Semi-Annual Progress Report

Project Number and Title: 4-2 Future-Proof Transportation Infrastructure Through Proactive, Intelligent, and Public-involved Planning and Management
Research Area: Thrust 4 Connectivity for enhanced asset and performance management
PI: Jin Zhu, Ph.D. Assistant Professor, University of Connecticut
Reporting Period: *Oct 01, 2018 – Mar 31, 2019*
Date: *March 25, 2019*

Overview:

Overview and summary of activities performed during previous months:
During the last reporting period, the research team of project 4-2 initiated the development of an integrated ontology for transportation infrastructure planning. Ontology is a systematic description of entities with regards to their properties, relationship, and constraints expressed by axioms. Due to its ability in promoting sharing of knowledge bases, knowledge organization, and interoperability among systems, ontologies have been used extensively in many domains and studies (e.g., disaster management, business modelling, disease management). The transportation research domain has long been benefitting from the application of ontologies. Ontologies are particularly fitting to handle the challenges arising from the large volume and variety of transportation related data (e.g., survey, sensor data). Majority of the ontologies developed in transportation research domain focus on trip planning, trip disruption, traffic management, service monitoring, and urban freight transport problems. Despite the extensive use of ontologies in transportation research, an integrated ontology for transportation infrastructure planning is still missing. Transportation infrastructure planning can be defined as the process of making decisions concerning the potential changes required for transportation related infrastructures to improve the quality of life. Transportation infrastructure planning is a complex process. Without a structured ontology for information and knowledge management, there are major challenges in transportation infrastructure planning. First, transportation infrastructure planning documents and publications from different sources may have different focus and use different structures and terminologies. For example, different agencies may have different definitions and metrics to assess transportation infrastructure performance. Therefore, it will be difficult to compare or integrate transportation infrastructure performance data and establish uniform baselines across different agencies. Second, a robust transportation infrastructure planning requires information from varied sources such as online news, social media, technical reports, and many more. There is no way to effectively extract, analyze, and utilize the information from heterogeneous sources without a formal structure. To address these challenges, the researchers developed a six-step methodology to build the preliminary infrastructure planning and management ontology (Figure 1). The ontology is proposed in this study via two topic modelling techniques: Latent Dirichlet Allocation (LDA) and Non-negative Matrix Factorization (NMF). The models have been used to extract important and emerging concepts related to transportation planning from 20 transportation planning documents. These documents include transportation vision statements, long-range transportation plans, long-term regional transportation and land use plans, and many more. Majority of these documents were developed by state DOTs, while some were developed by legislative bodies at the city level. In some cases external consultants were assigned to provide support to develop such documents.

How these activities are helping achieve the overarching goal of the project:
The activities conducted is related to Task 1 in the proposed research: develop an integrated information platform to collect and analyze information for better transportation infrastructure planning and management from various sources. Information could include physical conditions of infrastructure obtained from inventory, inspections, and sensors, as well as future risks related to climate change and extreme weather and their impacts. Such integrated ontology could help: (1) increase interoperability across different transportation infrastructure plans published by different governmental and non-governmental agencies at different levels; (2) facilitate the collection and analysis of data from various sources including social media (e.g., Twitter). Essentially, the ontology would help to build an integrated framework for smart transportation planning by incorporating data from different sources into a smart knowledge management and decision making system. It

ultimately helps decision makers and planners to have a holistic approach to plan, build, and manage our transportation infrastructures.

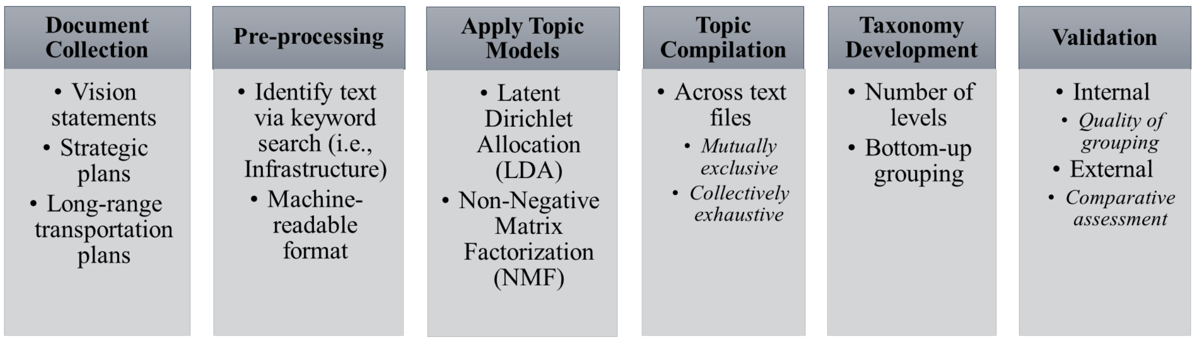

Figure 1. Ontology development steps

Accomplishment achieved under the project goals:
Two topic modelling techniques: LDA and NMF were implemented in this study to analyze the transportation infrastructure planning documents in order to identify different dimensions and importation concepts in the domain. LDA and NMF are two widely used topic modelling techniques. LDA is based on probabilistic graphical modelling while NMF relies on linear algebra. While LDA and NMF have significantly different mathematical underpinnings, both techniques are capable of returning the most pertinent topics in a document. In this study, we used both techniques for analysis and compared the results to ensure we capture all the important topics for transportation infrastructure planning. Table 1 shows a sample output from one document using NMF.

Table 1. A sample output of the application of topic models

Topic Model	**Topic Example**	**Contexts (from "Invest in Transit 2018-2023 Regional Strategic Plan")**
Non-negative Matrix Factorization (NMF)	Funding, agencies, projects, state, local	• The agencies are ready to deliver more improvement projects, but more funding is required to do that. • The agencies have filled some of the funding gaps with short term fixes and by working with local governments and agencies. • If supported by a diversity of state, federal, and local funding commitments, it would empower agencies to improve the systems. • Agencies will be specific and transparent about funding needs and usage of funds. • RTAs Project Management Oversight office monitors the implementation of capital projects and found in its most recent report that 95 of state bond funded projects were tracking on time and 100 were on budget notwithstanding delays related to state funding. • Diversify and increase transit capital funding sources through state and local funding commitments of new revenue sources or expansions of existing revenues (e.g. gas tax). • The region will continue to seek federal funding and apply it to regionally and nationally significant projects. • The agencies have a track record of delivering on large capital project commitments.

A bottom-up approach was used to aggregate analysis results from different documents and build the preliminary ontology. For example, different hazards such as tornado, wildfire, and heatwave were included in the contexts of topics related to natural hazards and their impacts on transportation infrastructure. Therefore, these specific hazards were identified as the bottom-level entities and were grouped together at a higher level of the taxonomy under natural risk. Since there were other topics identified extensively discussing man-made hazards and their impacts on transportation infrastructure, "natural risks" and "man-made risks" were identified as two parallel concepts and grouped together under "risk" category in the taxonomy (see Figure 2 for details). Through the study, four levels (i.e., level 0, level 1, level 2, level 3) in total were identified in the taxonomy to organize all the important information and concepts in a structured way. Level 3 is the lowest level with the finest level of granularity, followed by level 2, level 1, and finally level 0. Level 3 is not shown in Figure 2 due to the large magnitude of information.

Opportunities for training/professional development that have been provided:
Since the research results are still preliminary, no training/professional development opportunities have been provided yet.

Activities involving the dissemination of research results
The results are disseminated during the TIDC proposal review workshop on November 8-9, 2018 in Portsmouth, NH and a meeting with Conn DOT representatives on February 28th in Storrs, CT. A conference paper titled "Towards the ontology development for smart transportation infrastructure planning via topic modeling", as a result of this study, is accepted by the 36th International Symposium on Automation and Robotics in Construction. The PI will present the paper during the conference.

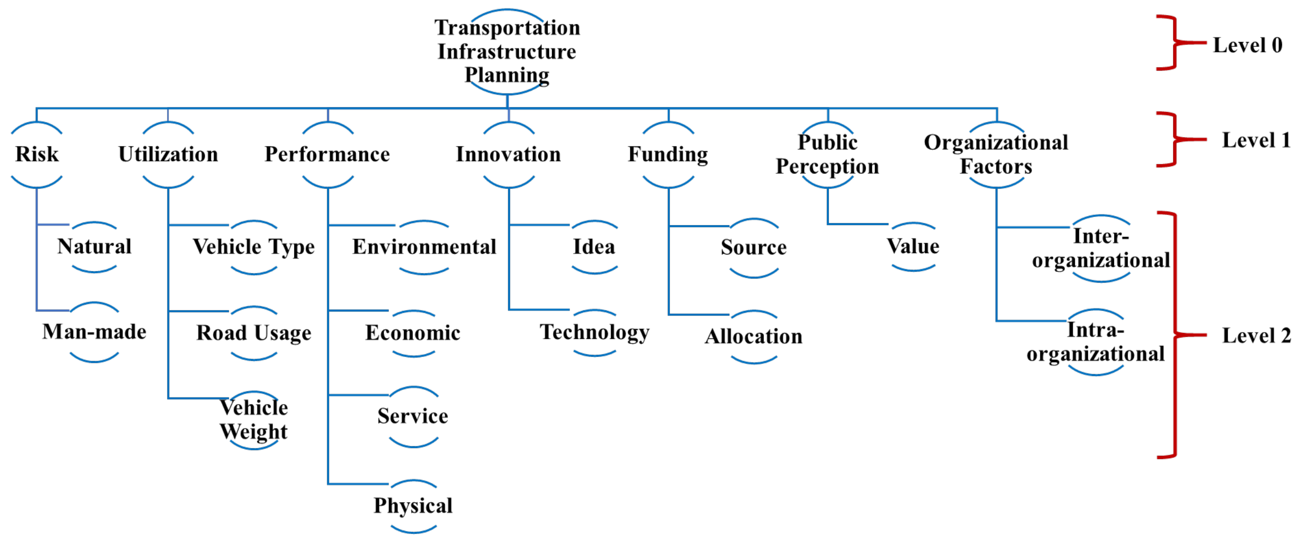

Figure 2. A preliminary future-proofing transportation infrastructure planning ontology

Participants and Collaborators:

Student Participants: Sudipta Chowdhury, PhD student, Civil and Environmental Engineering, graduate research assistant

Changes:

The research activities were conducted according to the plan submitted in the approved proposal. There were no significant changes or problems occurred so far.

Planned Activities:

In the coming six months, we plan to:

(1) Further refine the preliminary ontology by incorporating more documents and different types of data into analysis;
(2) Validate the ontology with subject-matter experts;
(3) Implement the ontology in a case study to test its capability in guiding the collection and analysis of information for better transportation infrastructure planning and management from various sources.