Transportation Research Part C 89 (2018) 53–82

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Review

Ontologies for transportation research: A survey

Megan Katsumi*, Mark Fox

University of Toronto 5 King's College Road Toronto, Ontario M5S 3G8, Canada

ARTICLE INFO

Keywords:
Transportation ontology
Knowledge representation
Reasoning
Interoperability
Formal logic
Semantic Web

ABSTRACT

Transportation research relies heavily on a variety of data. From sensors to surveys, data supports day-to-day operations as well as long-term planning and decision-making. The challenges that arise due to the volume and variety of data that are found in transportation research can be effectively addressed by ontologies. This opportunity has already been recognized – there are a number of existing transportation ontologies, however the relationship between them is unclear. The goal of this work is to provide an overview of the opportunities for ontologies in transportation research and operation, and to present a survey of existing transportation ontologies to serve two purposes: (1) to provide a resource for the transportation research community to aid in understanding (and potentially selecting between) existing transportation ontologies; and (2) to identify future work for the development of transportation ontologies, by identifying areas that may be lacking.

1. Introduction

Ontologies provide a means of knowledge representation; they capture a domain of interest by formally defining the relevant concepts in the domain, and the relationships between these concepts. The transportation domain stands to benefit considerably from the application of ontologies. Transportation data are varied and complex; they come from different organizations, sensors, surveys, and other means of data collection. The development of "Smart Cities", as well as more traditional applications such as research and planning face the challenge of how to integrate data from multiple, unrelated sources where the semantics of the data are imprecise, ambiguous and overlapping. This is especially true in a world where more and more data being used is being openly published on the Internet. Early successes in data "mash-ups" relied upon an independence assumption, where unrelated data sources were linked based solely on geospatial location, or a unique identifier for a person or organization. More sophisticated analytics projects that require the combination of datasets with overlapping semantics entail a significant effort to transform data into something useable. It has become increasingly clear that achieving interoperability among separate datasets requires an attention to the semantics of the underlying attributes and their values. An initial literature review reveals a number of existing transportation ontologies, however the relationship between them is unclear.

The goal of this work is to present a survey of transportation ontologies to serve two purposes: (1) to provide a resource for the transportation research community, to aid in understanding (and potentially selecting between) existing transportation ontologies; and (2) to identify future work for the development of transportation ontologies, by identifying areas that may be lacking. This overview is also intended to serve as an introduction to the application of ontologies and the opportunities they provide for transportation research and management. While there also exist ontologies that define concepts that are foundational within the transportation domain such as time and space, as well as closely related concepts in urban studies such as population and land use, such ontologies are out of the scope of this survey. Here, the focus is restricted to ontologies designed to capture transportation: its

* Corresponding author.
E-mail address: katsumi@mie.utoronto.ca (M. Katsumi).

https://doi.org/10.1016/j.trc.2018.01.023
Received 13 September 2017; Received in revised form 24 January 2018; Accepted 24 January 2018

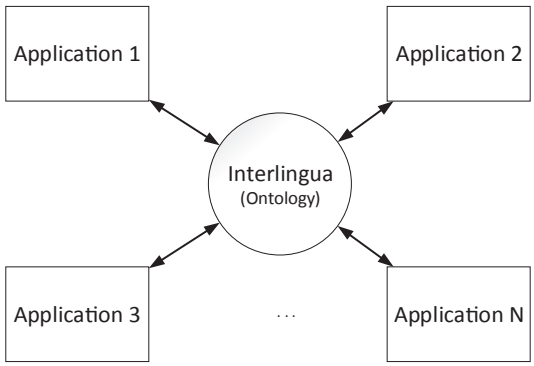

Fig. 1. Ontology as an Interlingua.

systems, infrastructure, and activities.

We begin by providing some background on the ways in which ontologies may be beneficial for the transportation research community. We then consider the cumulative scope of the identified ontologies; this serves as the basis for a comparison amongst the ontologies and also provides an indication of the scope of concepts that currently define the transportation domain. Following this we present an overview of the criteria that the ontologies will be assessed against for a more detailed comparison, before presenting and reflecting on the results.

2. Background

To motivate the relevance of this survey, a brief introduction to ontologies and the Semantic Web is necessary. We then discuss the value of ontologies, specifically for transportation research.

2.1. Ontologies: What & Why?

The most widely used definition of an ontology was presented by (Gruber, 1993) and states simply: *"An ontology is an explicit specification of a conceptualization"*. In the literature, there exist a range of artefacts that are identified as ontologies, from basic glossaries to formal ontologies specified in highly expressive logics. For the purposes of this discussion the term ontology will refer only to formal ontologies with explicitly defined semantics. This semantics can be transcribed into machine-readable languages; as a result, ontologies are able to support the various knowledge management and reasoning services described below.

Integration: precisely defined concepts provide a sort of *interlingua* (illustrated in Fig. 1) that applications can use to share and exchange information. For example, in one system a vehicle may refer to a personal, household vehicle, whereas in a transit application a vehicle may refer to a bus or streetcar. The ontology provides a common language to both distinguish and relate these concepts between the applications. This integration serves to support what is referred to as *semantic interoperability* between applications. Semantic interoperability exists between two systems when they are capable of automatically and unambiguously exchanging meaningful information.

Data validation: when data are represented with an ontology, they can be easily validated against the definitions. For example, an ontology might define a tour as something that's performed by some agent. It might also specify that all tours must start and end at the same location. In this case, if a system observes some data where the start and the end of a trip are not the same location, this data will be inconsistent with our definition of a trip and can easily, automatically be recognized as such.

Inference: we can infer new information about some data based on the domain knowledge that is encoded in the ontology. As a very simple example, say the data tells us that "AC123" is a Transit Vehicle. The definitions in a transportation ontology may specify that all Transit Vehicles are Vehicles, so it can be inferred that "AC123" is a Vehicle; this sort of inference is useful in ensuring the completeness of query results. More complex sorts of inference are also possible; for example, the definitions in the ontology state that vehicles change locations through some occurrence of travel. Given some observation of a vehicle located at one place and then located at another place a short time later, the ontology can be used to infer that some travel activity has taken place. With some additional definitions and a layout of the road network, inference about which possible route(s) was taken may also be supported.

In a given application, an ontology may be employed for one or many of these abilities. In theory, a single ontology can support all of the described services. However, in practice it is often the case that an ontology is designed with a particular service (or set of services in mind). It is in part for this reason that sometimes there are different ontologies designed to represent the same domain; this survey includes eleven different ontologies that capture the transportation domain in various ways.