Towards the Ontology Development for Smart Transportation Infrastructure Planning via Topic Modeling

S. Chowdhury[a] and J. Zhu[b]

[a]Department of Civil and Environmental Engineering, University of Connecticut, 261 Glenbrook Road Unit 3037, Storrs, CT 06269-3037
[b]Department of Civil and Environmental Engineering, University of Connecticut, 261 Glenbrook Road Unit 3037, Storrs, CT 06269-3037
E-mail: sudipta.chowdhury@uconn.edu, jzhu@uconn.edu

Abstract
Media age has seen a huge amount of data flowing in from all directions, be it online news sources, social media, technical documents, and many more. There is a huge scope of these data sources for utilization in the transportation sector that can potentially improve the current practice of transportation infrastructure planning. In order to effectively capture, analyze, and utilize the information from various sources, ontologies are useful tools as they can provide clear and structured knowledge in the transportation domain. Majority of the existing transportation-related ontologies focus on traffic management and route planning. The objective of this paper is to initiate the development of an integrated ontology that can help with long-term planning and decision-making of transportation infrastructure by proposing a preliminary taxonomy in this domain. To this end, 20 transportation planning visionary documents published by government agencies were collected and analyzed using topic modelling techniques. Specifically, two topic modeling methods: Latent Dirichlet Allocation (LDA) and Non-negative Matrix Factorization (NMF) models were used to extract important and emerging concepts related to transportation infrastructure planning. Leveraging the important and emerging concepts, a preliminary taxonomy of transportation infrastructure planning was then developed and presented.

Keywords –
Smart Transportation infrastructure planning; Ontology; Taxonomy; LDA; NMF

1 Introduction

Ontology is a systematic description of entities with regards to their properties, relationship, and constraints expressed by axioms [1]. 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) [2-5].

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) [6-7]. Majority of the ontologies developed in transportation research domain focus on trip planning [1, 8], trip disruption [9], traffic management [10-11], service monitoring [12], and urban freight transport problems [13, 14]. Despite the extensive use of ontologies in transportation research, to the best of the authors' knowledge, there is no scholarly work available in literature that directly focuses on an integrated ontology for transportation infrastructure planning. In many of the existing studies, transportation infrastructure planning had been considered as an auxiliary element towards other functions of transportation (e.g., see [15]). An assumption had been made in many of these studies that a desirable level of transportation infrastructure quality was already achieved for other transportation planning purposes such as trip planning, traffic management, and service monitoring. However, infrastructure is the foundation for realizing transportation functions. With the help of an integrated transportation infrastructure planning ontology, transportation infrastructure can be planned and managed in a more holistic way. Thus, transportation infrastructures can be better equipped to serve traffic demand, public safety, and social needs.

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 [5]. There is no way to effectively extract, analyze, and utilize the information from heterogeneous sources without a formal structure.

To address these challenges, the objective of this study is to initiate the development of an integrated ontology for transportation infrastructure planning. 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.

As a first step in developing the ontology, a preliminary taxonomy for transportation infrastructure planning 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 concepts are then appropriately analysed and categorized to create a hierarchical taxonomy for transportation infrastructure planning.

2 Methodology

The methodology used to develop the preliminary taxonomy includes six steps. As shown in Figure 1, the first step was document collection, in which transportation planning related documents were collected.

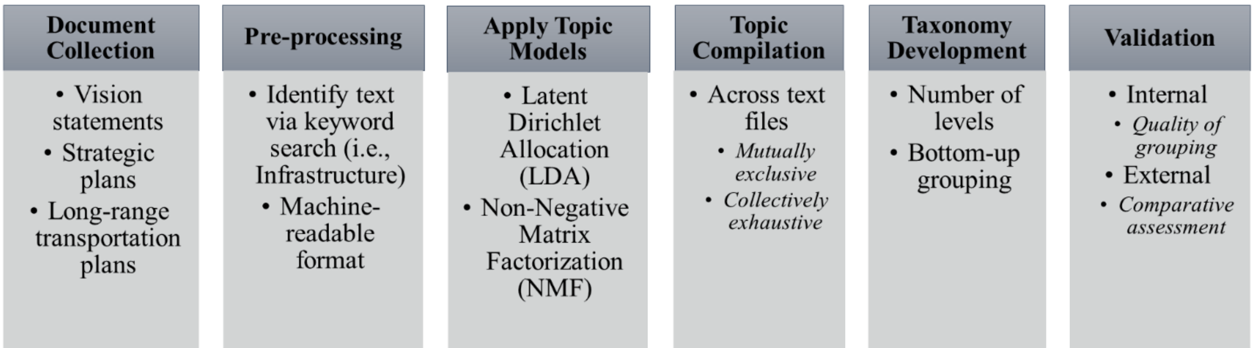

Figure 1. Taxonomy development steps

In the second step, useful texts were extracted from the collected documents and processed into a format that could be readily used for topic modelling. In the topic modelling step, two modeling techniques, i.e., LDA and NMF, were implemented on data extracted from each document to generate relevant topics. Then, different topics identified across all the documents were compiled and summarized in step four, i.e., topic compilation. The fifth step was taxonomy development, in which the structure of a preliminary transportation infrastructure planning ontology were developed. Finally, in step six, the proposed taxonomy was validated. Details of each of the steps in the proposed methodology are provided below.

2.1 Document Collection

As a first step, transportation planning documents were collected via a standard google search using key words such as transportation planning vision, transportation long-term plan, strategic transportation plan, etc. In total, 20 documents were identified through this process. These documents include transportation vision statements [13], long-range transportation plans [14], long-term regional transportation and land use plans [16-17], 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.

2.2 Pre-processing

Since transportation planning is such a broad term and covers many different aspects, the collected documents in the first step are often times long-winded and hence, contain a variety of information that may not be directly pertinent to transportation infrastructure planning. Hence, data pre-processing was conducted by searching the keyword *"infrastructure"* within each document. Paragraphs with the keyword "*infrastructure*" in them were extracted from each document under the assumption that such texts were more apposite to the focus of this study. Separate text files were created for each document. These text files were then further processed to eliminate the possibility of encoding errors when they were read via *Python* for topic modelling. For example, all the apostrophe characters were removed from the text files before topic models were applied.

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" [21])
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.

2.3 Apply Topic Models

The processed data was then analyzed using topic modelling. Topic modelling is essentially a form of text mining that utilizes either unsupervised or supervised statistical machine learning techniques to identify patterns in structured or unstructured text [18]. It employs the process of similarity by clustering the words and identifying topics. Two topic modelling techniques: LDA and NMF were implemented in this study. LDA and NMF are two widely used topic modelling techniques. LDA is based on probabilistic graphical modelling while NMF relies on linear algebra [19-20]. 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 to ensure we capture all the important topics for transportation infrastructure planning. Before delving into detail about the mechanisms of LDA and NMF, it is important to clarify that in this study we ran topic models for each text files individually. Therefore, each text file generated is considered as a collection of *documents*. Here, *document* means a complete sentence (shown as one line in text files). The topic modeling results generated for each text file by both LDA and NMF were essentially based on the analysis of the collection of *documents* in each file. For better representation, from now on *document* in the text files will be referred to as "Single Line Source (SLS)" throughout this paper.

The mechanisms and implementation procedures of LDA and NMF are provided below. Machine learning library Scikit-learn was used in conjunction with Python to implement both LDA and NMF for identifying topics and their respective contexts [21].

2.3.1 Latent Dirichlet Allocation (LDA)

Let us suppose there are D SLSs in a text file with W words. Assuming K topics to discover, LDA was implemented as

Step 1: Randomly assign each word in $d \in D$ to one of the $k \in K$ topics

Step 2: Identify topic representations of all the SLSs and word distributions of all the topics.

Step 3: For each SLS $d \in D$:

- Go through each word $w \in W$
  - For each topic $k \in K$:
    - Compute the proportion of words in SLS $d \in D$ that are currently assigned to topic $k \in K = p(k|d)$
    - Compute the proportion of assignments to topic $k \in K$ over all SLSs that come from the word $w \in W = p(w|k)$
  - Reassign $w \in W$ to a new topic based on the probability $p(k|d) * p(w|k)$

Step 4: Repeat step 3 a large number of times to reach a steady state solution.

2.3.2 Non-negative Matrix Factorization (NMF)

Let us consider a nonnegative data matrix where each column of X corresponds to a SLS $d \in D$ and each row to a word $w \in W$. Each entry (m, n) in data matrix X represents the number of times the mth word appears in the nth SLS. Given such a matrix X with rank r, it can be factorized into two matrices V and H such that

$$X\,(:,n) \approx \sum_{k=1}^{r} V(:,k)\,H(k,n) \quad with\, V, H \geq 0$$

The goal is to find the best possible factorized matrixes that minimizes the following objective function

$$\underset{V,H\geq 0}{minimize} ||X - VH||^2$$

Since the weights in the linear combinations are nonnegative (i.e., $H \geq 0$), only the union of the sets of words defined by the columns of V can be used to reconstruct the original SLSs. Hence, the columns of V can be interpreted as topics.

The primary output of both LDA and NMF is a set of topics, each of which consists of a cluster of words. However, with only words in a topic, it is sometimes challenging to interpret or determine what each topic is about. This is non-trivial especially in ontology or taxonomy development as understanding of the context of each topic is crucial for better representation of the categories, properties, and relations among the data, concepts, and entities. Hence, in this study, not only the topics but also the context of each topic (i.e., SLSs that are the origins of each topic) are extracted and evaluated. A sample output of the topic models is provided in Table 1. This example shows one of the topics extracted from "Invest in Transit 2018-2023 Regional Strategic Plan" that provides a regional transit strategic plan for Chicago and Northeastern Illinois using NMF [22]. This topic covers a set of 5 words and the 8 SLSs. With the help of the contexts, it is evident that this topic is about the funding of transportation infrastructure projects. More specifically, the diversity of funding sources and transparent use of funding are the two focuses of this topic.

2.4 Topic Compilation

Due to the varying nature of collected files, a vast array of topics were generated. In this step, first, topics generated via LDA and NMF from each text file were compared and compiled. Then, topics identified across different text files were complied. Two criteria were used for topic compilation: mutually exclusive and collectively exhaustive. Mutually exclusive ensures that no topics were repeatedly counted and collectively exhaustive ensures that all the important topics identified were included.

2.5 Taxonomy Development

Taxonomy development step provides a preliminary solution about what concepts and/or factors impact and should be considered in smart transportation infrastructure planning. Moreover, successful implementation of this step can provide a hierarchical structure to organize different concepts and/or factors. A bottom-up approach was used in this step to develop the taxonomy using topic information generated in the previous step. 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). In this 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. It should be noted that caution needs to be taken in this step, as subjective judgement is required for appropriate grouping.

2.6 Validation

As the last step, both internal and external validation of the preliminary taxonomy were conducted. First, internal validation was conducted by the authors to ensure the quality of topic modelling and taxonomy development. The topics identified as well as the concept grouping structure were cross checked. Second, external validation was conducted by comparing the taxonomy with other existing studies. For example: *mobility* was a key topic captured across multiple documents and thus was identified as a concept related to service performance of transportation infrastructure in the proposed taxonomy.

In related transportation literature [23], *mobility* is used as one indicator of service quality of transportation infrastructure as well. Therefore, it was validated that *mobility* was properly grouped with other concepts that are relevant to *service quality* of transportation infrastructure under the *performance* domain in the proposed taxonomy.

3 Proposed Taxonomy

This section presents the preliminary transportation infrastructure planning taxonomy developed in this study. Figure 2 shows the first three levels in the taxonomy. Level 3 is not shown in Figure 2 due to the large magnitude of information. A sample representation of all the levels is demonstrated in Table 2 using the example of *risk* domain. The details of the proposed preliminary taxonomy are explained as follows, organized by the seven categories of concepts in Level 1:

Risk

A few documents explored the potential of *natural* and *man-made hazards* on transportation infrastructure and risks associated with them. Among the man-made hazards, traffic crashes were considered to be the most important source and hence, were investigated in detail. Apart from traffic crashes; hazmat spills, cyber-attacks, terrorism, etc. were also explored in several documents as potential man-made hazards that could have a detrimental effect on transportation infrastructure.

It was observed that natural hazards and their associated risks were given more importance compared to man-made hazards. A variety of natural hazards, such as wildfire, drought, heatwave, flooding, landslides, avalanches, storm surges, and earthquakes were extensively discussed and investigated for their threats to the transportation infrastructure (e.g., prolonged heat waves could increase the premature deterioration of infrastructure).

Utilization

Utilization refers to the utilization of transportation infrastructure in terms of the type and weight of vehicles using the infrastructure, and the usage frequency of infrastructures. The information of utilization has significant impacts on current and future transportation infrastructure planning. For example: in Georgia's statewide transportation plan [24], it highly emphasized that flow of freight in terms of weight, the type of vehicle most commonly used for freight transportation, and the usage of roads could provide useful insights into the infrastructure needs. Hence, *vehicle type*, *road usage*, and *vehicle weight* need to be carefully monitored. *Vehicle type* can refer to different classes of vehicles specified by The Federal Highway Administration (FHWA), *road usage* can refer to estimating Average Annualize Daily Traffic (AADT), and *vehicle weight* can refer to the maximum allowable weight that each class of vehicle can carry. *Vehicle type*, *road usage*, and *vehicle weight* can holistically provide useful information to maintenance requirement and frequency, as well as future needs of transportation infrastructures.

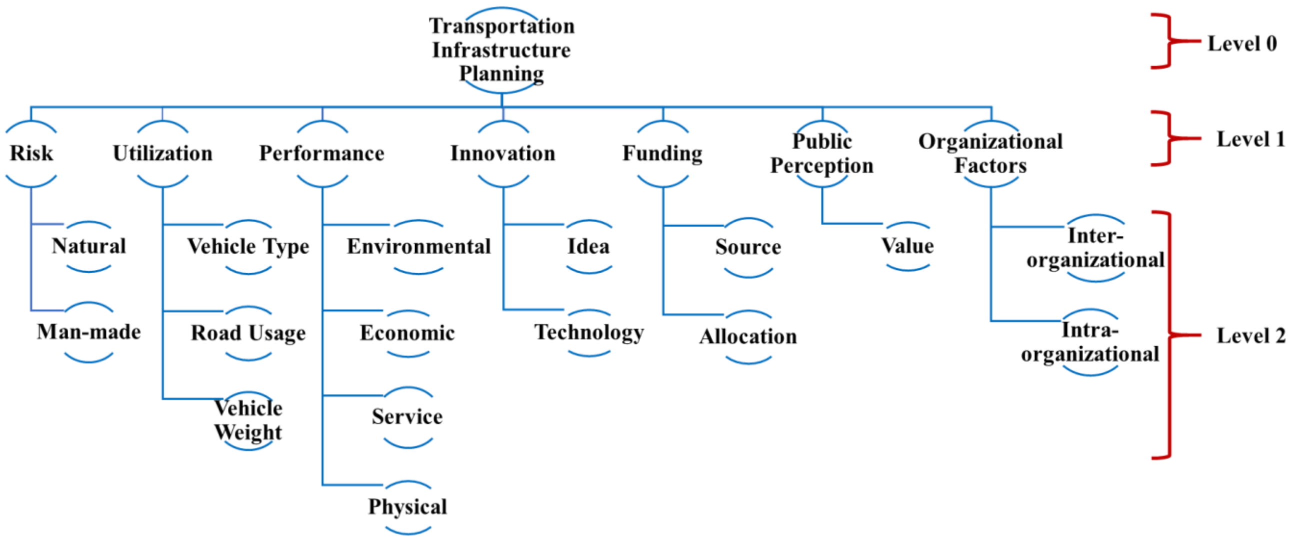

Figure 2. Illustrating the proposed transportation infrastructure planning taxonomy

Performance

Performance refers to the capability of current transportation infrastructures to fulfill the user needs.

Performance has been categorized into four classes in the proposed taxonomy, i.e., *environmental*, *economic*, *service*, and *physical*. In terms of *environmental* performance, different state DOTs or local legislative bodies provided different benchmarks such as air quality, water quality, greenhouse gas/carbon emission, and preservation of open space. Among them, greenhouse gas/carbon emission had been the focal point in majority of the documents that discussed the environmental performance of transportation infrastructures. With regards to *economic* performance, employment growth, supply chain and logistics reliability, inter-industry economic competition, etc. were identified as the major indicators for transportation infrastructure. *Service* performance is arguably the predominant measure of transportation infrastructure performance that was emphasized in almost all the documents. There were a number of transit service performance measures that had been discussed in detail such as ride quality, vehicle mobility, accessibility, intermodal coordination, transit security, connectivity, flexibility, and quality of transit fare system. On the other hand, service measures in terms of personal mobility such as walkability were also highlighted in a few documents. There were a few additional service measures that were not associated with transit but with quality of transportation infrastructure in general. For example, quality of road signs, quality of lightings, and coverage of emergency service patrol were discussed as indicators of transportation infrastructure quality in many documents. Finally, *physical* performance measures refer to the physical conditions of transportation infrastructures. *Physical* performance measures refer to the condition of transportation infrastructures such as bridge, culvert, airport, parking lots, rapid transit, heavy rail, and urban rail. The physical performance measures indicators can range from visual inspection ratings to complex performance index (e.g., pavement condition index (PCI)).

Innovation

Innovation has brought a lot of changes to today's transportation infrastructure. In the proposed taxonomy, innovation has been identified as an important concept in Level 1 and can be represented by two classes, i.e., *idea* and *technology*. *Idea* refers to new lines of thoughts to improve the quality of transportation system. For example, complete streets (i.e., streets that are designed and maintained to enable convenient and safe travel and access for all users regardless of modal choice) and greenway corridors (i.e., corridors that are specifically designed for recreational or pedestrian use ensuring environmental, wildlife habitat, and water resources benefits) are new and innovative ideas that were introduced in many of the documents to future-proof transportation infrastructure [17].

In terms of *technology*, it was found that automated and connected vehicles, novel data management framework, alternative fuel, alternative energy (e.g. solar power) etc. had been gaining attention and adopted widely across USA and hence, need to be exhaustively investigated regarding their impacts on transportation infrastructure planning and development. For example, with more electric vehicle on the road, there is a need to integrate the electric charging infrastructure development into transportation planning.

Funding

Funding was identified to be a critical factor in transportation planning. It is often challenging to acquire adequate funding to maintain and improve the existing transportation infrastructure, as well as to develop new infrastructure. Leveraging the information from topic modelling results, there are two level-2 concepts under funding: *source* and *allocation*. *Source* refers to the authorities and financial entities that provide money for transportation infrastructure related ventures. Three types of major sources were found from the documents, i.e., public, private, and public-private partnerships. *Allocation* refers to allocating the monetary funds to different areas or activities. For example, allocation of funding can be based on different levels of transportation development and improvement needs in different areas (e.g., rural, urban, or sub-urban). Allocation of funding could also be based on types and urgency of activities (e.g., preservation, modernization, or expansion). Different funding allocation rules might be applied to make sure the money is used in a most effective way to the most needed areas and activities.

Public Perception

Public perceptions and opinions are influencing transportation infrastructure more and more. One example found in the topic modelling result from this study is that, people (especially the younger generation) are becoming more conscious of the transportation choices they have and how these choices affect not only their lives but also the environment and society. They are becoming more aware of the detrimental effects of auto-dependency and want a transportation system that promote healthy living. Moreover, there is a growing consensus among the people that reducing auto-dependency help preserve the environment in many ways as well. Hence, it was observed from multiple documents that there had been a growing demand in recent years for developing and improving biking, hiking, and walking infrastructures. This phenomenon has been termed as *value* in the proposed taxonomy under public perception category. There are other types of public perceptions such as complaints that might affect transportation infrastructure planning, which were not captured in the

20 documents analysed in the current study. More documents will be included in the analysis to further develop the concept of public perception and its impacts on transportation infrastructure planning in future work.

Organizational Factors

Organizational factors relate to the culture of the transportation planning agencies (e.g., DOT) towards better management practices. These organizational factors greatly affect the adoption and implementation of any transportation planning policies and decision-making process behind. There are two major classes of organizational factors: *intra-organizational* and *inter-organizational* factors. *Intra-organizational* factors include an agency's commitment to not only improve internal communications and engagement processes, but also ensure that plans, programs, and projects are in line with its long term strategic goals. For example, San Francisco Municipal Transportation Agency's strategic plan emphasize that all agency plans, programs, and projects of the agency must be adjusted accordingly to fit the scope of the strategic plan [25].

On the other hand, *inter-organizational* factors refer to the collaboration and coordination attitude of transportation planning agencies with other transportation partners. A high level of coordination and collaboration among state government leaders, congressional delegation, land use planning governing bodies, industry stakeholders, tribes, non-profit organizations, and the general public were found to be significant in ensuring robust transportation infrastructure [e.g., 26].

Table 2. Sample construction of transportation planning taxonomy

Level 0	**Level 1**	**Level 2**	**Level 3**
Transportation infrastructure planning	Risk	Natural	Wildfire, drought, heatwave, flooding, landslides, avalanches, storm surges, earthquakes
		Man-made	Hazmat spills, cyber-attacks, terrorism, traffic crash

4 Conclusions

In order to develop an integrated ontology for transportation infrastructure planning, a taxonomy was proposed in this study as the first step. The transportation infrastructure planning taxonomy was developed using topic modelling techniques based on textual information from 20 transportation planning documents published by government agencies. The taxonomy and the ontology that will be built could potentially address the interoperability challenge of transportation plans at various levels and across different regions. In addition, they could provide a formalized structure to collect, organize, analyse, and utilize information from different sources to help transportation infrastructure planning. The ultimate goal of this study is to create a smart transportation infrastructure planning platform that can help planners to make informed infrastructure planning and management decisions using information and knowledge effectively with the help of the ontology.

The current study has several limitations. First, both the number and scope of documents included in the topic modelling analysis in this study are limited. For future study, the authors will collect more data from more sources, including online social media data to implement topic modelling and identify emerging concepts related to transportation infrastructure planning. The taxonomy will be further developed based on new concepts and information captured. Second, in order to grow the taxonomy proposed into an integrated ontology, more elements including attributes, relationships, rules, and restrictions of concepts will need to be investigated extensively in future studies.

Acknowledgement

The authors acknowledge the support received from U.S. DOT under its Transportation Infrastructure Durability Center (TIDC). The views expressed in this paper are those of the authors. The findings, conclusions or recommendations either inferred or specifically expressed herein do not necessarily indicate acceptance by US DOT.

5 References

[1] Houda, M., Khemaja, M., Oliveira, K., & Abed, M. (2010, May). A Public Transportation Ontology to Support User Travel Planning. In Research Challenges in Information Science (RCIS), 2010 Fourth International Conference on (pp. 127-136). IEEE.

[2] Babitski, G., Probst, F., Hoffmann, J., & Oberle, D. (2009). Ontology Design for Information Integration in Disaster Management. GI Jahrestagung, 154, 3120-3134.

[3] Wang, Z., & Lei, J. (2013, June). A Study of Spatial Information Sharing Model Based on Geological Hazard Domain Ontology. In Computational and Information Sciences (ICCIS), 2013 Fifth International Conference on (pp. 268-270). IEEE.

[4] Upward, A., & Jones, P. (2016). An Ontology for Strongly Sustainable Business Models: Defining an Enterprise Framework Compatible with Natural and Social Science. Organization & Environment, 29(1), 97-123.

[5] Chowdhury, S., Emelogu, A., Marufuzzaman, M., Nurre, S. G., & Bian, L. (2017). Drones for Disaster Response and Relief Operations: A Continuous Approximation Model. International Journal of Production Economics, 188, 167-184.

[6] Katsumi, M., & Fox, M. (2018). Ontologies for Transportation Research: A Survey. Transportation Research Part C: Emerging Technologies, 89, 53-82.

[7] Jain, G. V., Jain, S. S., & Parida, M. (2014). Public Transport Ontology for Passenger Information Retrieval. International Journal of Transportation Engineering, 2(2), 131-144.

[8] Corsar, D., Markovic, M., Edwards, P., & Nelson, J. D. (2015, October). The Transport Disruption Ontology. In International Semantic Web Conference (pp. 329-336). Springer, Cham.

[9] Bermejo, A. J., Villadangos, J., Astrain, J. J., & Cordoba, A. (2013). Ontology Based Road Traffic Management. In Intelligent Distributed Computing VI (pp. 103-108). Springer, Berlin, Heidelberg.

[10] Herzog, A., Jacobi, D., & Buchmann, A. (2008, September). A3ME-an Agent-based Middleware Approach for Mixed Mode Environments. In Mobile Ubiquitous Computing, Systems, Services and Technologies, 2008. UBICOMM'08. The Second International Conference on (pp. 191-196). IEEE.

[11] Benvenuti, F., Diamantini, C., Potena, D., & Storti, E. (2017). An Ontology-based Framework to Support Performance Monitoring in Public Transport Systems. Transportation Research Part C: Emerging Technologies, 81, 188-208.

[12] Anand, N., Yang, M., Van Duin, J. H. R., & Tavasszy, L. (2012). GenCLOn: An Ontology for City Logistics. Expert Systems with Applications, 39(15), 11944-11960.

[13] Bouhana, A., Zidi, A., Fekih, A., Chabchoub, H., & Abed, M. (2015). An Ontology-based CBR Approach for Personalized Itinerary Search Systems for Sustainable Urban Freight Tansport. Expert Systems with Applications, 42(7), 3724-3741.

[14] Malloy, D. P. (2015). Connecticut’s Bold Vision for a Transportation Future. Available from https://www.ct.gov/dot/lib/dot/documents/dcommunications/ctdot_30_yr.pdf

[15] De Oliveira, K. M., Bacha, F., Mnasser, H., & Abed, M. (2013). Transportation Ontology Definition and Application for the Content Personalization of User Interfaces. *Expert Systems with Applications*, *40*(8), 3145-3159.

[16] MaineDOT. (2010). Statewide Long-Range Transportation Plan 2008 – 2030. Available from <https://www.maine.gov/mdot/publications/docs/plansreports/connectingmainefulldocument.pdf>

[17] City of Portland. (2018). 2035 Comprehensive Plan. Available from https://www.portlandoregon.gov/bps/2035-comp-plan.pdf

[18] Grimmer, J. (2010). A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases. Political Analysis, 18(1), 1-35.

[19] Khanzadeh, M., Chowdhury, S., Marufuzzaman, M., Tschopp, M. A., & Bian, L. (2018). Porosity Prediction: Supervised-learning of Thermal History for Direct Laser Deposition. Journal of manufacturing systems, 47, 69-82.

[20] Gillis, N. (2014). The Why and How of Nonnegative Matrix Factorization. Regularization, Optimization, Kernels, and Support Vector Machines, 12(257).

[21] Chowdhury, S., Khanzadeh, M., Akula, R., Zhang, F., Zhang, S., Medal, H., Marufuzzaman, M., & Bian, L. (2017). Botnet Detection Using Graph-Based Feature Clustering. Journal of Big Data, 4(1), 14.

[22] Regional Transportation Authority (2018). Invest in Transit 2018-2023 Regional Strategic Plan. Available from https://www.rtachicago.org/sites/default/files/documents/Invest%20in%20Transit%202018-2023%20Regional%20Strategic%20Plan%20ScreenView%20for%20Web.pdf

[23] Compares, T. F. (2003). Measuring Transportation: Traffic, Mobility and Accessibility. ITE journal, 73(10), 28-52.

[24] GDOT. (2007). 2005-2035 Georgia State-wide Transportation Plan. Available from http://www.dot.ga.gov/InvestSmart/Documents/SSTP/SWTPFinalReport.pdf

[25] San Francisco Municipal Transportation Agency. (2018). Strategic Plan. Available from https://www.sfmta.com/sfmta-strategic-plan

[26] NMDOT. (2015). The New Mexico 2040 Plan: NMDOT’s Long Range, Multi-Modal Transportation Plan. Available from http://dot.state.nm.us/content/dam/nmdot/planning/NM_2040_Plan.pdf