing perspectives: a researcher's, an educator's, a professional's, a novice's, or an expert's. Several key applications are highlighted below that will be covered in more detail in subsequent publications.

### 5.1. BoKScoreCard application

The BoKScoreCard provides a means for projecting or 'mapping' individual elements of the knowledge domain (actors, activities, and artifacts) onto the GIS&T BoK.

In the parlance of our BoK ontology, artifacts, actors, and activities can be collected in BoK Modules, and the BokScoreCard notes the degree to which each reflects and relates to concepts in the BoK ontology. This can be either *asserted* (e.g., a researcher judging her/his own expertise) or *inferred* (e.g., a researcher's expertise determined through text analysis of her/his publication history).

Once a domain element is mapped onto the BoK via its BoKScoreCard, its location in the reference system can be compared to that of other artifacts. The results of such a comparison can even be visualized. In the initial prototype, a person who is mapping a particular artifact uses a web-based form to assign a level of expertise in each concept in the BoK 1st Edition. This level rubric (developed in Prager and Plewe 2008) represents the degree of knowledge and ability that

a person has, or a course teaches, or a job requires, from basic 'familiarity,' through a few degrees of competence, to the highest 'research and development' level. This prototype only represents the level of expertise the user claims or *asserts*, which has obvious limitations; tools for *validating* and *certifying* those assertions need to be developed in the future.

Once multiple artifacts have been entered, the data can be analyzed, to summarize trends (e.g., 'what do most introductory GIS courses in the US cover?') or make comparisons (e.g., 'how do my abilities match the requirements of this job?'). As an example, [Figure 11](#) shows a comparison of the GIS&T expertise of two of this paper's authors (i.e., domain actors) as mapped on a visual representation of the BoK1, using a bivariate choropleth technique. Here, bright green and bright red boxes are concepts for which one author (i.e., 'green author') or the other (i.e., 'red author') claims much more proficiency than his colleague, brown boxes are concepts at which both authors claim to be equally knowledgeable, and white boxes are concepts at which neither author considers himself very proficient.

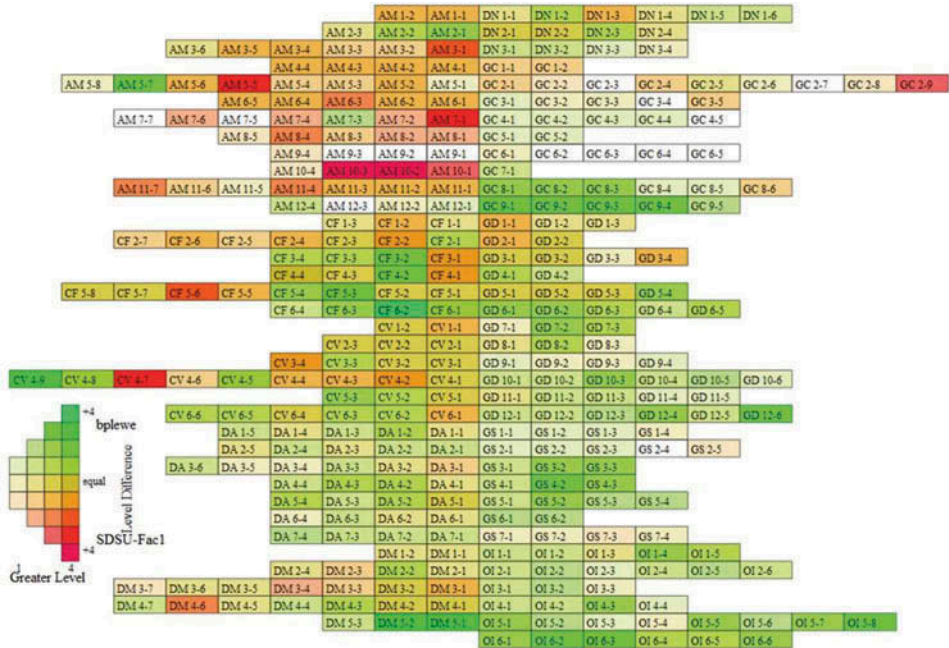


Figure 11. A comparison of the asserted proficiency of two people, using their BoKScoreCards mapped onto a structural visualization of the GIS&T BoK1.

**5.2. BoKVis application**

BoK1 had been developed in a completely top-down manner, by a team of experts that organized the topical components of the GIS&T into a hierarchical tree structure. The linear page sequence of the printed BoK1 document (DiBiase *et al.* 2006) naturally fails to fully convey that structure. The initial goal of BoKVis has been to better convey the existing hierarchical structure. To that end, the current BoKVis application implements three main hierarchy visualization methods: tree graph, indented list, and tree map (Figure 12a, b, c). While those methods completely conform to the top-down hierarchy, a similarity graph implementation is also being developed. Its geometric layout breaks out of the hierarchy by arranging BoK elements according to the similarity of their text content (Figure 12d). This bottom-up layout is then complemented by line work that reflects the original BoK hierarchy. Note how different layouts vary in their use of space, which has implications for the ability to convey text-heavy content and for overall usability. The latter is being addressed through an ongoing human subject study. Whichever layout method is being used, BoKVis is intended to provide a base map functionality onto which user-driven content can be overlaid (e.g., Figure 12).

**5.3. BoK VPE**

In many ways, the use of virtual persistent environments (VPEs) to explore and virtualize concepts is in line with the constructivist view of education. In VPEs, a student will be able to create three-dimensional realizations of concepts and explore them in ways that are not possible in any other medium. Among these realizations is the ability to interact on a





Figure 13. A chatbot (left) being interrogated by an avatar (middle). The kiosk on the right allows for redirection to holodecks of individual knowledge areas.

personal level with the use of avatars. Our initial entry into this environment was based on the now well-established virtual world known as Second Life<sup>®</sup>. Within Second Life, we created a virtual environment not unlike the holodeck technology suggested in the *Star Trek Next Generation* television series. When participants' avatars enter the environment in Second Life, they are greeted by a digital robot, known as a chatbot and are able to query a database of selected GIS&T-related questions which are organized off-game using artificial intelligence techniques surrounding the ALICE project and now called Pandorabots (Wallace 2005). A kiosk near the chatbot contained all 10 of the knowledge areas, and each could be touched by the avatar, the result of which would be that they would be instantly transported to another area of the simulation that contained the subject-specific chatbot as well as a context-sensitive holodeck containing links to outside websites, videos, and other multimedia sources (Figure 13).

While Second Life offered many advantages and proved to be a useful test case for a VPE, it had a number of flaws, most critically the fact that it couldn't receive 'real-time' updates from our web services. Unity 3-D is a promising technology we are currently exploring, whose environment enables the animation of virtual 3-D objects, the development of in-game and on-screen user interfaces and the creation of extensive interactive scenarios using standard transportable languages like Java Script, C#, and Boo (a dialect of Python) (Unity 2012). Unity itself is a promising technology for game development, but is normally designed for single-user environments. Its Jibe interface includes multiple open source objects designed to leverage the best for gaming, web service communication, and communications.



## 6. Conclusion and future development

The premise of this paper is that there is a critical need to improve the manner in which the construction, dissemination, and use of knowledge is organized within and across the traditionally disparate concerns of scientific research, academic education, technical training, and professional practice. Within this context, we have introduced the foundation of a new computational framework for the GIS&T BoK2 that structures core concepts and knowledge artifacts, while leveraging Semantic Web standards for formal ontologies. At the heart of our contribution lies a redesigned ontology for the GIS&T BoK2, one that extends the hierarchical structure of the original BoK to focus on *concepts* that are related through part-whole hierarchies and associations. The central ideas are encapsulated as follows: the notion of a *knowledge corpus* as the aggregate of all the *artifacts* that are produced and consumed by the *activities* of a *knowledge domain*, and the notion of a *reference system* within which artifacts are located and by which their relationships can be expressed. The system is deployed with Jena Semantic Web technologies that drive a RESTful web service that is accessed by a series of applications enabling visualization, exploration, maintenance, and contextualization of GIS&T concepts and entities. We believe this new computational framework is differentiated by its deployment of such familiar geographic notions as navigation, projection, and overlay analysis from other efforts to create domain BoKs. In terms of updating and maintaining the GIS&T BoK, our framework places emphasis on a bottom-up approach, with particular focus on enabling continuous community-wide input through a collaborative visual wiki environment. These will be supported by computational approaches to knowledge discovery and aggregation, including natural language processing, topic modeling, and artificial neural networks. As demonstrated in this paper, we also significantly advance the BoK vision through an expanded view of how it could be deployed and used. With the notion of a reference system at the BoK's core, the competencies of individuals, organizations, and particular institutional efforts (e.g., curricula) can now be *asserted* through the concept of a BoKScoreCard or *inferred* through computational means. Visualization is put forth as a key strategy for helping the domain community understand the structure and content of the BoK itself and, more importantly, help to convey relationships existing among knowledge artifacts through their overlay onto the BoK base map. The BoK vision is thus broadened toward evolving into an enabler of knowledge exploration, analysis, comparison, and synthesis. Update and maintenance is a critical component of this new system and we will begin piloting the BoK Visual Wiki in the near future. Over the past 2 years, we have presented our research at over a dozen conferences and meetings throughout the United States and Europe to a wide range of national and international organizations in geography, geographic information science, cartography, CyberGIS and remote sensing, and photogrammetry. Collaboration with these organizations is a key strategy for realization of the next version of the GIS&T BoK and the continued development of its new framework. In summary, we believe this new computational framework has the potential to evolve into the centerpiece of a CI ecosystem for the GIS&T domain by acting as the foundation of a space in which we explore, compare, and contrast knowledge artifacts, generate hypothesis, compare algorithms, and conduct research.

## Acknowledgments

The authors wish to thank the anonymous reviewers for their helpful comments and suggestions. We would also like to thank Professor Shaowen Wang for his leadership in bringing to fruition this special issue on CyberGIS. We are grateful to the NSF Geography and Spatial Sciences for supporting this research under Grant Award 0964563.

## References

- ACM/IEEE-CS Computer Science Curricula, 2013. Available from: <http://www.sigart.org/CS2013-EAAI2011panel-RequestForFeedback.pdf>
- Apache Jena, 2012. Available from: <http://jena.apache.org/>
- Blei, D.M., Ng, A.Y., and Jordan, M.I., 2003. Latent dirichlet allocation. *Journal of Machine Learning Research*, 993–1022.
- CyberGIS Project, 2012. Background [online]. Available from: <http://cybergis.cigi.uiuc.edu/background> [07 September 2012].
- DiBiase, D., et al., 2009. *Supporting curriculum development in geographic information science and technology: the GIS&T Body of Knowledge*. Chapter 62 in Madden, M., Bathesda, MD: Manual of Geographic Information Systems, ASPRS, 1261–1271.
- DiBiase, D., et al., eds. 2006. *Geographic information science & technology Body of Knowledge, University Consortium for geographic information science*. Washington, DC: AAG.
- Doan, A., et al., 2003. Learning to match ontologies on the Semantic Web. *The VLDB Journal*, 12, 303–319/Digital Object Identifier (DOI) 10.1007/s00778-003-0104-2
- Couclelis, H., 2010. Ontologies of geographic information. *International Journal of Geographical Information Science*, 24 (12), 1785–1809.
- Egenhofer, M., 2002. Toward the semantic geospatial web. *Paper presented at the 10th ACM international symposium on advances in geographic information systems*. McClean, VA, USA.
- Hebeler, J., et al., 2009. *Semantic web programming*. Indianapolis: Wiley Publishing, Inc.
- Hirsch, Christian, John Hosking, and John Grundy. 2009. Interactive visualization tools for exploring the semantic graph of large knowledge spaces. *Proc. of the IUI Int. Workshop on Visual Interfaces to the Social and the Semantic Web*, Sanibel Island, USA. 2009.
- Kuhn, W., 2003. Semantic reference systems. *International Journal for geographic information science*, 17 (5).
- Maedche, A., and Staab, S., 2001. Ontology learning for the Semantic Web. *IEEE Intelligent Systems*, 16 (2), 72–79.
- McGuinness, D.L., and Van Harmelen, F., 2004. *OWL web ontology language overview* [online]. Available from: <http://www.w3.org/TR/owl-features/> [Accessed 7 May 2012].
- Miles, A., and Bechhofer, S., 2008. SKOS simple knowledge organization reference [online]. Available from: <http://www.w3.org/TR/2008/WD-skos-reference-20080125/> [Accessed 7 May 2012].
- Newman, B., 2003. Agents, artifacts, and transformations: the foundations of knowledge flows. In: C.W. Holsapple ed. *Handbook on knowledge management 1: Knowledge matters*. Berlin: Springer-Verlag, 301–316.
- Painho, M. and Curvelo, P., 2011. Ontology-based alternative paths for GIS&T curricula. In: D.J. Unwin, David J., Kenneth E. Foote, Nicholas J. Tate, and D. DiBiase, eds. *Teaching geographic information science and technology in higher education*. Chichester: Wiley, 97–115. doi: 10.1002/9781119950592.ch7.
- Prager, S.D., 2011. Using the GIS&T Body of Knowledge for curriculum design: different design for different contexts. In: Unwin, David J., Kenneth E. Foote, Nicholas J. Tate and D. DiBiase, eds. *Teaching geographic information science and technology in higher education*. Chichester: Wiley, 63–80.
- Prager, S.D., and Plewe, B., 2009. Assessment and evaluation of GIScience curriculum using the geographic information science and technology Body of Knowledge. *Journal of Geography in Higher Education*, 33, S1–S24.
- Riechert, T., et al., 2010. Knowledge engineering for historians on the example of the catalogus professorum lipsiensis. *The Semantic Web – ISWC 2010. Lecture Notes in Computer Science*, 2010, 6497/2010, 225–240.
- Reinhardt, W., 2012. *Views on the body of knowledge*. Avignon, France: AGILE.
- Rip, F., 2008. GIS&T Body of Knowledge: basis for e-learning, certification and curriculum planning? *11th AGILE international conference on geographic information science 2008*.
- Sahami, M.M., et al., 2011. Setting the stage for computing curricula 2013: computer science – report from the ACM/IEEE-CS joint task force. *SIGCSE’11, Proceedings of the 42nd ACM technical symposium on computer science education*.
- Skupin, A., 2009. Discrete and continuous conceptualizations of science: implications for knowledge domain visualization. *Journal of Informetrics*, 3, 233–245.

- Skupin, A., Biberstine, J., and Börner, K., 2013. Visualizing the topical structure of the medical sciences: a self-organizing map approach. *PLoS ONE*, 8 (3), e58779.
- Toppen, F., and Reinhardt, W., 2009. A European approach towards the UCGIS geographic information science and technology Body of Knowledge, a discussion paper. *12th AGILE international conference on geographic information science 2009*.
- Unwin, D., et al., eds., 2011. *Teaching geographic information science and technology in higher education*. John Wiley and Sons.
- Wang, S. 2010. A CyberGIS framework for synthesis of cyberinfrastructure, GIS and spatial analysis. *Annals of the Association of American Geographers*, 100 (3), 2010.
- Wang, S, Armstrong, M. P., and Bennett, D. A., 2002. Conceptual basics of middleware design to support grid computing of geographic information. *Paper presented at the 2nd international conference on geographic information science*. Boulder, CO, USA.