Distributed Ontology Architecture for Knowledge Management in Highway Construction

T. E. El-Diraby[1] and K. F. Kashif[2]

Abstract: The ongoing plethora of rehabilitation in the infrastructure domain requires more planning and integration during design and construction. To achieve this, there is a need for developing and using semantic (ontology-based) mechanisms for the exchange of development knowledge among all project stakeholders. This paper presents a distributed ontology architecture for knowledge management in highway construction. With every other utility tied to the highway geometry, the architecture is intended to be the base for a cross-discipline knowledge exchange in the infrastructure domain. The architecture presents highway knowledge on three levels: domain knowledge (an umbrella for infrastructure shared knowledge), application knowledge (representation of highway-specific knowledge), and user knowledge (an enterprise-specific representation of highway knowledge). The proposed architecture models highway concepts using six major root concepts: project, process, product, actor, resources, and technical topics (attributes and constraints). The architecture was developed using rigorous knowledge acquisition and ontology development techniques. It was developed as an extension for the *e-COGNOS* ontology. The architecture was validated through input from domain experts.

DOI: 10.1061/(ASCE)0733-9364(2005)131:5(591)

CE Database subject headings: Highway construction; Information management; Infrastructure; Knowledge-based systems.

Introduction

Semantic Web is an extension of the current web in which information is accessed based on meaning (not hypertext), better enabling computers and people to exchange knowledge (not data). Ontology is one of the main components of semantic web (Berners-Lee 2001). In the context of information management, ontology is a "formal explicit specification of a shared conceptualization" (Gruber 1993). Ontology provides a mechanism to categorize/classify domain knowledge items/information into inter-related concepts. Ontology structures information in the form of concept hierarchies (taxonomies), axioms, and semantic relationship. This structuring allows natural language to be presented in an unambiguous form to the computers.

Ontologies representation of information is superior to resource description framework (RDF). The RDF defines knowledge in terms of triples, consisting of a concept (resource) with its properties and relations (predicate) and certain restrictions (values). In contrast, a typical ontology structure "O" is formed if the following set is specified (Maedche 2002):

$$O = \{C, H^c, R, R^t, A^o\} \tag{1}$$

where O=ontology structure; C=disjoint set whose elements are called concepts: $X_1, X_2, X_3 \ldots . X_n$, H^c=class hierarchy representing a directed relation such that $H^c \subseteq C \times C$, where H^c (C_1, C_2) means that C_1 is a subconcept of C_2; R^t=disjoint set whose elements are called taxonomical relations: $Y_1, Y_2, Y_3 \ldots . Y_n$, R =nontaxonomic relation between concepts: $R \rightarrow C \times C$; and A^o =set of ontology axioms, represented in first order logic.

The five structural elements of ontology $\{C, H^c, R, R^t, A^o\}$ provide adequate information representation necessary for semantic web.

There is a need to develop semantic presentations for infrastructure knowledge. This will facilitate web-based integrated development in the infrastructure domain. Like many other industries, several research projects are underway to formalize a set of construction ontologies (see for example, Katanuschkov et al. 2002). Nonetheless, little work has been done in the infrastructure domain.

No single ontology will be able to fully cover a domain nor will it satisfy the needs and preferences of each user (Gruber 1993; Guarino and Welty 2000; Maedche et al. 2003b). Consequently, the research team focused on developing a distributed highway ontology (HiOnto) architecture that will allow for future evolutions (Maedche et al. 2003a). HiOnto is based on the *e-COGNOS* ontology (Lima et al. 2003). *e-COGNOS* models construction concepts using six major classes, which have been retained by HiOnto. Furthermore, HiOnto was designed to be consistent with existing classification systems (for example, BS6100, Master Format, and UniClass) to enhance its usability. Currently, HiOnto includes a comprehensive taxonomy of highway construction terms (about 2,800 concepts), a mapping to *e-COGNOS* domain ontology, a set of subdomain application ontologies (including relationships and axioms), and provision for links with proprietary enterprise models.

[1]Assistant Professor and Director, Dept. of Civil Engineering, Center for Information Systems in Infrastructure & Construction, Univ. of Toronto, Toronto, Canada, M5S 1A4.

[2]Graduate Research Assistant, Dept. of Civil Engineering, Univ. of Toronto, Toronto, Canada M5S 1A4.

Note. Discussion open until October 1, 2005. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on October 13, 2003; approved on November 17, 2004. This paper is part of the ***Journal of Construction Engineering and Management***, Vol. 131, No. 5, May 1, 2005. ©ASCE, ISSN 0733-9364/2005/5-591–603/$25.00.

J. Constr. Eng. Manage., 2005, 131(5): 591-603

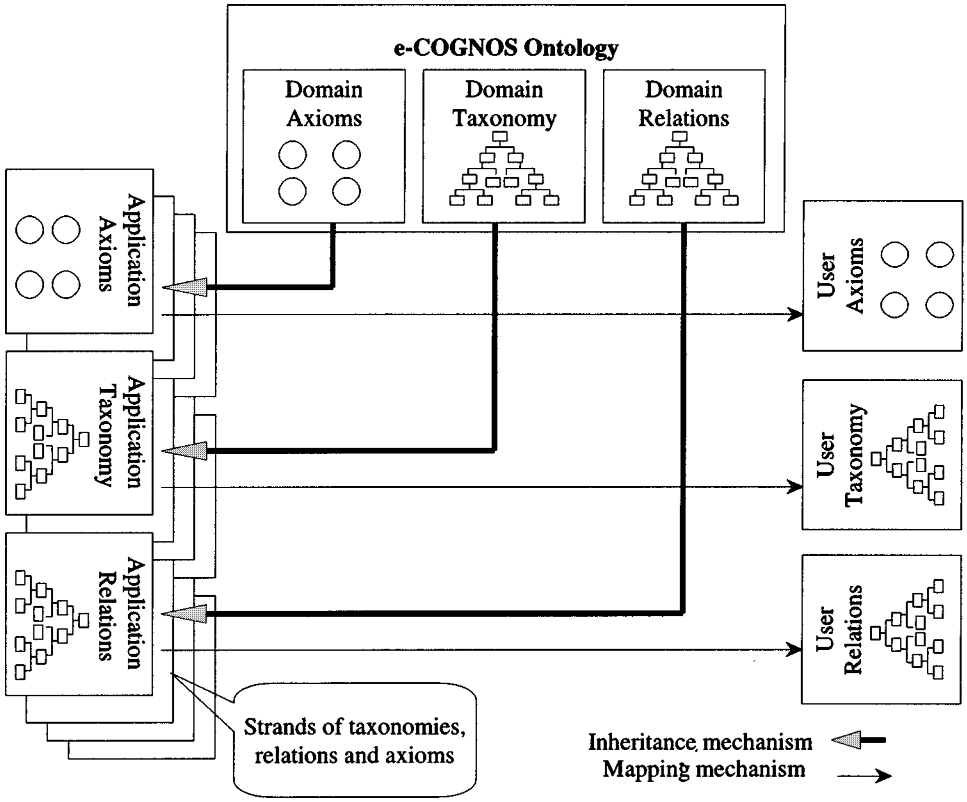


Fig. 1. HiOnto architecture

The architecture presented here pertains to highway construction. Given that highways are the geographic backbone of any infrastructure system, HiOnto can be the basis for an extended infrastructure construction ontology. On the down side, the dominance of highways in infrastructure systems makes it very hard to model highway concepts without extending research scope into a multitude of other domains (which range from tunneling, to urban planning to intelligent transportation systems). The research project has developed a set of major concepts in these related domains and linked them to their counterparts in HiOnto. Future research should extend to the linkage between HiOnto and other related domains.

In the immediate term, this architecture helps researchers in the field to create shared representation of relevant design and construction concepts in a semantic way. This will allow more effective collaboration during project design and construction. It also provides a needed foundation for building web services to support integrated project communications among its shareholders. In the long term, application software that is based on consistent representation of knowledge could have positive impact on the efficiency of the industry supply chain and overhead reduction. This will be achieved through avoiding overspecification; enhancing the interface between design and construction; using more standard and preassembled components; involving suppliers in the design process; and using more off-site manufacturing and assembly (Atkin et al. 1995).

HiOnto Architecture

HiOnto was designed as an object-oriented architecture. It includes a full ontological description of the main entities of highway construction. In HiOnto, an entity refers to a project, a process, a product, an actor, or a resource. A set of subdomain ontologies were developed within HiOnto spanning three levels of abstraction: domain, application, and users (Guarino 1997; Uschold and Jasper 1999). Each level either inherits or maps the concepts, relations, or axioms of the preceding level(s). Fig. 1 shows the general architecture of HiOnto, which is explained below.

Domain Level

This is the most abstract level in HiOnto. It represents an umbrella for assuring interoperability between HiOnto and any future application ontologies in the infrastructure domain (for example, telecommunication, gas utilities, or electrical utilities). This level contains the *e-COGNOS* domain ontology, which defines six main domains for describing construction concepts: project (PJ), process (PR), product (PD), actor (AC), resource (RE), and technical topics (TT). Within these abstract domains, *e-COGNOS* ontology defines a set of taxonomies for major construction concepts (*C*), their higher level relationships, and some of the most fundamental axioms (we use underlined text to donate relationships). For example, it defines the following processes: design process, bidding process, field construction process, billing process, and payroll process. The basic ontological model of *e-COGNOS* ontology is a process centered, whereby actors are involved in processes that are part of a project. Such processes utilize resources and are constrained by boundary conditions (such as code, company objectives, weather), which are detailed in the technical topics domain.

Some of the basic relations at the domain level of HiOnto include the following:

1. The relationship "is part of" assigns a project to a process: "is part of" (PJ, PR).
2. The relationship "output" assigns a product to a process (as its output): "output" (PD, PR).
3. The relationship "involves" assigns an actor to a process: "involves" (AC, PR).
4. The relationship "utilizes" assigns resources to a process (as its input): "utilizes" (RE, PR).
5. The relationship "constrained by" assigns some constrain (from TT) to a process: "constrained by" (TT, PR). These

major concepts and relationships are also bound by some axioms. Examples are shown as follows.

Temporal Control Axioms

At all times, each entity has a state, a stage, and a situation (all could be denoted by σ)—each has a specific time (t) where it starts. These σ's are discrete ordered events (if one σ is followed by another σ, then one cannot assume an intermediate σ between these two, unless the structure of all σ's is changed). All σs evolve monotonically away from the initial σ:

$$\forall(a,\sigma,t,t'),\ \text{start}\ (\sigma,t) \wedge \text{start}\ (\text{do}\ (a,\sigma),t') \supset t < t'$$
$$\rightarrow \text{start}\ ((\text{do}\ (a,\sigma),t) < \text{start}\ (\text{do}\ (a,\sigma),t')) \supset t < t'$$

where t=time; a=a situation.

Aggregation Axioms

The state of a process (σ_{PR}) is the collection of the state of all actors (σ_{AC}), products (σ_{PD}), and resources (σ_{RE}) involved in the process

$$\sigma_{\text{PR}} = \{(\sigma_{\text{PD}_1}, \sigma_{\text{PD}_2}, \ldots \sigma_{\text{PD}_m}) \cup (\sigma_{\text{AC}_1}, \sigma_{\text{AC}_2}, \ldots \sigma_{\text{AC}_n}) \cup (\sigma_{\text{RE}_1}, \sigma_{\text{RE}_2}, \ldots \sigma_{\text{RE}_k})\}$$

Causality

"The causal and temporal structure of states and [subprocesses] of a [process] should be explicit in the representation of the [process], then (Gruninger and Fox 1996)"

- $\forall(a,\sigma_1,\sigma_2) \rightarrow \sigma_1 < do(a,\sigma_1) \equiv \sigma_1 \leqslant \sigma_2$
- $\forall(a_1,a_2,\sigma_1,\sigma_2) do(a_1,\sigma_1) = do(a_2,\sigma_1) \supset a_1 = a_2$
- $\forall(\sigma_1,\sigma_2) \rightarrow \sigma_1 < \sigma_2 \supset \neg\sigma_2 < \sigma_1$

Application Level

The second level of HiOnto architecture contains a set of application ontologies. While domain level provide an umbrella of generic terms that are relevant to construction engineering at large, this level models the specific domains of highway construction (see Fig. 1). This level contains a set of distributed ontologies for project, process, product, actor, and resource (the sub domains of entity). These ontologies cover, for example, certain processes (such as testing, design, planning), common actors [such as field engineer, heating, ventilation, and air-conditioning (HVAC) consultant], or common products (such as bridge, tunnel). Additional ontologies were developed to describe several technical topics such as cost, productivity, quality, safety, constructability, soil tests, design code, design specifications, etc.

Each ontology is composed of three major components (think of them as DNA strands): taxonomy, relationships, and axioms. The entity ontologies form a grid (a 5×3 matrix whose rows are the five subdomains of entity and the columns are the three DNA strands). Process ontologies represent the key row in the matrix, while ontologies for project, actor, product, and resource present the required supporting ontologies to put process ontologies in perspective. In this schema, an actor could have one role in process *A* and another role in process *B*. A resource may be produced by process *C* and consumed by process *D*.

Orthogonal to this matrix are the ontologies for Technical topics, which could be applicable to any of the previous ontologies through proper relations. In general, technical topics ontologies model the attributes and constraints of entity ontologies. The following example shows how to represent some aspects of design process (attributes/constraints):

1. A design process (D_{pr}), which is part of the process ontology, has to conform to a set of design objectives (D_o), which are parts of technical topics ontologies: $\forall D_{\text{pr}} \exists D_o | \text{conforms\_to}(D_{\text{pr}}, D_o)$
2. A design process has to comply with specific design codes (D_c), which are part of technical topics: $\forall D_{\text{pr}} \exists D_c | \text{complies\_with}(D_{\text{pr}}, D_c)$

User Level

The third level is the user ontology (see Fig. 1). These are enterprise-specific ontologies. They basically connect the enterprise model, where specific roles are assigned to real people and real projects (more like instants), to the application ontology. Specific relations and axioms at this level could take the following format:

1. VP (whose name and attributes are defined in the enterprise model) to be the head of safety committee (which is a concept in the safety ontology): Head_of (VP, Safety_Committee);
2. required annual growth equal 10%: set (Target_Annual_Growth, 10%); and
3. between Time $t1$ and $t2$, enterprise A is the sole supplier of Material B: Between $(\text{A}, t1, t2) | \text{supplier\_of}(\text{A}, \text{B})$.

Taxonomy Development (R^t and H^c)

A comprehensive highway taxonomy had to be developed as a common platform over which application ontologies could interact. Taxonomical relations (is a, is kind of, is a type of) enable every sub-concept to inherit the properties of its super concept. It should be noted that in HiOnto it was assumed that "is-part-of" and "is-kind-of" relations convey the same meaning as conveyed by "is-a" relation. Future research will attempt to explicitly use Fuzzy Sets to better define the variances between these semantic relationships.

The use of taxonomies allows computers to derive new knowledge from existing knowledge. In the following example, based on the first and second sentences, a computer system could infer the third sentence:

1. Roads and highways are construction products;
2. Arterial roads are one type of highway; and
3. Arterial roads are construction products.

A glossary of about 4,000 terms was built by extracting terms from well-established sources, such as construction handbooks, textbooks, research papers, and informal interviews with experts. Through a process-centered analysis and using a set of competency questions (CQ), concepts were categorized into *e-COGNOS* main domains: process, project, product, actor, resource and,

J. Constr. Eng. Manage., 2005, 131(5): 591-603

Table 1. Process Competency Questions

Competency questions (CQ)	Answers to CQ	Abstraction
What processes are involved?	Designing/layout development	Process
Who performs the processes?	Engineer/designer	Actor
Where/when is the process performed?	Project X; location Y, time Z	Project, technical topics
What are the results of the processes?	A design/drawings/specs./work plan, etc.	Products
What is needed to perform the processes?	Designer/software	Resources
What affects/controls process performance?	Productivity/safety/sustainability/ code/geography etc.	Technical topics

technical topics. The CQ is a commonly used technique to assure consistent categorization and modeling of ontologies. Table 1 illustrates a subset of the CQ used for categorizing the concepts related to "the process of designing horizontal curves."

Using CQ, concepts in HiOnto are categorized along three dimensions: context, intent, and extent (Davey and Priestley 1990). The context dimension divides the taxonomy to major subdomains (traffic engineering versus highway planning, for example). Within each context, the intent dimension classifies concepts according to their attributes—for example, dividing highway engineering into highway design and highway construction. Finally, the extent dimension examines the length of subtrees (how many concepts can be children of a super class?)

The final version of HiOnto includes a taxonomy of approximately 2,800 unique concepts (most of them have additional keywords/synonym terms). It is not feasible to list all products here. Figs. 2–7 represent an overview of some of the major domains, which are explained below.

Projects

Several application level projects were established, for example: bridge project (includes new bridge project, bridge maintenance project, bridge rehabilitation project, and bridge demolition project) and highway project (includes new highway project, highway maintenance project, highway rehabilitation project, and highway demolition project). Several related concepts (such as tunnel project), although not relevant to HiOnto scope, were defined. Their exact attributes have to be further developed in future research.

Processes

At the domain level, this research project added the following processes to the *e-COGNOS* ontology: analysis process, and testing process to be the parent processes for similar processes in the application level. In the application ontology level, 281 processes were identified. These include, for example: highway project management process, highway construction process, highway design process, and highway estimation process. Further down the taxonomy, the highway design process includes: geometric design process, pavement design process at the preliminary, and details levels.

Products

Three hundred eighty four highway related products were identified at the application level. HiOnto defines both physical and managerial products. Fig. 2 shows an overview of the major classes in this domain. Highway construction complex describe major highway physical product (such as, roads, bridges, and tunnels). Fig. 3 shows the basic sub elements of the bridge product. Highway basic products used in developing these complex ones are also described (such as urban street furniture elements). Construction aids used to build these products are also defined (such as caisson, scaffolds, and dewatering systems). Finally, manage-

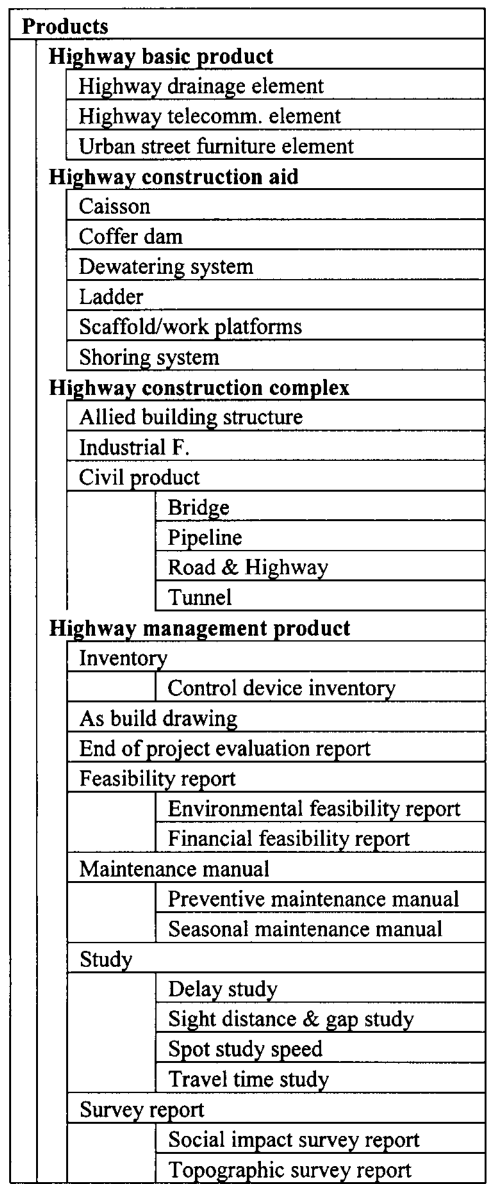


Fig. 2. Highway products (partial list)

J. Constr. Eng. Manage., 2005, 131(5): 591-603

Bridge	
	Arch bridge
	Canal bridge
	Cantilever bridge
	Covered bridge
	Curved bridge
	Floating bridge
	Girder bridge
	Lattice bridge
	Opening bridge
	Pedestrian overpass bridge
	Pipeline bridge
	Portal frame bridge
	Railway bridge
	Skew bridge
	Stayed girder bridge
	Strutted bridge
	Suspension bridge
	Temporary bridge
	Transporter bridge
	Twin bridge
	Vaulted bridge

Fig. 3. Subclasses of bridge product

ment products of related managerial processes are also defined in this domain (for example, budget, schedule, safety report).

Actors

A new major subdomain was added to the domain level (*e-COGNOS* ontology): other actors. *e-COGNOS* includes two major actors: organizations and personnel. This subdomain will help in defining any future abstract/theoretical construct—for example, simulation objects, and a driver-vehicle unit, which are major actors in any traffic analysis process. In total, the ontology includes a total of 117 highway-specific actors (under organizations and personnel).

Resources

Four hundred forty one application level resources were identified. Most of the additions were made in the equipment and materials subdomains. The self-explanatory Fig. 4 shows an overall view of this domain, which includes an extensive list of materials and equipment used in highway engineering. Fig. 5 shows the details of a sample sub domain: earth moving equipment.

Technical Topics

This domain defines the technical aspects (attributes, parameters, features, constraints, supporting concepts, mechanisms) of entity domains. They cover domains such as different transportation modes (ground, air, ship), transportation services (regular and special). It also describes the parameters used in transportation planning (access time, spot speed, and wait time). It also defines the tests used in highway construction and the applicable codes. Figs. 6 and 7 show samples of technical topics domains.

Resources		
	Highway equipment	
		Accessories equipment
		Bending equipment
		Compaction equipment
		Concreting equipment
		Cutting equipment
		Drilling equipment
		Earth moving equipment
		Hauling equipment
		Lifting equipment
		Painting equipment
		Piling equipment
		Pumping equipment
		Retaining device
		Scaffold equipment
		Spraying equipment
		Storage equipment
	Highway material	
		Bituminous materials
		Cementitious materials
		Ceramic materials
		Coating and paints
		Glass
		Lumber and wood materials
		Masonry materials
		Metallic materials
		Polymeric materials
		Pozzolonic materials
		Pulp and paper materials
		Rocks and soils
	Personnel	
		Laborers
		Professional

Fig. 4. Resources (partial list)

HiOnto Relationships and Axioms

Similar to object oriented methods, in developing ontologies, concepts are defined by nouns and relationships are identified by verbs (Gomez-Perez 1998). For example, an entity, say civil engineer, may have several attributes, e.g., name, role, profession, etc. Consequently the relations: has-name, has-role, has-profession, etc., could be established to link the entity to the concepts corresponding to the attributes. In addition, axioms (or rules) are used to constrain the development of concepts, their interrelationship, and their updates. HiOnto inherited the axioms of *e-COGNOS* (Lima et al. 2003).

To represent the application-level relations and axioms, the research team developed an object model for the highway domain that presented our basic view of the domain. Fig. 8 shows a partial view of this model, which is process-centered. Ground transportation engineering includes: Railroad engineering, pipeline system engineering, waterway system engineering, and highway engineering. Highway engineering includes: highway design, highway construction, highway maintenance, highway operations, highway rehabilitation, and highway traffic engineering. Highway traffic engineering includes traffic design, traffic operation, and intelligent transportation systems (ITS). Highway operations influence the need for highway maintenance. Traffic engineering

- **Earth moving equipment**
  - Breaker
  - Bucket excavator
  - Bulldozer
  - Ditcher
  - Dragline
  - Excavation equipment
    - Digger
    - Trencher
  - Grab bucket
  - Grader
  - Loader
  - Paving machine
  - Ripper
  - Scraper
  - Snow melting machine
  - Spreader

Fig. 5. Earth moving equipment

- **Traffic analysis parameters**
  - 30th highest hourly volume
  - Access time
  - A lot time
  - Average speed
  - Check in time
  - Designated speed
  - Down time
  - Driving skills
  - Dwell time
  - Excess time
  - Hourly volume
  - Ideal capacity
  - Intersection capacity
  - Lay over time
  - Monthly volume
  - Physical characteristics
  - Possible capacity
  - Practical capacity
  - Running speed
  - Sight distance
  - Spot speed
  - Stopped time
  - Transfer time
  - Uninterrupted capacity
  - Volume forecast
  - Volume measuring device factor
  - Wait time

Fig. 6. Sample technical topic

- **Soil tests**
  - Soil field tests
    - Field compaction test
    - Field moisture content test
    - Field permeability test
    - Plate bearing test
    - Soil consolidation test
  - Soil laboratory tests
    - California bearing ratio test
    - Laboratory permeability test
    - Moisture content test
    - Particle size distribution test
    - Soil atterberg limit test
      - Liquidity limit test
      - Plastic limit test
    - Soil chemical test
    - Soil drawdown test
    - Soil proctor test
    - Soil sedimentation test
    - Soil shrinkage test
    - Soil viscosity test
    - Tri-axial test
    - Vane shear test

Fig. 7. Sample technical topic

defines transit modes, which is part of the passenger transportation mode. Ground transportation modes provide transportation services to the community. These services could be public or private and could be regular or special services. Transportation planning influences highway design. Traffic engineering data influences transportation planning, which has impacts on the national ground transportation system. A set of axioms have been also developed. The following sample process illustrates the use of the three levels of HiOnto, its relationships, and axioms.

Testing Process

Testing processes are some of the most common processes in any construction system. Fig. 3 shows the object model for the testing process as it spans the three levels. The major relations/axioms in this model include:

1. Each highway testing process (H_{Tpr}) has a highway test object (H_{To}) that is either a highway product or a highway resource:

$$\forall H_{\mathrm{Tpr}} \exists H_{\mathrm{To}} \in (\text{highway\_product} \vee \text{highway\_resource}) \supset \text{has}(H_{\mathrm{Tpr}}, H_{\mathrm{To}})$$

2. Each highway testing process (H_{Tpr}) has a highway testing method (H_{Tm}) that is either a destructive or nondestructive method:

$$\begin{aligned}\forall H_{\mathrm{Tpr}} \exists H_{\mathrm{Tm}} \in &(\text{destructive\_testing\_method} \\ &\vee \text{nondestructive\_testing\_method}) \\ &\supset \text{has}(H_{\mathrm{Tpr}}, H_{\mathrm{Tm}})(\forall H_{\mathrm{Tpr}} \in \text{testing\_process}) \\ &\times (\exists H_{\mathrm{Tm}} \in \text{testing-method} \\ &= \{\text{destructive testing method}, \\ &\text{nondestructive\_testing\_method}\}) \\ &\text{has}\ (H_{\mathrm{Tpr}}, H_{\mathrm{Tm}})\end{aligned}$$

J. Constr. Eng. Manage., 2005, 131(5): 591-603

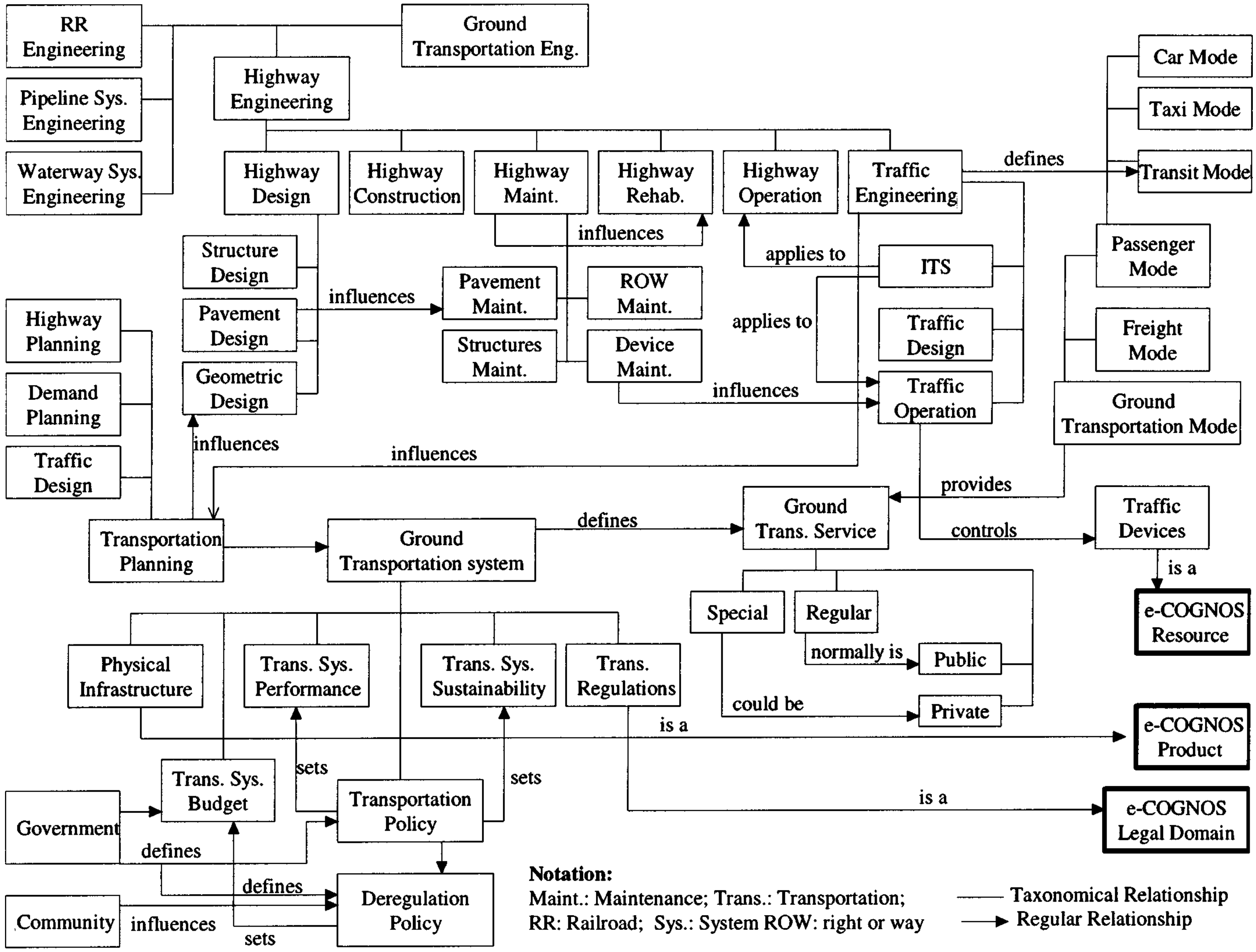


Fig. 8. HiOnto relationships: abstract object model

3. Each highway testing process has to conform to a testing code

$$\forall H_{\mathrm{Tpr}} \exists T_c \supset \mathrm{conforms}(H_{\mathrm{Tpr}}, T_c)$$

4. Each highway testing process has a testing safety standard (S_s) (that constrain its procedures): $\forall H_{\mathrm{Tpr}} \exists S_s \supset \mathrm{has}(H_{\mathrm{Tpr}}, S_s)$
5. Test results may be approved, not_approved, under_development, or under_evaluation.
6. Each Test Result (T_r) has to be approved. An approved test results (A_{Tr}) entitles that the results be in conformance with Test Code (C_t) and test specification (S_t) and be certified by the authorized actor (A_{AC})

$$\forall H_{\mathrm{Tpr}} \exists A_{\mathrm{Tr}} \Leftrightarrow \mathrm{conforms}(T_r, C_t) \wedge \mathrm{conforms}(T_r, S_t)$$
$$\wedge\, \mathrm{certified}(T_r, A_{\mathrm{AC}})$$

Fig. 9 shows how all the entities in the application level are inherited from the domain level. It also shows how all the elements in the application level can be linked to a specific instant in the enterprise model (for example, field quality engineer supervises soil tests).

Validation

The issue of validating ontologies is not easy and has been addressed by leading aritificial intelligence (AI) experts in different ways (please review: Bench-Capon 1990; Schreiber 1992; Gruber 1993; Visser 1995; Uschold and Jasper 1999; Chandrasekaran and Josephson 1997). In fact, by their very nature, the creation of ontologies (as a form of formal models) challenges our conceptualization of the subject domain; subsequently, they lend themselves to further change (Boyd 1976). After reviewing these and other research in AI and ontology evaluation, Visser and Bench-Capon (1997) developed a comprehensive analysis of the issue of ontology effectiveness and representation. They concluded that the following criteria should be considered during ontology evaluation:

1. Epistemological adequacy: The degree to which the ontology resembles the cognitive sentence: (1) Epistemological clarity: Do all concepts and relations in the ontology have a clear and unequivocal meaning? (2) Epistemological intuitiveness: Do the ontological concepts and relations provide a vocabulary that matches the intuition of the experts in the domain?

J. Constr. Eng. Manage., 2005, 131(5): 591-603

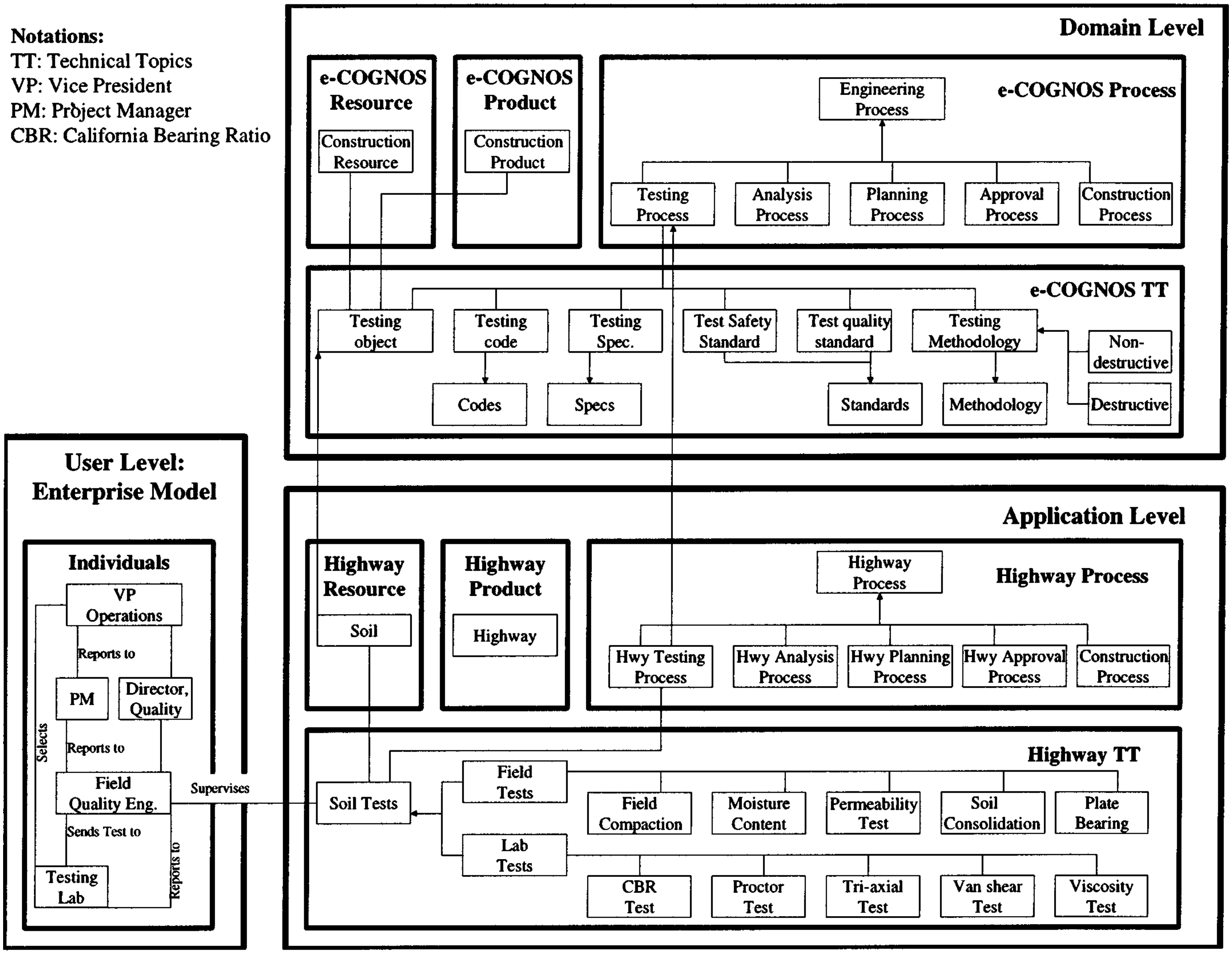


Fig. 9. Three level presentation of testing process

(3) Epistemological relevance: Are all the methods, concepts and relations in the ontology relevant for the domain? (4) Epistemological completeness: Does the ontology cover all relevant concepts and relations that may be relevant for any combination of task, method and subdomain? Are there any entities that cannot be modeled in the ontology? (5) Discriminative power: Does the ontology have enough discriminative power in that it provides distinctions at a sufficiently high granularity level (*viz.* sufficient detail)?

2. Operationality: Effort required to implement ontological concepts and relations in a representational language: (1) Encoding bias: Does the ontology rely on symbol-level choices? An ontology should be specified at knowledge level, an encoding bias results when a representation choice is made purely for the convenience of notation or implementation. (2) Coherence: Is the ontology coherently defined in that it is internally consistent? (3) Computationality: Does the ontology provide a basis for (computational) representation, and is this representation computationally adequate?
3. Reusability: The degree in which the ontology can be reused to conceptualize new concepts: (1) Task-and-method reusability: Is the ontology dependent on certain (types of) tasks and methods? (2) Domain reusability: Is the ontology dependent on certain (types of) subdomains?

Table 2. Validation Strategy

	Validation issue/Concern number									
Validation tool	1.a	1.b	1.c	1.d	1.e	2.a	2.b	2.c	3.a	3.b
Competency questions	X	—	X	—	—	—	X	—	—	X
Separating design & implementation	—	—	—	—	—	X	—	X	—	—
Professors involvement	X	X	X	X	X	—	—	—	X	X
Industry survey	X	X	X	X	X	—	—	—	X	X

J. Constr. Eng. Manage., 2005, 131(5): 591-603

In benchmarking the best practice of this and other domains with reasonable history in ontology development and use (for example, biomedical and automotive), our attempts to address these issues started from the design of research methodology, to requirements analysis, to ontology development and implementation, and validation. The following steps were taken (see also Table 2):

1. Use of competency questions: as shown in Table 1, these are a set of questions that had to be answered before any concept was to be added to the ontology including under what title, what kind of domain-neutral definition should be used to define the concepts, how to drive domain-specific types of this concepts and where to classify them. This was important to satisfy Criteria 1.a, 1.c, 2.b, and 3.b.
2. Separating design from implementation: the ontology was developed using plain English terms (using *Excel*) to make sure that it can be used without a specific ontology language. Each major concept (like PR, PD) was also represented in *UML* format. Then the ontology was implemented using web ontology language (OWL). The use of *OWL* was particularly important as it will allow this ontology to link to other upper level ontologies such as IEEE's Cyc. This decision helped with Criteria 2.a.
3. Involvement of experts throughout the development: an ontology is basically a formal model for a domain theory (academic view) and its use (by industry). To assure theoretical consistency, during the different phases of ontology development, three professors (in the areas of highway and transportation planning and design) were involved in evaluating the work through workshops. This was crucial for the coherence of the ontology (Criteria 2.b and 1.b) and also for Criteria 3.a and 3.b.
4. Expert survey: The final validation was achieved through interviews with domain experts. Purposive sampling technique was used to choose the experts. In this technique, a set of criteria is used to select the experts. In our case, these included: years of experience in the field; and the diversity of expertise (i.e., knowledge of design and construction). The survey was designed to address the following concerns: (1) Navigation/traversing: It has been found that ontologies, with their sheer size, may cause problems related to structuring, retrieval, and maintenance (Klinker et al. 1990). (2) Abstraction consensus: It is very hard, within such a wide spectrum of concept, to reach consensus among users about the core concepts that could model a domain knowledge. Traditionally, this has been addressed through iterative development of ontologies and the engagement of a set of domain experts (3–10) for in-depth analysis, in contrast to mass surveys (the interested reader is referred to the wealth of research into ontology development and validation. For example, Gruber 1993). (3) Representation: HiOnto represents transportation engineering in general and highway construction in particular. It is important to test if HiOnto contains a sufficient number of concepts so that it could adequately represent its domain.

At the final stage of ontology validation, 13 industry experts were interviewed. The chief reason for selecting one-on-one interviews was the fact that ontology, as a knowledge management tool, is relatively new concept in construction. On average each interview took 2 h 40 min to complete. The experts were briefed for 30–45 min about the sources of gathering different concepts and how they are structured to form hierarchies. The survey included 12 questions that were designed to serve the following objectives (a six-point scale was used to record experts' response, with 1 being the most favorable in each case):

1. Navigational ease through locating concepts: four questions were included in this section. First, experts were asked to find about 20 concepts in the ontology. This helped in assuring Criteria 1.a, 3.a, and 3.b. Having to find concepts on their own, experts became more aware of the classification scheme adopted in HiOnto and asked probing questions. Table 3 (in the Appendix) shows the results of the main question in this section. The average response to this question was 2.2.
2. Categorizing concepts: as experts became more aware of the domains, we asked them to categorize 15 additional concepts that were not in the taxonomy (these were intentionally left out). This was meant to test Criteria 1.e, 3.a, 3.b. Tables 4 and 5 in the Appendix show the results of the main question in this section. The average response to this question was 2.6.
3. Overall assessment: finally, as the experts became fully aware of the ontology and the conflicting needs of categorization, we asked them to make a general assessment about the taxonomy. They were asked specifically if they thought that the ontology was biased to one domain, if they agreed with the classification system adopted in HiOnto and its six major classes. Table 6 in the Appendix shows the results of the main question in this section. The average response to this question was 2.1.

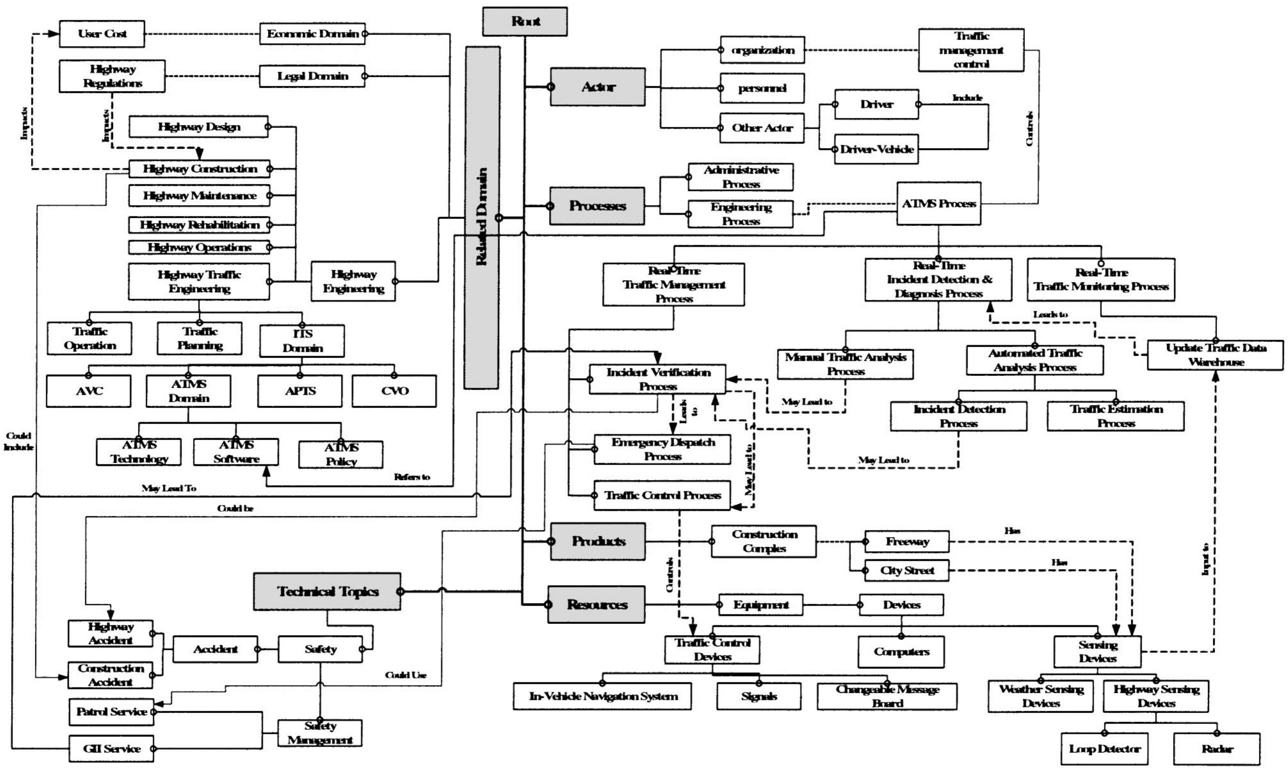
Deployment and Testing

One of the most important tools that helped evaluate the validation of the proposed ontology was actual deployment and testing by construction companies. A formal evaluation was conducted at four leading European contractors as part of the final evaluation of the *e-COGNOS* project. First a survey (similar to the aforementioned survey) was conducted with four leading IT experts and end users of the system. This survey used a scale of 1–6, where 1 is very hard (worst) and 6 is very easy (best). In order not to confuse the reader, the expert ratings were transferred to the previous scale, where 1 is best and 6 is worst. Their assessment for the most part was positive. For example:

1. How easy was it to find concepts (a list of 28) in the Ontology? Average rating = 2.5.
2. How effective was the categorization of the concepts? Average rating = 1.5.
3. How easy was it to categorize concepts (a list of 24) in the Ontology? Average rating = 2.9.
4. How easy was it to navigate through the Ontology? Average rating = 1.
5. How familiar were the terms used in the Ontology? Average rating = 2.
6. How representative are the terms used? Average rating = 3.
7. Overall, did the ontology cover the main domains of Construction sector? Average rating = 2.

These results were seen as very positive because the experts deal with all elements of construction systems (not limited to highway engineering). It is well recognized that human–software interaction enhances over time. Given that this was the first time end users used the system, it is expected that they, in future use, would have more familiarity with the terms and find the ontology easier to use. Following that, each organization provided a formal evaluation of the use of the ontology (and *e-COGNOS* tools) in the following applications:

J. Constr. Eng. Manage., 2005, 131(5): 591-603



Notation:
APTS: Advanced Public Transit Systems
ATMS: Automated Traffic Management Systems
AVC: Automatic Vehicle Classification
CVO: Commercial Vehicle Operator
GII: Global Information Infrastructure

Fig. 10. Automated traffic management systems domain (partial view)

1. Knowledge extraction: The prime need here being the extraction of content and metadata from existing documents in different formats.
2. Knowledge classification: Including categorization and clustering (finding similarities between knowledge items).
3. Search (manner and knowledge items): The search, discovery, and ranking of knowledge items with respect to both the manner in which these are done and in terms of the different types of knowledge items considered.
4. Knowledge sharing and assessment: Publishing of valuable documents (making one's solutions available to others).
5. Knowledge dissemination: Knowledge dissemination in terms of company news, project news, and specialized knowledge available through distribution channels on subscription or user profiles.
6. Creation and synthesis of new solutions: The creation and synthesis of new solutions from existing ones (or related information).
7. Knowledge maintenance: Crawling of dedicated knowledge sources and the watching of changes on them, and the automatic notification of change to impacted knowledge items.
8. User profiling: Personalized user interfaces and shared virtual workspaces being of medium priority, and private configurable workspaces of low priority.
9. System administration: Setup of access and rights control.

Overall, ontology based tools were received very positively by end users and the system evaluators at each organization. The exact evaluation of *e-COGNOS* performance against these criteria varied from one organization to another based on the available tools in each of these organizations. The interested reader is referred to (*e-COGNOS evaluation results* 2003) for detailed explanation of the assessment of each of these criteria at each organization.

Corroborating all of this is the results of the deployment of the ontology and its applications. At the time of this paper, the ontology (and other *e-COGNOS* tools) was fully deployed and implemented by a leading United Kingdom contracting organization. They, eventually, compared the functionality of *e-COGNOS* and two other knowledge management tools that are being used at their branches. In many cases *e-COGNOS* was seen as superior (mainly due to its use of ontology), in others it had mixed reviews due to the unfamiliarity of users with its interface. The main conclusion of this deployment is that, despite the lack of user friendly interfaces in *e-COGNOS* tools, the ontology provides a very promising tool, based on which more effective knowledge management systems can be built (please see *e-COGNOS evaluation results* 2003 for an interesting comparison of the functionality and user friendliness of *e-COGNOS* and other knowledge management/search tools).

It is recognized that, as a theoretical model of a very complex domain, ontology evolves with use. Further research needs to be conducted to expand the ontology, link it to existing systems and address end users requirements for easier interface (*e-COGNOS evaluation results* 2003).

Exchange of Infrastructure Information

This section provides an outlook into the future usability of HiOnto beyond highway construction. The main role of ontolo-

J. Constr. Eng. Manage., 2005, 131(5): 591-603

gies in this regard is to bridge the interoperability gap between different domains. Based on the proposed HiOnto, several other infrastructure ontologies will be developed to model telecommunication, gas, and electric utilities. A mapping protocol will be developed to support the exchange of data between these ontologies, based on the fact that they share the same geometry of highways. Furthermore, these ontologies will be connected to IFC-based computer aided design (CAD) systems and geographic markup language (*GML*)/*ISO*-Technical Committee 211 (*TC211*)-based geographic information system (GIS) systems. Knowledge (not merely geographical data) about physical products (highway, pavement, gas line, bridge, vehicle) will be exchanged in a seamless way between CAD and GIS systems. This will allow a user to define the location of infrastructure elements through the GIS interface, to be able to explore their engineering details through the CAD system, and to access any technical/business attributes of such elements through their ontological representation.

To illustrate the potential for cross discipline integration through HiOnto, an application ontology for automated traffic management systems (ATMS) was developed. Along with highway construction, ATMS is part of the highway engineering domain. With real-time data management becoming a major factor in controlling construction activities and enhancing safety, it is important to link the two domains. The ATMS data are normally presented through GIS, video feeds, and software systems.

Three major processes are paramount in ATMS: real-time traffic monitoring, real-time incident detection and diagnosis, and real-time traffic control. Each process may have additional subprocesses. Three major actors are identified: traffic management center, driver, and driver–vehicle unit (see Fig. 10).

Three major products are included in the ATMS ontology: freeway, highway sensing devices, and traffic control devices. On a GIS file, the freeway will be presented as a graphical object. On a CAD file, a segment of that freeway will be shown with all engineering details, including traffic control devices and highway sensing devices. All these graphical objects will have reference to other relevant logical objects like process, actor, and software. A traffic engineer could specify the location of a changeable message board on a GIS file, based on the overall scheme of ATMS. This will automatically show on the CAD file of this portion of the freeway on the construction engineer's computer. If the number or specifications of this board changes, a cost object will keep track of these changes. If the construction sequence of the installation of this board is changed in a way that impacts project cost, another object will take account of that. During construction, if a construction accident takes place, some one could call emergency service (911 Service in North America), which will trigger the ATMS including the alteration of traffic control devices and emergency dispatch.

Conclusions

This paper presented a distributed ontology architecture for highway construction (HiOnto). The architecture is built as an extension to the *e-COGNOS* ontology. The architecture classifies highway concepts into processes, projects, products, actors, resources, and technical topics (boundary conditions and technical details). The ontology was developed using *Web Ontology Language* OWL, which is designed as the main tool for enabling semantic web. This will allow easier link to other domain ontologies.

This is the first ontology that aims at covering the whole highway construction domain. The ontology was developed in a distributed way to allow for future expansions and customization of its terms and relationships to suit proprietary data structure systems of each individual organization. In this regard, the proposed architecture allows each organization to link its enterprise model directly to the concepts of the ontology.

In addition to proposing a semantic knowledge representation model (the ontology), the contribution of this research includes applications of such semantic systems in knowledge management. The research presented a system for integrating traffic management with highway construction as a means to illustrate the possible impacts of semantic knowledge representation. With highways being the backbone of all other infrastructure, the proposed architecture could act as the basis of an infrastructure ontology.

Like any theoretical modeling initiative, the proposed architecture is not the only way to model highway engineering and construction. In fact, validating an ontology is not an easy endeavor—as they normally attempt to encapsulate tacit knowledge (in the form of taxonomies, relationships and axioms) among different stakeholders with, normally, varying backgrounds, interests, and experiences. To this end, we attempted to address this by focusing on developing a flexible architecture that lends itself to modifications and future additions; using competency questions to assure consistent knowledge representation and concept classification; and engaging experts at all levels of development.

Summary

The paper presented a distributed ontology architecture for managing highway construction knowledge in a semantic way. Six major concepts were used to classify construction knowledge: project, process, products, actor, resource, and technical topics (boundary conditions). The architecture also allow users to link a specific enterprise model to the proposed architecture.

Appendix. Summary of Survey Guide

Section 1: Ease of Navigation (Four Questions)

Q 1: Please navigate through the ontology and find the following concepts (Table 3) then rank the ease (or difficulty) of finding them on a scale of 1–6.

Section 2: Classification/Abstraction Efficiency (Four Questions)

Q 5: The following concepts (Table 4) were classified under the shown super class, do you agree with that? Please rank your answer on a scale from 1 to 6.

Q 6: The following concepts (Table 5) **are not** included in HiOnto, please:

1. Suggest a classification for them (under which super class should they be classified).
2. How easy was it to classify them? Please rank your answer on a scale from 1 to 6.

J. Constr. Eng. Manage., 2005, 131(5): 591-603

Table 3. Navigation Ease

Concepts	Average
Ministry of transportation	1.25
Bank	1.40
Project change management process	1.92
Mobilization process	3.20
Horizontal curve design process	1.73
Igneous rock	1.58
Total quality management	2.09
Electronic message board	3.45
Slurry seal coat	2.90
Transportation system performance	2.29
Highway bridge design	1.22
Dial-a-bus service	2.25
Subway transit mode	1.86
Public art/sculpture	3.27
Traffic signal	2.13

Table 4. Classification Efficiency

Super class	Concept	Average
Actor	Research institutes	2.17
	Municipal Government	1.17
	Highway maintenance manual	2.00
Product	Toll plaza/toll station	2.33
	As built drawing	2.08
Resources	Pile driver	1.50
Process	Soil stabilization process	4.50
	Traffic control design	1.50
	Electronic message board	3.33
Technical topics	Dial-a-bus service	2.14
	Highway devices	2.10
	Exclusive bus lane	1.89
	School transport service	2.00
	Highway rehabilitations	1.30
	Transportation system performance	1.8

Table 5. Question Number 6—Categorization Efficiency

Concepts	Average
Actual cost of work performed	2.58
Critical activity	2.21
Soil movement	3.08
Household	3.50
Trip matrix	3.27
Traffic corridor	4.00
Pothole	3.46
Decision making	2.64
Highway (operational) reliability	3.40
Automatic train control	3.29
Transportation master plan	2.27
Three block system	3.00
Traffic congestion	3.70
Traffic pollution	3.45
Genetic algorithm	3.80

Table 6. Overall Assessment

	Rating scale from best to worst					
Questions	1	2	3	4	5	6
How easy was it to navigate through the taxonomy?	—	X	—	—	—	—
How familiar were the concepts used?	X	—	—	—	—	—
How representative are the concepts used?	—	X	—	—	—	—
Overall, did the taxonomy cover the main domain of highway engineering. Was it biased to a certain subdomain?	—	X	—	—	—	—

Section 3: Overall Assessment (Four Questions)

Based on your navigation, please answer the following (Table 6). Rate your comments on a 6 point scale (1–6) where 1 is the best and 6 is the worst.

Notation

The following symbols are used in this paper:

A^o = set of ontology axioms, represented in first order logic;
A_{AC} = authorized actor;
A_{Tr} = approved test results;
AC = actor;
ATMS = automated traffic management systems;
C = disjoint set whose elements are called concepts: $X_1, X_2, X_3, \ldots. X_n$;
C_t = test code;
D_c = design code;
D_o = design objective;
D_{pr} = design process;
EN = entity;
EN_e = emergency entity;
EN_l = planned entity;
EN_o = change order entity;
EN_r = rework entity;
H^c = class hierarchy representing directed relation such that $H^c \subseteq C X C$;
H_{Tm} = highway testing method ( );
H_{To} = highway test object ( )P=process;
H_{Tpr} = highway testing process ( );
O = ontology structure;
PD = product;
PJ = project;
PR = process;
R = nontaxonomic relation between concepts: $R \rightarrow C X C$;
R^t = disjoint set whose elements are called taxonomical relations: $Y_1, Y_2, Y_3 \ldots. Y_n$;
RE = resource;
S_s = testing safety standard;

J. Constr. Eng. Manage., 2005, 131(5): 591-603

S_t = test specification;
T_c = testing code;
TT = technical topics;
t = time; and
σ = Stat, stage, or situation.

References

Atkin, B., Flanagan, R., Marsh, A., and Agapiou, A. (1995). *Improving value for money in construction: Guidance for chartered surveyors and their clients*, The Royal Institution of Chartered Surveyors, London.

Bench-Capon, T. J. M. (1990). *Knowledge representation: An approach to artificial intelligence*, APIC Series, No. 32, Academic, London.

Berners-Lee, T. (2001). *Weaving the web*, Harpur, San Francisco.

Boyd, J. (1976). "Destruction and creation." ⟨http://www.belisarius.com⟩ (August 1, 2004).

Chandrasekaran, B., and Josephson, J. R. (1997). "The ontology of tasks and methods." *Working Notes of the AAAI Spring Symp. on Ontological Engineering*, Stanford Univ., Stanford, Calif.

Davey, B. A., and Priestley, H. A. (1990). *Introduction to lattices and order*, Cambridge Univ. Press, Cambridge, U.K.

e-COGNOS evaluation results. (2003). "e-COGNOS consortium." ⟨www.e-cognos.org⟩ (August 3, 2004).

Gomez-Pereze, A. (1998). "Knowledge sharing and re-use." *Handbook of applied expert systems*, J. Liebowitz, ed., CRC, Boston.

Gruber, T. (1993). "A translation approach to portable ontology specifications." *Knowledge System Laboratory, Technical Rep. No. KSL 92-71*, Computer Science Dept., Stanford Univ., Stanford, Calif.

Gruninger, M., and Fox, M. S. (1996). "The logic of enterprise modeling." *Modeling and methodologies for enterprise integration*, P. Bernus and L. Nemes, eds., Chapman and Hall, Cornwall, U.K.

Guarino, N. (1997). "Understanding, building, and using ontologies—A commentary to Using explicit ontologies in KBS development by van Heijst, Schreiber, and Wielinga." *Int. J. Comput.-Human Studies*, 46(2–3), 293–310.

Guarino, N., and Welty, C. (2000). "Towards a methodology for ontology based model engineering." *Proc., IWME-2000-International Workshop on Model Engineering*, Sophia Antipolis and Cannes, France, June 12–16.

Katanuschkov, P., Gehre, A., and Scherer, R. J. (2002). "An engineering Ontology framework as advanced user gateway to IFC model data." *Proc., European Conf. on Product and Process Modeling*, Prortoroz, Slovenia.

Klinker, G., Bhola, C., Dallemagne, G., Marques, D., and Mcdermott, J. (1990). "Usable and reusable programming constructs." *Proc., 5th AAAI Knowledge Acquisition for Knowledge Based Systems Workshop*, Banff, Canada, SRDG, Univ. of Calgary, Calgary, Canada.

Lima, C., Fies, B., El-Diraby, T. A., and Lefrancois, G. (2003). "The challenge of using a domain Ontology in KM solutions: The e-COGNOS experience." *Proc., 10th ISPE Int. Conf. on Concurrent Engineering: Research and Applications*, Madeira Island, Portugal.

Maedche, A. (2002). *Ontology learning for the semantic web*, Kluwer Academic, New York.

Maedche, A., Motik, B., Stojanovic, L., Studer, R., and Volz, R. (2003a). "An infrastructure for searching, reusing and evolving distributed ontologies." ⟨http://kaon.semanticweb.org.⟩ (June 12, 2003).

Maedche, A., Motik, B., Stojanovic, L., Studer, R., and Volz, R. (2003b). "Ontologies for enterprise knowledge management." *IEEE Intell. Syst.*, 18(1), 26–33.

Schreiber, G. (1992). "Pragmatics of the knowledge level." PhD thesis, Univ. of Amsterdam, Amsterdam, The Netherlands.

Uschold, M., and Jasper, R. (1999). "A framework for understanding and classifying ontology applications." *Proc., IJCAI-99 Workshop on Ontologies and Problem-Solving Methods (KRRS)*, Stockholm, Sweden.

Visser, P. R. S. (1995). *Knowledge specification for multiple legal tasks: A case study of the interaction problem in the legal domain*, Computer/Law Series, No. 17, Kluwer Law International, The Hague, The Netherlands.

Visser, P. R. S., and Bench-Capon, T. J. M. (1997). "A comparison of four ontologies for the design of legal knowledge systems." *Artificial Intelligence Law*, Springer Science, 6(1), 27–57.

J. Constr. Eng. Manage., 2005, 131(5): 591-603