DEPARTMENT OF CIVIL ENGINEERING
UNIVERSITY OF TORONTO

A Knowledge-Enabled System for Coordinating the Design of Co-Located Urban Infrastructure

December 2004

A Proposal By:
Hesham Osman

Supervisor:
Prof. Tamer El-Diraby

1 EXECUTIVE SUMMARY

The following document is intended to portray a snapshot of the research currently underway to create a knowledge-enabled semantic system for coordinating the design of co-located urban infrastructure. Work was commenced in May 2004 and extensive literature reviews, preliminary case study analysis, requirements analysis, and a proposed preliminary ontology were performed to date. As such, this document presents excerpts of the work accomplished so far and proposes a detailed methodology, scope, and schedule for the remaining portions of the research.

The main focus of the research is on the representation of knowledge (both tacit and explicit) used in the domain of urban infrastructure design, in a formal, semantic and reusable way. The research proposes the use of *Ontologies* – a formal explicit model for a domain – to be the main enabler in this endeavor. The knowledge available in urban infrastructure is immense, thus to focus the scope of our knowledge representation, the process of routing/alignment of buried co-located buried urban infrastructure is highlighted. Immense potential exists for reusing the tacit knowledge of domain experts as well as learning from street evolution patterns by analyzing the efficiency and sustainability of using underground space in congested urban areas.

The document contains 7 main sections and 3 appendices. The main sections outline the research motivation, problem statement, objectives, scope, methodology, contributions and management. The appendices contain excerpts from the literature review and some of the work accomplished so far.

2 INTRODUCTION & MOTIVATION

Our urban infrastructure is decaying at a rate that far exceeds our ability to replace. Larger proportions of infrastructure budgets are being allocated to infrastructure renewal than to the development of new infrastructure. Infrastructure rehabilitation projects – especially in old urban areas – are becoming essential for sustaining their vitality. With the increasing number of these projects taking place, designs should become more effective and be conducted in a more coordinated fashion. Traditionally, buried infrastructure was located with very little regard to important aspects like the need for maintenance, the need for replacement, disruption to community, etc… These needs are very difficult to perceive in an insular design environment that only takes into consideration the current requirements of the system being designed. Novick (1990) identified seven key issues that should be studied to find long-range solutions to the problems facing urban infrastructure. One of these issues involves "*The preparation of new design approaches that focus on ease of inspection, repair, and reconstruction in place, all under traffic*".

The main driving impetuses behind this research are:

1- The large number of infrastructure projects being carried out in our cities. These projects involve both the construction of new systems and the rehabilitation or replacement of old systems.
2- The large number of stakeholders involved in urban infrastructure projects. In comparison to building construction projects, infrastructure projects will usually involve a much larger number of stakeholders. This can be attributed to:
    a. The deregulation of infrastructure in the early 90's has led to utilities that are managed by a multitude of agencies.
    b. Local communities and businesses are directly affected during and after construction. They play a key role in project decision making.
    c. Private sector investment in infrastructure has placed financial institutions as a key stakeholder in the infrastructure development process.

3- The potential benefits of feeding back the knowledge gained during the construction and operation of co-located urban infrastructure to guide the development of new projects in a more enhanced manner.
4- Municipal codes governing buried infrastructure that solely address technical, engineering, environmental, and safety requirements need to be augmented with guidelines that address issues of the sustainable and optimized use of underground space. Co-located urban infrastructure should be located taking into consideration the requirements, constraints and aspirations of the future.

3 PROBLEM STATEMENT

There is a need to embed knowledge from the lessons we have learned from the past to enhance the process of designing co-located urban infrastructure in order to improve the design, construction and operation of urban streets and the congested mass of utilities buried underneath. This enhancement should take into account sustainability issues like life cycle costs and long term impacts on the environment, businesses and surrounding communities.

Thus, there exists a great need to formalize new design guidelines that consider both the codependent nature and the sustainability concerns that underlie the design of colocated urban infrastructure. If we design utilities that are too finely tuned to present needs alone, they will soon become outdated. By considering urban underground space as a scarce resource, the process of routing buried urban utilities should focus not only on meeting the requirements of today, but should allow future generations to easily maintain, expand and improve our infrastructure systems.

4 OBJECTIVES

4.1 Develop an ontology to encapsulate the knowledge available in the field of urban infrastructure design

Ontologies aim at developing a shared model of a domain by representing both tacit and explicit knowledge. Ontologies incorporate all pertinent concepts relevant to the domain of interest as well as the relationships and axioms/rules that govern the concepts. Most modeling efforts in the civil/architectural domain were focused on building product models, with special attention given to representing architectural components of buildings (Eastman, 1999). The need to extend modeling efforts in the civil engineering realm to include civil infrastructure has been identified as one of the most needed upcoming steps in model-based exchange (Froese, 2003). A rigorous review of the literature revealed very scarce initiatives to model concepts related to civil infrastructure.

The ontology will collect/organize the main concepts/relationships pertaining to the field of urban infrastructure design. The ontology will encapsulate the application-specific knowledge pertaining to the design process of routing buried infrastructure. As such the ontology can be classified to lie between a domain and application ontology. In the broader sense, the ontology is specific to the domain of "Urban Infrastructure Design" but cannot be completely classified as a domain-specific ontology as it is intended to be used for encapsulating the tacit knowledge involved in routing buried infrastructure. On the other hand it cannot be completely classified as an application-specific ontology, as its use is not intended for a particular organization.

4.2 Propose a decision-making model for the sustainable routing of co-located buried urban infrastructure

The decision making model should be cognizant of the following issues:

<u>1- Sustainability issues:</u> By considering our urban underground space as a scarce asset that needs to be preserved, routing practices should take into consideration the needs of future generations.

<u>2- Dependencies between infrastructure systems:</u> Different infrastructure systems can be considered heterogeneous due to their varied design guidelines, constraints, owners, age and life cycle. Although these systems are different, they share a fundamental common attribute; they all share the same limited underground space. It is this common location that leads to a great deal of inter-relationship/codependency. As such, routing criteria for a specific system should not follow an insular perspective, but should take into account the relationships/dependencies between colocated systems.

<u>3- Land use patterns:</u> Urban infrastructure routing should be cognizant of the existing and expected land use in the surrounding area. In an urban setting, a high density commercial zone will be much more sensitive to frequent street reconstruction than a low density residential area. A common utility tunnel is an example of a viable design alternative that could be implemented in areas very sensitive to street closures (Adnan & Heng, 2003).

<u>4- Traffic Requirements:</u> The most fundamental aspect that affects almost every member of society is the disruption and delays caused by the need to maintain, inspect, repair, rehabilitate, or replace buried infrastructure. This need has been well articulated by Novick (1990): "*The preparation of new design approaches that focus on ease of inspection, repair, and reconstruction in place, all under traffic*".

4.3 Implement the ontology and decision making model in an interactive web-based environment

The final product of this research is an implementation that makes use of the ontology and the decision criteria. It is envisioned that a web-based GIS system be created for the following purpose:

- To demonstrate the data sharing capabilities across different organizations that the ontology can provide.
- To demonstrate the interoperability opportunities across CAD and GIS media.
- To act as a knowledge repository for knowledge pertaining to the routing of urban infrastructure.
- To act as dynamic knowledge repository for urban infrastructure knowledge in general.

Using GIS is suitable for the problem of urban infrastructure routing in several ways. First of all the problem is a spatial problem involving geometric constraints. The use of a spatial information system like GIS that provides spatial analysis capabilities is very suitable. Secondly, GIS is preferred over CAD systems due to their underlying data model which provides links geometric features to attribute data. This combination allows for simultaneous spatial/attribute data analysis and decision-making.

5 Scope

5.1 Within Scope

<u>1- Focus on buried urban infrastructure:</u> The main focus of the research is on urban infrastructure/utilities. Design guidelines, constraints and the problems facing infrastructure development vary significantly from urban to rural context. The problems associated with congestion, coordination and sustainability are much more profound in an urban context. In congested urban settings almost all utilities are buried underground to maximize the use of street space.

<u>2- Focus on the design stage of projects:</u> The ontology and the decision making model are created to support the design process. Commitments made during this early stage have the

most profound impacts on project outcomes. As such, the efficient management of design coordination has the highest potential to improve overall project outcomes.

3- Focus on the tacit knowledge used in the process of routing buried urban infrastructure: This process is characterized by the following:

- It is a process that is common to all utilities.
- The alignment decision is entails a long-term commitment. Whereas utilities can undergo several material upgrades during their lifecycle, changes to the alignment of networks are highly unlikely especially in congested urban areas (Ross Pattenden, personal communication).
- Formal knowledge in this domain is very scarce. Most of the knowledge resides as best practices/lessons learned by domain experts.

4- Micro-level analysis of utilities: The problem or routing urban utilities will be investigated at the micro or street level. As such the analysis will focus on how to enhance the routing of one or more utility in a particular street. Macro or network level analysis is beyond the scope of the research.

5- Includes both transmission and distribution lines: The analysis will include utilities that serve as main transmission lines (i.e. Not providing service to units) and distribution lines (both main and service lines).

6- Focus on the notion of *sustainable routing*: By considering underground space as a scarce and limited resource, the decision-making model should incorporate criteria for the sustainable use of urban underground.

7- Includes both development design and rehabilitation design: The envisioned system will support design coordination for the following scenarios:

- New development: This involves the case of a new urban development. It is assumes that no buried utilities have been installed yet.
- A new utility in an existing development: This involves the case of introducing a new buried utility in an area that has existing buried infrastructure.
- An existing utility in an existing development: This involves the case where an existing buried utility is repaired, rehabilitated or upgraded.

8- System built on the assumption of free information flow: The envisioned system is created on the tenet that all stakeholders involved:

- Do not object to sharing information.
- Have the required information available at the required time.
- Store data in a consistent format, or that can be converted to a consistent format.

5.2 Beyond Scope

1- Construction methods: Various types of construction methods for buried utilities and the effects these methods have on the project will not be considered in this research because:

- It is not directly related to the process of design coordination or utility routing.
- Extensive work has been performed in this arena and it is very well documented in the literature.

2- Technical details and design issues for specific utilities: The envisioned system is not intended to be used as a technical design tool for an individual utility. As such the system will not attempt to represent the immense body of knowledge pertaining to technical design issues. The system will not assist in designing a better infrastructure network (this is the job of the designer in their respective domains). Instead, the main focus is on aligning this network in such a way that minimizes its conflicts with other systems.

3- Scheduling/phasing details regarding project construction: Aspects of project schedule, work sequencing and construction phasing are beyond the scope of this research.

6 METHODOLOGY & VALIDITY CONCERNS

The following section addresses the proposed methodology for the research along with an analysis of how the methodology will address various validity concerns.

6.1 Research Methodology

The methodology depicted in Figure 1 outlines how each proposed tool will be used to serve the 3 main research objectives. The various types of tools presented in this section can be classified into main tools, software tools, and validation tools as shown in Figure 1.

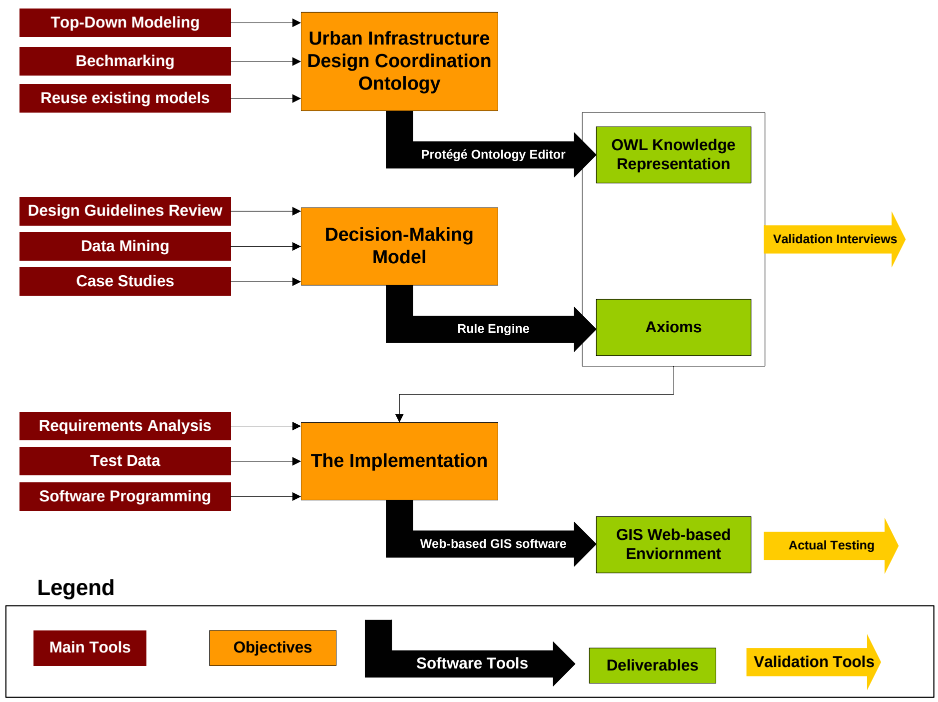

Figure 1 Research Methodology

6.1.1 Literature Review

The literature review will cover the following topics:

- Ontologies.
- Knowledge representation and knowledge-based systems.
- Design processes of urban infrastructure.
- Requirement Analysis
- Data Mining
- Technical guidelines and design manuals for buried utilities of various sectors.

The following section provides a succinct review pertaining to ontologies and the design of urban infrastructure. The detailed review conducted to date has been summarized for brevity.

6.1.1.1 Ontologies

Of the most cited definitions of ontology is that of Gruber (1995): “A formal explicit specification of a shared conceptualization”. The ontology is considered ‘formal’ in the sense that it should be machine readable, ‘explicit’ in the sense that it is used to describe the type of concepts used and any constraints on their use, and ‘shared’ in the sense that a group of people is accepting the definition and that all group members share their consensus.
Uschold and Gruniger (1996) report 3 important areas where ontologies could be used:

1- Communication: By reducing conceptual and terminological confusion by providing a unified framework within an organization, ontologies enable the shared understanding and communication between people with different needs and viewpoints.
2- Interoperability: Ontologies can be used for the creation of integrated environments for different software tools. In this regard, ontologies can be used as inter-lingua or translators between software to enable interoperability between heterogeneous systems.
3- Systems Engineering: Ontologies can be used to support the design and development of software systems. Their use improves the specification, reliability and re-use of the software used by different project partners. In an informal approach, ontologies facilitate the process of identifying the requirements of the system and understanding the relationships between the components of the system.

The new generation of the world-wide-web is the *semantic web*. This is envisioned an extension to the current web where information is given a well-defined meaning, enabling computers and humans to work in cooperation (Berners-Lee et. al., 2001). Ontologies play a central role in the Semantic Web: They provide formal models of domain knowledge that can be exploited by intelligent agents (Knublauch et. al., 2004).

6.1.1.2 Urban Infrastructure Design

The process of infrastructure design can be considered to consist of preliminary design and detailed design. The preliminary design of water distributions systems (and all similar linear systems) consists of 3 main tasks (Mays, 2000):
1- Alignment: In deciding upon appropriate alignment for a pipeline, important considerations include ROW, constructability, access for future maintenance, separation from other utilities. Many municipalities adopt standardized locations for utility pipelines (such that water lines will be generally located 15ft north and east of the street centerline). Such standards complement alignment considerations.
2- Subsurface Conflicts: Issues of consideration regarding utility information:

- Exact depth of utilities, especially when crossing
- Lateral separation between sewer lines and water lines.
- Exact dimensions of utilities: Especially with electrical/telecommunications conduits, plans indicating a single line might in fact depict a huge concrete encased structure
- The sensitivity of telecommunications: A buried ¾” inch telephone line might turn out to be a fiber optic cable that, if severed during construction will result in heavy fines being levied by the communications utility.

3- Rights-of-Way: Must be considered for both construction and future maintenance. If the alignment will be crossing private lands, the proper easements must be acquired.

- Temporary easements assure that installation can safely take place
- Permanent easements must make sure that no construction takes place surrounding the pipe’s alignment to ensure the integrity and accessibility to the pipeline.

The main focus of this research is on the aspect of alignment. Guidelines that govern utility alignments can be broadly classified into two main groups:

- Layout Guidelines: These general guidelines set by municipalities to regulate where each utility should be placed in the ROW
- Clearance Guidelines: These are general guidelines set by individual utilities or other regulatory bodies to regulate the minimum clearances/covers each utility line should have to ensure the safe operation and maintenance of the system. Table 1 shows some of the guidelines in Ontario.

The aforementioned guidelines were created to fulfill a set of design/operation criteria. Generally speaking, criteria for the routing/alignment of buried urban infrastructure can be broadly classified into explicit and implicit criteria. Explicit criteria include the guidelines, regulations and procedures that govern the routing/alignment of buried infrastructure.

Table 1 shows some examples of these explicit criteria. Although these criteria are usually published and well-known, each set of criteria is specific to the individual infrastructure domain (water, wastewater, electricity, gas, telephone, etc…). As such, this knowledge - although formalized - is highly fragmented. By combining all these criteria in a single consistent repository, knowledge from different fields can be shared across all infrastructure domains in a transparent fashion.

Table 1 Examples of some guidelines that govern the routing of buried infrastructure

Procedure / Guideline / Regulation	Issued By	Stipulation
Procedure F-6-1: Procedures to Govern Separation of Sewers and Watermains	Ministry of Environment Ontario	A minimum separation distance of 2.5m should always be kept between sewers and watermains
Support of Gas pipelines in the vicinity of excavation	Enbridge & Union Gas	For gas distribution pipelines the minimum separation between any buried utility and a gas line is 1 foot for parallel utilities and 50mm for crossing utilities.
CSA Standard C22.3 No. 7-94 Underground Systems	CSA	Hydro services must have a minimum cover of 65cm

Implicit criteria are not explicitly articulated in a published document but can be of tremendous benefit if taken into consideration. These criteria can be further classified into:

Formal Implicit Criteria: These include best practices that are systematically employed by designers in routing buried infrastructure. They have evolved from informal criteria and have become an industry *de facto*. These criteria tend to be more situation-based compared to the more general explicit criteria.

Informal Implicit Criteria: These include the lessons that have been learned from previous projects but have not yet reached an industry consensus to be recognized as a best practice. These criteria tend to be much more project-specific and cannot be used as a rule of thumb across all projects. These are the most difficult set of criteria to elicit due to their tacit nature.

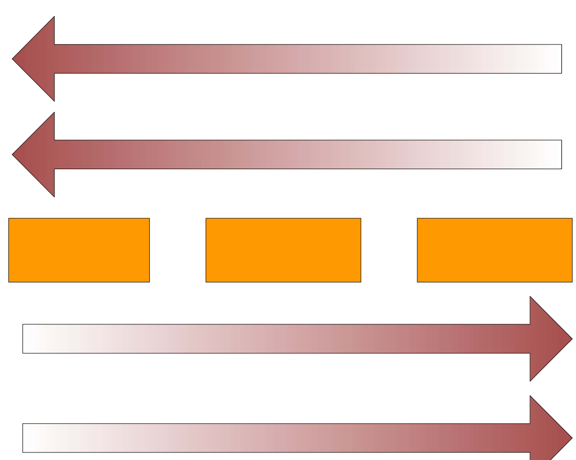

Figure 2 Comparison between the decision criteria

6.1.2 Benchmarking

Several initiatives have been undertaken by industry groups, researchers and software developers to model concepts related to civil infrastructure. Most initiatives have focused on representing the physical concepts and overlooked a great deal of logical/soft concepts. The modeling approach followed by this research will depend on benchmarking and reusing portions of these initiatives.

As such, the ontology will attempt to reuse the most comprehensive of these product models to be the core or the envisioned infrastructure ontology. The main contribution of the ontology lies in its definition of the soft/logical concepts and the relationships between various concepts (the semantics).

6.1.3 Top-down Modeling

Top-down modeling approaches depend on human knowledge and include a good deal of modeling art. Here a researcher or an industry group, based on their collective expertise, develop a theory about what data are most common in the industry and their interrelationships. This procedure starts at the highest level of abstraction and proceeds towards the lowest level. The approach is usually an iterative one that involves continuous updating and refinement based on analysis.

The bottom-up approach will also be utilized via a number of tools (T3, T5, T6, T7). This approach involves individual knowledge discovery and model re-evaluation. As such, the process of creating the ontology and discovering the knowledge using the aforementioned tools is a two-way process.

6.1.4 OWL Representation

The OWL Web Ontology Language is an XML-based langauge endorsed by the World-Wide-Web consortuim designed for use by applications that need to process the content of information instead of just presenting information to humans (OWL, 2004). The devloped ontology will be implemented using OWL as it is currently considered the most recognized language for representing ontologies. OWL has 3 increasingly expressive sublanguages (OWL-Lite, OWL-DL and OWL-Full). It is expected to represent the ontology in either OWL-Lite or OWL-DL due to the complicated reasoning capabilities of OWL-Full. The Protégé Ontology editor (Protégé, 2004), and the OWL plug-in will be utilized for this task.

6.1.5 Preliminary Case Studies

Select case studies involving the rehabilitation and replacement of existing infrastructure have been investigated. So far 3 case studies have been considered. The purpose of undertaking these preliminary case studies is to:

1- Understand how infrastructure rehabilitation and reconstruction occur in a congested urban setting.
2- Understand the constraints that govern the routing of buried urban infrastructure.
3- Gain a better understanding of the involved stakeholders, their roles, requirements, and common practices.
4- Create a preliminary model of the decision making criteria based on the aforementioned information.

6.1.6 Data Mining

Based on the preliminary model that will be formalized into the ontology, data mining will be used to extract knowledge from guidelines/codes/documents pertaining to infrastructure development. This form of mining commonly referred to as *text mining* attempts to extract unknown patterns and associations from large text repositories.

In the construction realm very few attempts have been made to mine data. Soibelman & Kim (2002) developed a framework for applying knowledge discovery in databases and applied it to a case study for the Resident Management System database provided by the U. S. Corps of Engineers. The process was applied to identify the causes of construction activity delays. Caldas & Soibelman (2002) utilized a machine learning algorithm named Support Vector Machines (SVM) to perform text mining operations on 4,000 sets of documents in a construction company. The objective of the mining operations was to automatically classify documents into their representative categories. Using the SVM technique resulted in 92% accuracy in classification.

Potential Candidates:

- The repository of emails belonging to one of the research industry collaborators (See Section **Error! Reference source not found.**). Mining this type of data repository will elicit more of the informal implicit criteria due to the specific nature of the data.
- The set of *InfraGuide* documents. InfraGuide brings together a large network of organizations and individuals in the infrastructure domain who contribute their experienced knowledge of municipal infrastructure. To date, Infraguide has published 37 best practices relating to all facets of infrastructure development (InfraGuide, 2004). All documents are readily available in an electronic format. Mining this type of documents will elicit more explicit/formal implicit criteria due to the general nature of the document.

6.1.7 Detailed Case Studies

In total, it is expected that 7 detailed case studies will be investigated. The detailed case studies will go into more depth after acquiring basic knowledge from the first set of preliminary cases. These detailed case studies will elicit a large portion of the tacit knowledge in the decision criteria.

6.1.8 Validation Interviews

The interviews will be conducted with a set of experts in the design of urban infrastructure. The list of research collaborators form the industry (Section **Error! Reference source not found.**) will be increased as new contacts become available. The purpose of the interviews will be to validate the model (ontology) as well as the knowledge (axioms) elicited via the criteria analysis. This validation will ensure the external validity of the research.

6.1.9 Requirement Analysis

Prior to beginning the software implementation, a requirements analysis exercise will be conducted to explicitly articulate what the proposed software will accomplish. This methodology is well-established in software engineering and is considered the corner-stone of software development (IEEE, 1994). The requirement analysis will address:

- The system’s objectives: What business objectives will this system help its users achieve.
- The potential benefits of the system.
- A description of the business process and how the system will be used in this context. Use case scenarios will be developed for this.
- Details pertaining to users’ roles and responsibilities and how the system will fit into what they do.
- Interactions with other systems (if any).
- A clear and concise statement of functionality.
- Security issues related to data access.
- The anticipated system performance.

6.1.10 Actual Testing

This tool will be used to test the GIS-based portal. Members of the ORCGA are excellent candidates for testing the system. Testing metrics that will be evaluated include:

- Suitability for application.
- Accuracy of implementing the embedded knowledge
- Degree of flexibility in acquiring new knowledge.
- Speed of handling large amounts of data.

7 CONTRIBUTION

The main contributions of this research include:

1- Formalizing a model for infrastructure products: Research in this area has been very scarce with most product modeling efforts focused on the building construction industry. Froese (2003) identifies the need to extend product models to include civil infrastructure as one of the major upcoming steps needed for model-based exchange.

2- Creating a generic model for representing knowledge pertaining to civil infrastructure: The developed Ontology provides a conceptualization for knowledge in civil infrastructure. Its generic framework (owed to the top-down modeling approach used), will enable knowledge to be seamlessly added without the need to re-create a model.

3- Formalizing the explicit knowledge used to route buried infrastructure: This knowledge is spread across different infrastructure industries that work in isolation. This research will combine these guidelines to provide a unified and formal representation of this knowledge that will be made available to all stakeholders involved.

4- Identifying the tacit knowledge used to route buried infrastructure: By analyzing case studies and data mining, the explicit knowledge that experienced designers use in routing urban infrastructure will be exposed. This is significantly important due to the scarcity of literature as well as the superficial nature of current design guidelines addressing this issue.

5- Creating a prototype portal for sharing infrastructure information among utility owners: The final product of this research will serve as a prototype portal for exchanging infrastructure related information among utility owners. Due to security and interoperability barriers, such a portal has never been implemented. Its implementation as a prototype housed in an academic workspace will create an excellent opportunity for evaluating the usefulness of such a system.