Exploring eGov Cooperation and Knowledge Sharing using Geospatial Ontologies in a Semantic Wiki

Gary Berg-Cross, Ph.D.
Engineering Management & Intgegration (EM&I)
Suite 350 455 Spring park Place
Herndon VA 20170
703-742-0585
Gary.Berg-Cross@em-i.com

Abstract

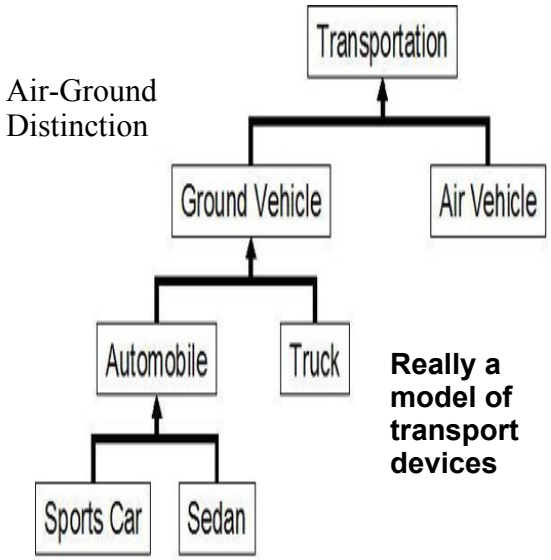
Traditional Wikis are currently used to support eGovernment (eGov) goals of better use and sharing services across the government through the capture of notes and activities for such things as enterprise architecture (EA) and service oriented architecture (SOA). However, the job of gaining a unified view of an enterprise's knowledge assets across government remains difficult to implement in practice because current Wiki approaches are too ad hoc to provide integration. Combining semantic technology with the Wiki ease of use seems likely to help organize eGov webs of collaborative knowledge and information seems. However, the resulting complexity makes meaningful implementation a challenge especially in light of the need to migrate existing eGov information to a new form. This paper describes a Semantic Wiki for geospatial information and knowledge for use by the Spatial Ontology Community of Practice (SOCoP). A Semantic Wiki is being used to help foster collaboration among researchers, technologists & users of spatial knowledge representations and reasoning with the goal of developing spatial ontologies for use by across in the Semantic Web. The intent is to use the Wiki to support open collaboration with open standards for increased interoperability of spatial data across government. The paper discusses plans, the initial approach and examples of early work and why the combination of community expertise and the semantics of geospatial knowledge may provide a good venue to study capture sharing of knowledge via the Semantic Web and its technologies.

1. Introduction

EGovernment (eGov) has firmly established itself as a valuable Information and Communications Technology (ICT) approach alongside of 'e-commerce' and 'e-business'. For example, experience with eGov portals and intranets have shown them to be valuable repositories of corporate knowledge. More recent work at the federal level continues as promoted by the eGov Act of 2002 which established an Intergovernmental Committee on Government Information (ICGI) and Data Integration Pilots, Federal Enterprise Architecture work such as the Data Reference Model (DRM) and its Data Management Strategy to enable Intergovernmental Data Exchange. This paper briefly reviews three key efforts in the eGov thrust –enterprise architecture (EA) , service-oriented architecture (SOA) and Wikis. It is argued that progress in each is dependent on improved semantics some of which can be tested in focused communities of practices using ontologies to enable better semantics.

2. Foundations for eGov Architecture and Supporting Collaboration with Wikis

In the last 10 years two important architectural approaches have become influential in government/eGov circles as a way to bridge diverse interests break down institutional barriers. These are enterprise architecture (EA) and service-oriented architecture (SOA). EAs and their Business, Information and Technical component architectures support eGov by helping to control ad hoc applications and data modeling across the government. EA components and their construction is defined and organized by EA frameworks, such as the Zachman Framework & FEAF, which are build on a mix of IM ERA and BP models. However, as argued in Berg-Cross (2007) EAs have several problems. One is that properties of a target EA are clearer than the path to them. SOAs are one possible vision. However, EA visions tends to be strategic diagram, or simple top-level lists which don't adequately ground SOA implementation. Another problem is that the meta-models used to capture architecture are typically semantically weak (Sowa & Zachman 1992). Most EAs are based as much on natural language descriptions as structured models. For example, in the Zachman framework the EAs typically include strategic and requirement information at the top level and this is only informally tied to the conceptual and logical levels below it. As a result of the use of conventional IT formalisms, EA models leave implicit many of the details required to understand one architecture and integrate it with others. This means that EA products are somewhat ad hoc in how they organize their architecture to solve strategic priorities such as data sharing across the government. However, EAs are moving incrementally towards better semantics. One example of this is the previously mentioned Federal data reference model (DRM) which has added taxonomies for controlled meaning. But this is still somewhat "piecemeal" since it uses RDF just one set of semantic standards in the semantic web family with varying formality and semantic expressiveness. This shown in the example of a DRM taxonomic example below. The taxonomy is formalized on the right using XML and RDF, but what is formalized is a rather informal hierarchy of transportation concepts (really transportation devices). For example, the taxonomy covers ground transportation devices but doesn't allow trains and autos as a different sub-type. Nor does it distinguish between self powered. (Bikes, Wheelchairs) and powered vehicle. It represents a pseudo-formalization which is not based on a deep conceptualization and categorization of the domain in terms of distinguishing properties or systematic relations between levels. This is not an uncommon problem and reflects the lack of the necessary conceptual analysis going into EAs and Service models**.** These reflect different conceptualizations & analytic methods, not just differences in formalisms.



<?xml version="1.0"?> <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:xsd="http://www.w3.org/2001/XMLSchema#" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:daml="http://www.daml.org/2001/03/daml+oil#" xmlns="http://www.owl-ontologies.com/unnamed.owl#" xmlns:dc="http://purl.org/dc/elements/1.1/" xml:base="http://www.owl-ontologies.com/unnamed.owl"> <owl:Ontology rdf:about=""/> <owl:Class rdf:ID="Transportation"/> <owl:Class rdf:ID="AirVehicle"> <rdfs:subClassOf rdf:resource="#Transportation"/> Etc.

The emergence of web services and service oriented architectures also promise support for the integration of eGovernment services, processes and applications. With a network, service classes are not on the same machine. Now a service class can find what it needs via an explicit Service Description and a class send its information (passes values) to the other class via XML/XMLS serialization. SOA promises improved delivery of eGovernment services and there is a natural synergies between the disciplines of SOA and EA because SOA represents a target example of doing "good enterprise architecture". That is, SOA helps realize that architectural vision. Thus we may view SOA as a subset of EA, because SOA represents an architecture style of designing application architecture. As an architecture SOAs also need improved semantics to fulfill the needs of eGov for system and information interoperability. A number of efforts have been launched to meet this need. The Object Management Group (OMG), for example, developed Model Driven Architecture™ (MDA™). MDA enables for robust development of services by more sophisticated use of system models in the software development process, and it supports reuse of best IT modeling practices when creating families of systems as a way of modeling business process being supported by services. A major principle underling the OMG's view of MDA is that models need to be expressed in a well-defined notation to understand systems for enterprise-scale solutions. Such work is moving slowly towards an adequate smenatics, but looking at the continuum of Semantic Web formalism employable for SOA Berg-Cross (2007) provides examples like the previous taxonomic one showing that proper semantic analysis remain a problem with making services and data usable across the eGov enterprise.

Wikis are a third innovation being widely applied to a useful tool for collaborative note-taking. capture information about both EA and SOA. Compared to EA and SOA wikis are relatively informal and serve as easy repositories of information for community groups as typify by ONTOLOG & COLAB. COLAB work is supported .by the GSA Intergovernmental Solutions Division for the purpose of Networking Among Communities of Practice, including the Interagency Collaborative Expedition Workshops.. Traditional Wikis allow collaborative editing of content and have very simple markup/syntax. This enables user authoring via a web browser. The content of a Wiki is defined by a page graph including page metadata and the actual content of pages. Wikis tend to be used increasingly for eGov purposes, but traditional Wikis are like portals and can quickly become hard to communicate with data islands. So will Wiki systems enable collaborative authoring their manually defined/maintained set of categories do not provide an effective mechanism for search and retrieval of published eGov content. Their ad hoc nature also means that while they are open for people to browse and contribute, they are essentially closed for automated use. Recently, Wikis have been combined

with Semantic Web technologies to create what are called Semantic Wikis which allow users to annotate/tag Wiki content (pages or parts of pages) with more standardized metadata in the form that makes them processable to varying degrees by applications. To some, including the COLAB Wiki, combining the user-friendliness of Wiki s with Semantic Web technologies would be a step forward to for effective publishing of eGov information on the Web. This approach, broadly called Semantic Wikis is discussed in the next section.

Semantic Wikis

A page in a semantic Wiki can be thought of as have 3 parts – the original structured text, formal annotations in some controlled vobabulary and the link from these annotation tags to particular parts of the structured text. This makes Wiki page exchange possible on the basis of the annotations if the annotations are semantically well founded. A key challenge for Wikis is how to use the power of semantic technologies for organizing and retrieving Wiki knowledge while keeping them easy enough for a community to use. Wikis attack this problem by extending the concept of Wiki annotations. Semantic Wikis typically create a "knowledge layer" or overlay network structure that defines concepts, attributes, and relationships of the underlying content of the Wiki . Relationships become explicit as links. For example, RDFWiki (Palmer) provides users with a simple text-based interface to edit content and metadata and stores all data as RDF statements. The SemanticMedia Wiki (Krotzch et al, 2005) is an extension of MediaWiki (the Wiki platform used by the Wiki Pedia community (Wiki pedia.org) that allows users to add metadata understandable by automatic processes too, while Rhyzome (Souzis, 2004) allows users to express RDF statements through a simplified syntax (ZML). In some Semantic Wikis, such as SemWiki , users can make semantics annotations, that are not bound to a particular page. Rather, the annotations belong to the Wiki as a whole. As seen in these summaries, RDFS is often used to tag Wiki annotations. Semantic Wiki work is relatively new and particular annotations can differ in expressive power, simplicity, and meaning. What is needed is a elaborate conceptual model for semantic annotations, introduce a unique and rich Wiki syntax for these annotations, and how to best formally represent the augmented Wiki content.
RDFS is widely used in current work but lacks sufficient semantics for the difficult parts. For example, formal annotations need an agreement about which formal identifier stands for particular real-world artifact. With RDF, these problems get even worse because URIs are used for formal identifers which are a superset of URLs, which are locations of real-world web resources, e. g. an HTML page. This problem has been called the URI "crisis". Other problems with RDFS have been noted in Berg-Cross (2007) including its deficiencies as too informal an ontology language. For example its vocabulary allows annotating that <Human,type,Species> and <Amber,type,Customer> which are very different meanings that arise because RDFS hasn't distinguished between classes (Human) and instances (Amber). Better semantics such as provided by ontologies is needed to adequately ground the RDF vocabulary and expand it to needed concepts. Use of ontologies would help semantically interconnect and enhance data and service definitions and descriptions by alignment with domain and system reference ontologies. A community of practice (CoP) for geospatial ontologies is discussed in the next section as an area where suitable testing of this idea might be conducted

3. GeoSpatial Knowledge, Standards and SOCoP

Geospatial information has a wide range use across eGov and supports many missions (cross-cutting), including national security, law enforcement, health care, the environment, and natural resources conservation**.** At least 25 Federal departments and agencies independently collect or produce geospatial information, or invest in potentially duplicative geospatial capabilities. Some notion of the breath can be understood by a simple listing of some topic areas and concepts within the Geospatial Profile: boundaries, location, elevation oceans, transportation, economy, structure, imagery basemaps, inland waters, telecomm, climatology, meteorology etc. Coordination of such concepts is aided by the

Federal EA which has a Geospatial Profile as well as a Geospatial Line of Business. Each was set up by the FGDC Steering Committee to establish a more coordinated approach to producing, maintaining, and using geospatial data and services as well as to ensure sustainable participation from Federal partners to establish a collaborative model for geospatial-related activities and investments such as integrating:

- Maps and map visualization
- Features and feature geometries
- Geographic and other relationships
- Coordinate and other reference systems

However, as previously noted regarding EAs they lack formal semantics to make such collaboration and interoperability easy. A major effort has been investedin standards and reflecting its diversity geospatial information crosses a wide spectrum of applicable geospatial standards (via W3C, ISO, OGC, etc.). And in the geospatial realm "Standards are great, because there are so many of them." To understand why there are so may standards we can think of different conceptual levels for geospatial information (Lieberman, 2007) running from the way geography is perceived to a yet to be developed semantic web of geospatial information.

- Geography: perception of the terrain
- GIS: adding information to features Cartography: symbolic representation of the terrain
- GeoWeb: connecting features across the Web
- Google Earth: the terrain as video game
- GeoRSS: adding features to information
- Geospatial Semantic Web: forming and distributing rich geospatial relationships across the Web

Integrating these is challenging and requires related sets of standards. As an example the Open Geospatial Consortium (OGC[1]) has promoted a number of OpenGIS® Standards with specifications including Geographic Objects, Location Services, Coordinate Transformation, and several Feature standards. Of particular interest are Open geospatial Web Services" (OWS), which OGC has been developing using specifications for a suite of Web services and associated encodings to expose geospatial content and operations from distributed content repositories to remote clients across diverse platforms, These standards tested in a series of OWS demonstrated include:

GML - geographic markup language (an information model and XML schema) for encoding features (geometric representations of geography).

Web Feature Service - service providing access to collections of features

Web Map Service - service providing access to map layers (cartographically rendered features and images)

Catalog Service / Web - service supporting (spatial) discovery of geospatial datasets and services

Most recently the OWS-4 demonstration employed simple ontologies to help integrate information across different databases and the different standards. Ontologies are a foundational backbone technology for the Semantic Web because they prove a basis for the management of formalized knowledge within the technical context of distributed information across the Semantic Web. Recently, reflecting the broadening interest in the use of ontologies with the geospatial community a Spatial Ontology Community of Practice (SOCoP) was chartered as a Community of Practice under the Best Practices Committee of the Federal CIO Council. SOCoP's purpose and focus is to foster collaboration among researchers, technologists & users of spatial knowledge representations and reasoning towards the development of a set of core, common spatial ontologies for use by all in the Semantic Web. As a Community of Practice SOCoP encourages open collaboration and the use of open standards. SOCoP represents a strategic investment for ontology development, building on core ontological competencies, documenting best practices within the community. SOCoP developed ontologies will

[1] OGC is a non-profit, international, voluntary consensus standards organization that is leading the development of standards for geospatial and location based services.

offer increased interoperability of spatial data across government and create opportunities to partner with other cross domain and ontology COP groups. As an aid to collaboration across the community a SOCoP Wiki was established: (http://www.visualknowledge.com/wiki/socop) Initially hosted by Visual Knowledge® Software, Inc. this is a traditional Wiki, but is intended to use Visual Knowledge "Semantic Wiki" to become a fully integrated Web 3.0 development and execution platform for building semantic Suites, semantic blogs and high performance knowledge-driven applications. An initial step has been to provide a set of ontologies including:

- Domain Ontology such as for Airports and Airplanes
- Base Geospatial Ontology such as Geometries originally expressed in GML
- Filter Ontology for Spatial Relationships
- Feature Ontologies including AIXM, DAFIF, Gazetteer (Backe, Berg-Cross, Kolas and Moeller, 2007)

Each of these represents sound work on ontology or the lifting of weaker formalisms into an ontological form. This is somewhat analogous to what should be done for Wiki annotations lifting them above the ad hoc or even RDF form into a richer ontology. Thus the early SOCoP experience is a good basis that might transfer into Semantic Wiki design. This is even clearer when we consider what current SOCoP work supports:

- A user asking a query in the vocabulary of his or her own perspective
- Automatic query decomposition to original data source concepts
- Automatic discovery of appropriate data sources
- Ultimately, geospatial data interoperability that is transparent and useful to the user
- Use of different ontologies to address the problem modularly

Each of these capabilities is desirable in a Semantic Wiki.Does a currently supported "geospatial logistics" query such as "What airports within 50000 meters of Saint Louis can support a C5?" can also be adapted to support, "What SOCoP ontology posted in 2007 contains boundaries for familiar objects."

5. Summary and Next Steps for Ontological COP as a Semantic Wiki Testbed

SOCoP inherits some ontologies from prior work but has room for the community to debate, argue or vary on their value. In this sense it is intended as a iterative, living ontology relies chiefly on its community, though regulated by an executive group. As a strategy a community devoted to building ontologies is in a great position to attempt this if for no other reason than that its users are already familiar with ontological concepts and their development. SOCoP is particularly suited because the geospatial domain is widely relevant but has a theoretical core of concepts well developed and focused. For all of these reasons SOCoP represents a practical testbed for semantic Wikis. It is also useful for demonstrating ontology design that is:

- Meaningful - all named classes can have instances
- Formal –can be represented/put into a form amenable to automated processing
- Rigorous – stands up to rational analysis
- Correct - captured intuitions of domain experts
- Minimally redundant - no unintended synonyms
- Sufficiently axiomatized – include detailed constraining descriptions as axioms

Finally it can also demonstrate the utility of ontologies for eGov by enabling semantic translation and support standardization efforts such as the Federal Enterprise Architecture and its information models including information discovery, data repurposing, integration of data sources, grounding SOA and enabling the semantic web.

References

Backe, Kevin. Berg-Cross Gary, Kolas, Dave and Moeller, John. SOCoP, Introduction to Spatial Ontologies, presented at USGIF/NGA Tech Days - Emerging Technology Showcase, McClain Va, May 16, 2007

Berg-Cross, Gary 'Incremental Semantics for Service Oriented Architecture' SOA EGov COP May 2007

COLAB: An Open Collaborative Work Environment (CWE) to Support NetworkingAmong Communities of Practice, http://colab.cim3.net/cgi-bin/wiki.pl?WikiHomePage

Krotzch Markus, Denny Vrandecic, and Max Volkel. "Wiki pedia and the Semantic Web The Missing Links". In *Proceedings of Wiki mania 2005*, Frankfurt, Germany, August 2005.

Lieberman, Joshua Geospatial Semantic Web:Is there life after geo:lat and geo:long ? Presented at European Geoinformatics Workshop 2007

ONTOLOG Forum http://ontolog.cim3.net/Wiki /

Palmer Sean B. "RDFWiki ". http://infomesh.net/2001/rdfWiki /.

SICoP Semantic Wiki s and Information Management Working Group, http://colab.cim3.net/cgi-bin/Wiki .pl?SICoP/SemanticWiki sandInformationManagementWG

Souzis A. "Rhizome position paper". In Proceedings of the *1st Workshop on Friend of a Friend, Social Networking and the Semantic Web*, September 2004.

Sowa J. and Zachman, J. Extending and For- malizing the Framework for Information Systems Architecture. IBM Systems Journal 31(3). 1992

Max Völkel, Eyal Oren, **Towards a Wiki Interchange Format (WIF) - Opening Semantic Wiki Content and Metadata.** In *Proceedings of the Workshop on SemWiki* Workshop at ESWC, 2006

Wiki pedia.org. Wiki pedia Home Page. http://www.Wiki pedia.org