Fonseca, F. (2000) Users, Ontologies and Information Sharing in Urban GIS. in: *ASPRS Annual Conference*, Washington, D.C.

Users, Ontologies and Information Sharing in Urban GIS

Frederico T. Fonseca

National Center for Geographic Information and Analysis, University of Maine
Department of Spatial Information Science and Engineering, University of Maine
Boardman Hall
University of Maine
Orono, ME 04469-5711, USA
fred@spatial.maine.edu

ABSTRACT

Many applications involve tasks that require the sharing of geographic information. Interoperable solutions provide end users with a means to retrieve and share data. Data and knowledge exchange among users of urban information systems presents many challenges. This paper discusses the role of user communities in specifying the concepts behind their everyday work with geographic information and how the resulting knowledge can be reused. The result of the user specifications is an ontology that is used to build *ontology-driven geographic information systems*.

1 INTRODUCTION

Starting an urban geographic information system (GIS) project presents many challenges. Describing the detail-rich urban environment is one of them. To face this challenge, the use of existing knowledge from previous GIS projects is a necessity. Beyond that, the use of existing data is also desirable. But the lack of formal methods to reuse knowledge and data makes this task difficult. We discuss in this paper the role of user communities in specifying the concepts behind their everyday work with geographic information and how this knowledge can be reused.

Within a defined group of users, such as cities in a state, a well-defined conceptualization of an urban environment can be shared among all the cities of the state even if they are not neighbors. The concept of *geospatial information communities* (GIC) as a group of users that share a digital geographic information language and spatial feature definitions was introduced by McKee and Buehler (1996). Bishr *et al.* (1999) revised this concept considering a GIC to be a group of spatial data users and producers who share an ontology of the geographic world. This agrees with Guarino (1998) who distinguishes an application ontology from a generic knowledge base by considering the latter as a particular knowledge base that describes facts always true for a community of users.

Gruber (1992) defines an ontology as an explicit specification of a conceptualization. Guarino (1998), while agreeing with Gruber, presents a refined distinction between an ontology and a conceptualization: an ontology is a logical theory accounting for the *intended meaning* of a formal vocabulary, i.e., its *ontological commitment* to a particular *conceptualization* of the world. The intended models of a logical language using such a vocabulary are constrained by its ontological commitment. This commitment and the underlying conceptualization are reflected in the ontology by the approximation of these intended models.

Although the use of ontologies started with Artificial Intelligence, ongoing research on ontology can be found throughout the computer science community, in such areas as computational linguistics and database theory. It covers fields ranging from knowledge engineering, information integration, and objected-oriented analysis to such applications as medicine, mechanical engineering, and GIS. Frank (1997) discusses the use of explicit ontologies in systems development. Using ontologies to build GIS applications can help data integration and avoid problems, such as inconsistency between *ad-hoc* ontologies built into the system. However, there is a gap between the ontologies and the software components. To allow transfer of knowledge from ontologists (i.e., the specialists in the area of the application) to software engineers it is necessary to focus on the consistent part of an ontology instead of highlighting differences between ontologies. It is also necessary to exclude the historical and philosophical point of view. Both the engineering and the cognitive views of the world are necessary to produce the small theories that account for the behavior of certain parts of reality. The first is necessary to integrate engineering knowledge into the system, the second to make it understandable to the user at the interface (Frank 1997). Ontology plays an essential role in the construction of GIS, since it allows the establishment of correspondences and interrelations among the different domains of spatial entities and relations (Smith and Mark 1998).

2 RELATED WORK

The use of ontology in information system building is thoroughly discussed by Guarino (1998), and specifically in GIS building by Frank (1997) and Smith and Mark (1998). Gruber (1991) suggested that ontologies play a software specification role. Also Nunes (1991) pointed out that the first step to build a next-generation GIS would be the construction of a systematic collection and specification of geographic entities, their properties, and relations. However, philosophers and software engineers have different perspectives about ontologies. In this section we review work that deals with the use of ontology in the development of information systems, and with knowledge sharing.

2.1 Ontology and Software Development

According to Guarino (1998), there is a difference in the definition of ontology in the philosophical sense and in the way the term is used in the Artificial Intelligence (AI) field. In AI, ontology is seen as an engineering artifact that describes a certain reality with a specific vocabulary, using a set of assumptions regarding the intended meaning of the vocabulary words. Meanwhile, in the philosophical arena, ontology is characterized as a particular system of categories reflecting a specific view of the world. Smith (1998) notes that since, to the philosopher, ontology is the science of being, it is inappropriate to talk about a plurality of ontologies, as engineers do. To solve this problem, Smith (1998) suggests a terminological distinction between referent-based or reality-based ontology (R-ontology) and elicited or epistemological ontology (E-ontology). R-ontology is a theory about how the whole universe is organized and corresponds to the philosopher’s point of view. An E-ontology fits the purposes of software engineers and information scientists and can be defined as a theory about how a given individual (or group or language or science) conceptualizes a given domain. There are as many proper E-ontologies as there are GICs.

In order to build software components from ontologies, it is reasonable to assume that ontologies are available on the market. As ontology development technology evolves, the benefits of ontology use will outweigh the costs of developing them. With the success of this technology, large-scale repositories of ontologies will be available in diverse disciplines (Farquhar *et al.* 1996), and previous work has been developed based upon this availability assumption (Kashyap and Sheth 1996). As Frank (1997) assumes, we believe that there is a commercial production of ontologies, and that these ontologies are good enough to be used. This position is not shared by Guarino (1998), however, who believes that the available amount of ontological knowledge is modest, although of good quality. Kemp and Vckovski (1998) consider that although certain types of geographic phenomena, like discrete objects, have been the object of ontology study, spatially continuous phenomena, like temperature and soil moisture, have received little attention. Guarino (1998) suggests that it is more feasible today to use very generic ontologies, although this solution has the drawback of limiting the degree of reusability of the software components and knowledge. The other option is to use an ontology library containing specialized ontologies of domains and tasks. The translation of this library into software components reduces the cost of conceptual analysis and ensures the ontological adequacy of the information system. The solution presented here tries to shorten the gap between these two solutions by allowing navigation through ontologies of different specialization levels.

3 SPATIAL KNOWLEDGE

The issue of exchanging data and knowledge among GIS users is different from other information systems. The location component in present in geographic information is the basic difference. If we consider a hierarchy of information communities and the correspondent ontologies, reuse of data is done more horizontally and reuse of knowledge is done vertically and horizontally. It is more likely that a GIC needs data of neighboring community than from a distant one.

3.1 Ontology-Driven Geographic Information Systems

In ontology-driven geographic information systems (ODGIS), an ontology is a component, such as the database, cooperating to fulfill the system's objectives. The first step to build an ODGIS is to specify the ontologies using an ontology editor. The editor should be able to translate the ontologies into a formal language. The translated ontologies are available to be browsed, and therefore can be employed to provide the user with information about the knowledge embedded in the system. They can also be used as classes that contain data and operations that constitute the system's functionality.

ODGIS are built using software components derived from diverse ontologies. These software components are classes that can be used to develop new applications. Being ontology-derived, these classes embed knowledge extracted from ontologies. We present here how the system is built and how it is used. The system architecture for an ODGIS is shown elsewhere (Fonseca and Egenhofer 1999).

Representing geographic entities, either man made features or natural differentiations on the surface of the earth, is a complex task. As Smith (Smith and Mark 1998) argues, they are not merely located in space, they are tied intrinsically to space. For instance, boundaries that seem simple can in fact be very complex; an example is the contrast between country boundaries, which can be somewhat fuzzy, and land parcel limits whose boundaries are crisp. A user that develops an application can make use of the accumulated knowledge of experts that have specified an ontology of boundaries instead of dealing with these complex issues. The same is true for ontologies that deal with geometric representations, land parcels, environmental studies. Users can create new ontologies using the existing ones as a source. The user can specialize an ontology by substituting some of its members or can extend it through the inclusion of new members or relations.

Once the ontologies are specified they can be translated into software components. These components can be seen as frameworks or interfaces. Classes should be provided to perform the activities specified in the ontology. The framework can be used in the conceptual phase of the system. To develop applications the classes and the framework are necessary. The framework comprises the whole ontology and its component parts. The application developer can combine classes from diverse ontologies and create new classes that represent the user needs. This way a class that represents a land parcel for a specific city can be built from Land parcel component specified in Urban ontology, from polygon specified in Geometry ontology and from crisp boundary specified in Boundary ontology. So the real class is *land parcel of city X*, but it plays many roles that together give the class its unique characteristic.

3.2 The Role of Users in Information Sharing

The base of ODIS is the willingness of users to share knowledge and data. The reasons to do so can be economic or regulatory. Reusing data can decrease dramatically the costs of developing a GIS project. Reusing knowledge may decrease costs but, more importantly, it may mean the success of a project (Huxhold 1991). Neches *et al*. (1991) suggest that it is difficult to lower these costs and it is better to focus research on sharing the acquired knowledge. Sharing is the only way to build qualitatively bigger knowledge-based systems, because we can rely on previous labor and experience. Some high-level government institutions recommend the use of mechanisms that enhance the possibility of data sharing (Arctur *et al*. 1998).

For interoperability to take place, an agreement on the terminology in the shared area must occur through the definition of an ontology for each domain (Wiederhold 1994). Ontologies are crucial for knowledge interoperation and they can serve as the embodiment of a consensus reached by a professional community (Farquhar *et al*. 1996). Sharing the same ontology is a pre-condition to data sharing and data integration. Kashyap and Sheth (1996) consider that there should be an ontological commitment revealing the agreement between the generic user querying the database and the database administrator that made data available. In ODGIS, the agreement is expressed through the use of *elected ontologies* that are used to build new ontologies, from which the software components will be derived. An alternative to an explicit ontological commitment is the semantic approach. Bergamaschi et *al*. (1998) propose the derivation of a global schema to overcome the absence of a common shared ontology through the use of clustering techniques. The solution of semantic heterogeneities is done through description logic. Rodriguez *et al.* (1999) present a similarity assessment among ontologies using a feature-matching process and semantic distance calculations.

To clarify the idea of who are the users, it is better to group them in geospatial information communities (GIC). Bishr (1999) considers that the definition of a GIC should not be restricted to the data model sharing, but we should use common ontologies as the high-level language that holds those communities together. Therefore, a GIC is a group of users that share an ontology. In the solution that is presented here, we allow the GIC to commit to several ontologies and give the users means to share the common ontologies through the use of classes derived from ontologies.

3.3 Knowledge and Data Sharing in Urban Environments

Smith (1995) introduces bona fide and fiat objects, a classification based on boundary characteristics of geographic features. Bona fide objects have boundaries that exist independently of all human cognitive acts. These boundaries are parts of the geographic objects, like riverbanks or coastlines. Fiat boundaries, on the other hand, correspond to no genuine heterogeneity on the side of the geographic entities. They are abstract, created by acts of human decision, and usually related to laws or political decrees.

The urban environment is very complex and although it has some natural occurrences like rivers that are bona fide objects, it is essentially made of fiat objects. Even features such as rivers, when crossing urban environments, have their boundaries shaped by people and can be considered as fiat objects

Smith (1995) observes that most of us live in a hierarchy of *fiat* objects. These objects are specified in ontologies related to political divisions. National land use ontology can be defined in this fashion, initially at a high level. More detailed ontologies can be specified for other levels in the political hierarchy. At the national level, only large political divisions are recognized, along with federal legislation aspects on property ownership. At the next level, a state land use ontology would deal with specific details related to state legislation. Counties and municipalities can refine these notions further, adding specific knowledge-related local characteristics. Knowledge sharing is more likely to occur over this vertical axis, through the action of national associations like URISA, which can gather information and specify high-level ontologies. Guarino (1998) classifies these high-level ontologies as top-level, domain, or task ontologies. This kind of knowledge is easily transferable to information communities all over this vertical axis.

On the other hand, data sharing is more related to *bona fide* objects. The geographic location of these objects has a greater influence on data exchange than the political hierarchy, which determine *fiat* objects. When perceived at a large scale, even *fiat* objects, such as cities tend to behave more like *bona fide* objects with distinct physical characteristics. A county-level environmental study will consider data on neighboring towns, and a regional study will also consider geographically close states.

4 CONCLUSIONS AND FUTURE WORK

This paper presented the role that user communities play as the agents to specify ontologies that are used to create software components. These components created from ontologies enable knowledge and information sharing. These software components are derived from ontologies using an object-oriented mapping. The translation of an ontology into an active geographic information system component leads to *ontology-driven geographic information systems* (ODGIS).

This work requires an ontological commitment from users and information providers. User associations can be used as anchor points to start the production of ontologies. The result of this work can be shared later inside and outside the user community.

Further study should investigate how to incorporate approaches that allow composition of ontologies developed by different user communities, for instance, through the use of a context algebra to compose diverse ontologies (Wiederhold and Jannink 1999) and the matching of synonym, hyponym and hypernym terms (Kashyap and Sheth 1996; Mena *et al*. 1996; Mena *et al*. 1998).

REFERENCES

Arctur, D., Hair, D., Timson, G., Martin, E., and Fegeas, R. (1998) Issues and Prospects for the Next Generation of the Spatial Data Transfer Standard (SDTS). *International Journal of Geographical information Science* 12(4): 403-425.

Bergamaschi, S., Castano, S., Vermercati, S., Montanari, S., and Vincini, M. (1998) An Intelligent Approach to Information Integration. in: N. Guarino, (Ed.) *Formal Ontology in Information Systems*. IOS Press, Amsterdan, Netherlands.

Bishr, Y. A., Pundt, H., Kuhn, W., and Rdwan, M. (1999) Probing the Concepts of Information Communities - A First Step Toward Semantic Interoperability. in: M. Goodchild, M. Egenhofer, R. Fegeas, and C. Kottman, (Eds.), *Interoperating Geographic Information Systems*. pp. 55-70, Kluwer, Norwell, MA.

Farquhar, A., Fikes, R. and Rice, J. (1996) *The Ontolingua Server: a Tool for Collaborative Ontology Construction*. Knowledge Systems Laboratory - Stanford University, Stanford, CA, Technical Report KSL 96-26.

Fonseca, F. and Egenhofer, M. (1999) Ontology-Driven Geographic Information Systems. in: C. B. Medeiros, (Ed.) *7th ACM Symposium on Advances in Geographic Information Systems*, Kansas City, MO, pp. 14-19.

Frank, A. (1997) Spatial Ontology: A Geographical Point of View. in: O. Stock, (Ed.) *Spatial and Temporal Reasoning*. pp. 135-153, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Gruber, T. (1991) The Role of Common Ontology in Achieving Sharable, Reusable Knowledge Bases. in: *Principles of Knowledge Representation and Reasoning*, Cambridge, MA, pp. 601-602.

Gruber, T. (1992) *A Translation Approach to Portable Ontology Specifications*. Knowledge Systems Laboratory - Stanford University, Stanford, CA, Technical Report KSL 92-71.

Guarino, N. (1998) Formal Ontology and Information Systems. in: N. Guarino, (Ed.) *Formal Ontology in Information Systems*. pp. 3-15, IOS Press, Amsterdam, Netherlands.

Huxhold, W. E. (1991) *An Introduction to Urban Geographic Information Systems*. Oxford University Press, New York, NY.

Kashyap, V. and Sheth, A. (1996) Semantic Heterogeneity in Global Information System: The Role of Metadata, Context and Ontologies. in: M. Papazoglou and G. Schlageter, (Eds.), *Cooperative Information Systems: Current Trends and Directions*. pp. 139-178, Academic Press, London.

Kemp, K. and Vckovski, A. (1998) Towards an Ontology of Fields. in: *Third International Conference on GeoComputation*, Bristol, UK.

McKee, L. and Buehler, K., Eds. (1996) *The Open GIS Guide*. Open GIS Consortium, Inc, Wayland, MA.

Mena, E., Kashyap, V., Illarramendi, A., and Sheth, A. (1998) Domain Specific Ontologies for Semantic Information Brokering on the Global Information Infrastructure.

in: N. Guarino, (Ed.) *Formal Ontology in Information Systems*. pp. 269-283, IOS Press, Amsterdam.

Mena, E., Kashyap, V., Sheth, A., and Illarramendi, A. (1996) OBSERVER: An Approach for Query Processing in Global Information Systems based on Interoperation across Pre-existing Ontologies. in: *First IFCIS International Conference on Cooperative Information Systems (CoopIS'96)*, Brussels, Belgium, pp. 14-25.

Neches, R., Fikes, R., Finin, T., Gruber, T., Patil, R., Senator, T., and Swartout, W. (1991) Enabling Technology for Knowledge Sharing. *AI Magazine* 12(3): 13-36.

Nunes, J. (1991) Geographic Space as a Set of Concrete Geographical Entities. in: D. Mark and A. Frank, (Eds.), *Cognitive and Linguistic Aspects of Geographic Space*. pp. 9-33, Kluwer Academic Publishers, Norwell, MA.

Rodriguez, A., Egenhofer, M. and Rugg, R. (1999) Assessing Semantic Similarity Among Geospatial Feature Class Definitions. in: A. Vckovski, K. Brassel, and H.-J. Schek, (Eds.), *Interoperating Geographic Information Systems - Second International Conference, INTEROP'99, Zurich, Switzerland. Lecture Notes in Computer Science* 1580, pp. 1-16, Springer-Verlag, Berlin.

Smith, B. (1995) On Drawing Lines on a Map. in: A. Frank and W. Kuhn, (Eds.), *Spatial Information Theory—A Theoretical Basis for GIS, International Conference COSIT '95, Semmering, Austria*. *Lecture Notes in Computer Science* 988, pp. 475-484, Springer Verlag, Berlin.

Smith, B. (1998) An Introduction to Ontology. in: D. Peuquet, B. Smith, and B. Brogaard, (Eds.), *The Ontology of Fields*. pp. 10-14, NCGIA, Bar Harbor, ME.

Smith, B. and Mark, D. (1998) Ontology and Geographic Kinds. in: *International Symposium on Spatial Data Handling*, Vancouver, Canada, pp. 308-320.

Wiederhold, G. (1994) Interoperation, Mediation and Ontologies. in: *International Symposium on Fifth Generation Computer Systems (FGCS94)*, Tokyo, Japan, pp. 33-48.

Wiederhold, G. and Jannink, J. (1999) Composing Diverse Ontologies. in: *8th Working Conference on Database Semantics (DS-8)*, Rotorua, New Zealand.