

The iCity Ontology: Towards a Data Standard for Transportation Planning

iCity-ORF Research Days Webinar Series

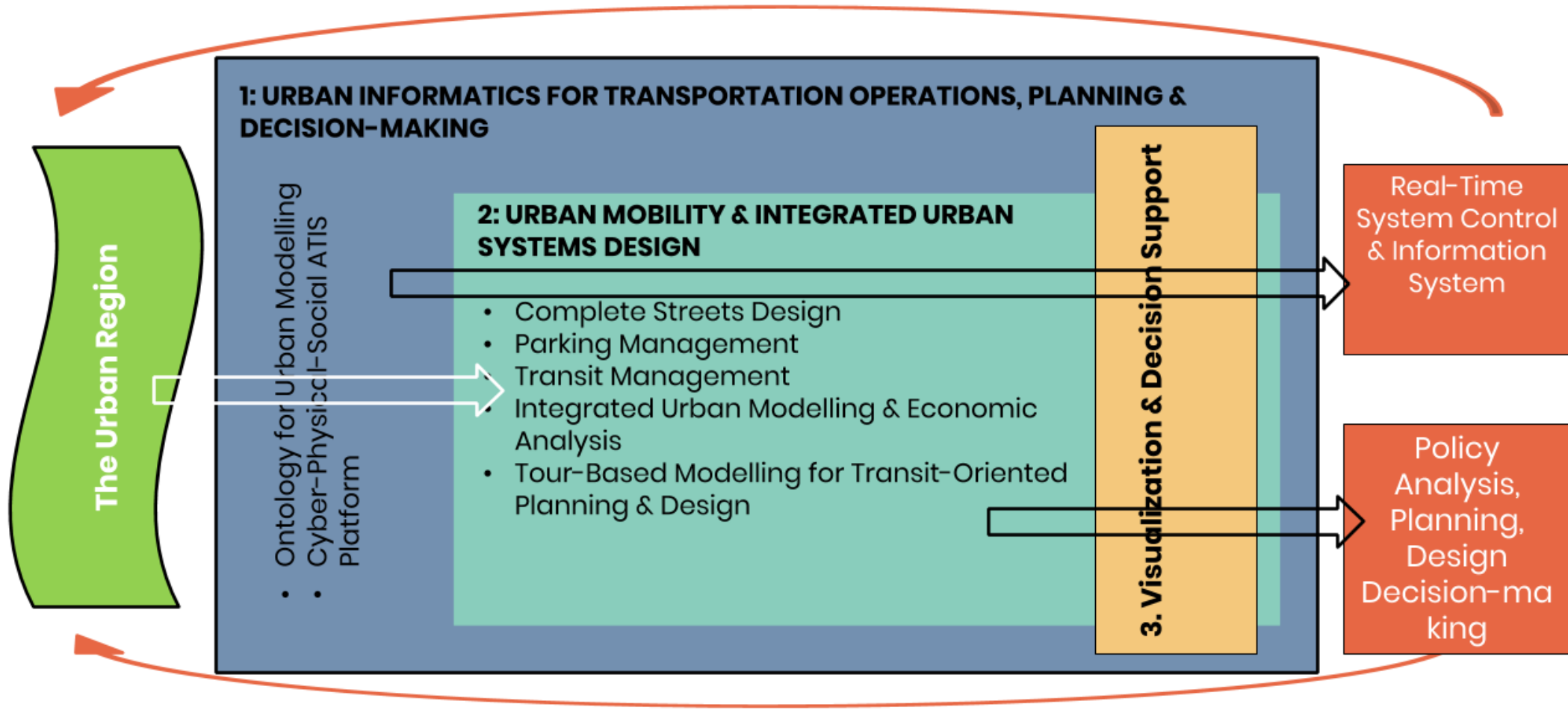
Ontologies & Platforms: June 8, 2020

Megan Katsumi (katsumi@mie.utoronto.ca)

Mark Fox (msf@mie.utoronto.ca)

iCity-ORF

Urban Informatics for Sustainable Metropolitan Growth



iCity Ontology

Our Focus:

- How is it all related?
- Can it be combined?
- How can it be shared?



Wait – what's an ontology?

- A specialized model
 - What are the core concepts and properties that span the domain's data?
 - To what extent can we generalize them in a useful way?
 - What are the key distinctions?
 - Can we formally define these concepts?
- Provides a precise, formal representation that is machine-interpretable
 - More than a reference model (vocabulary) for a domain
 - Supports:
 - Data reuse
 - Data validation
 - Semantic integration
 - Inference

Example: Definition of a Route

- **Simulation Route:** a route represents a possible path of travel. It **begins** and **ends** at some distinct **nodes** in the transportation network and connects the start and end points by **accessing** some set of **arcs**.
- **Wayfinding Service Route:** a route represents a possible path of travel. It **has a start** and an **end node** and some **associated geometry**.

SimulationRoute

$\sqsubseteq \exists \textit{hasStart}. \textit{Node}$

$\sqcap \exists \textit{hasEnd}. \textit{Node}$

$\sqcap \exists \textit{accesses}. \textit{Arc}$

Wayfinding

$\sqsubseteq \exists \textit{hasStart}. \textit{Node}$

$\sqcap \exists \textit{hasEnd}. \textit{Node}$

$\sqcap \exists \textit{hasShape}. \textit{Geometry}$

Data to Knowledge

Semantic Mapping to Ontology

Travel simulation routes

ID	Start	End	Arc	Attr
route1234	N501	N320	A7952	...
...

{ID} → Simulation_Route

{Start} → Node

Simulation_Route hasStart {Start}

...

Wayfinding Route Records

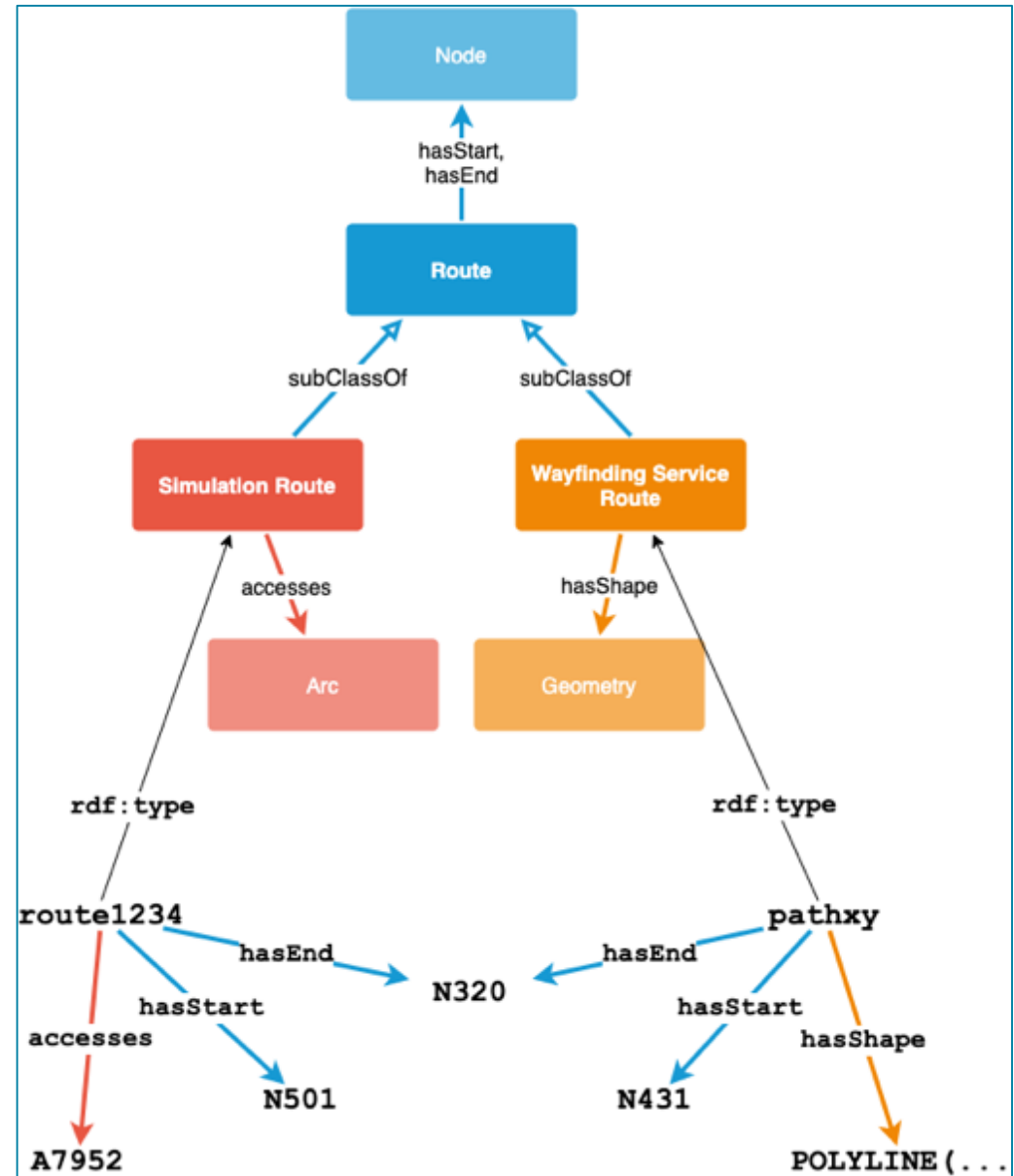
User	Route	o	d	Shape
jsmith	pathxy	N431	N320	POLYLINE (...)
...

Route → Wayfinding_Service_Route

o → Node

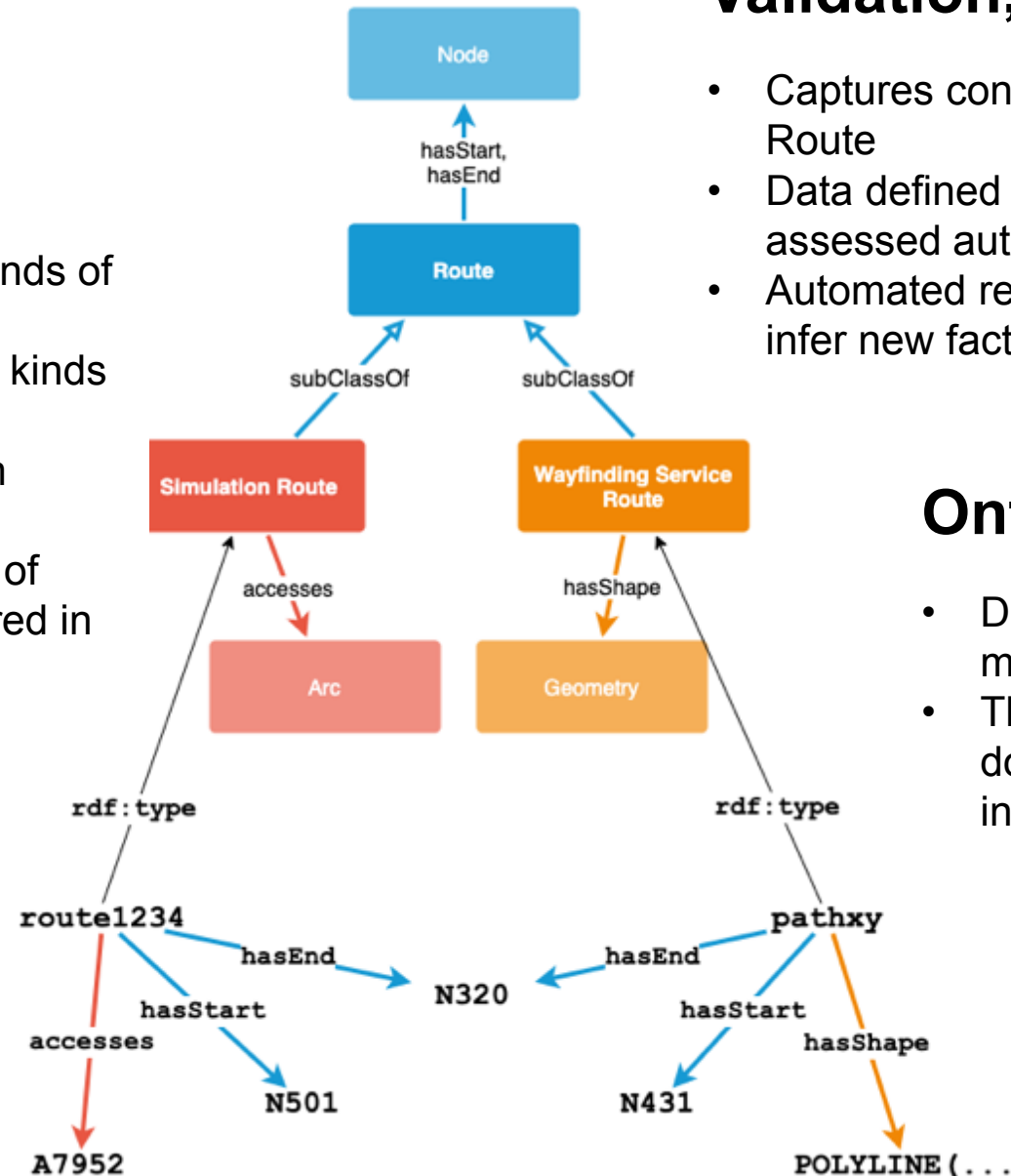
...

Knowledge Graph



Ontology for Integration

- Supports the definition of multiple kinds of route
- Explicitly identifies how the different kinds of routes are related
 - What's common between them
 - What's different between them
- Data sources using either definition of route can be understood and captured in an integrated knowledge base



Ontology for Data Validation, Inference

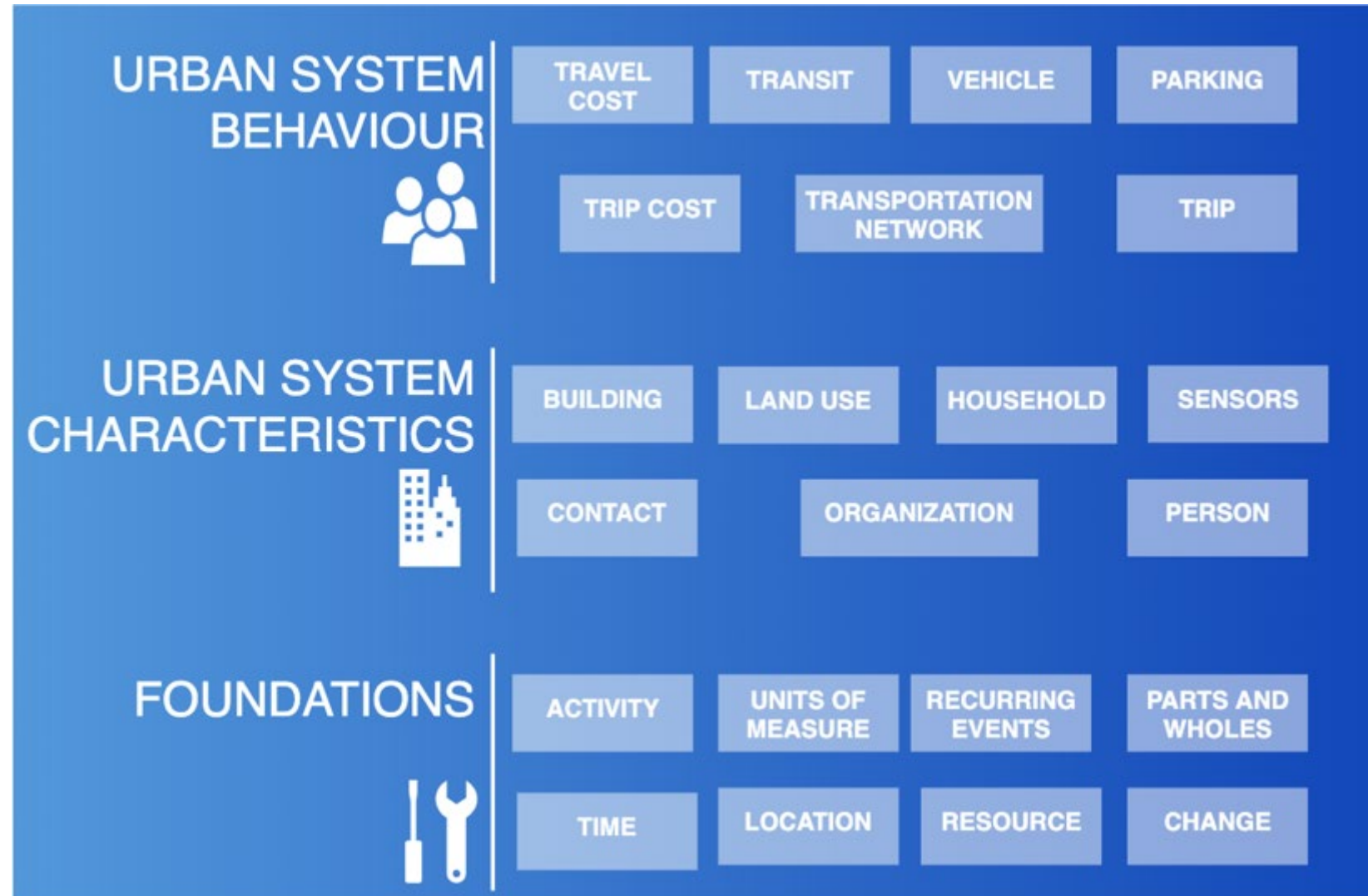
- Captures constraints on each type of Route
- Data defined with the ontology can be assessed automatically
- Automated reasoners can be applied to infer new facts about the data

Ontology for Data Reuse

- Data defined with the ontology has meaning embedded
- The representation serves as documentation for how it should be interpreted

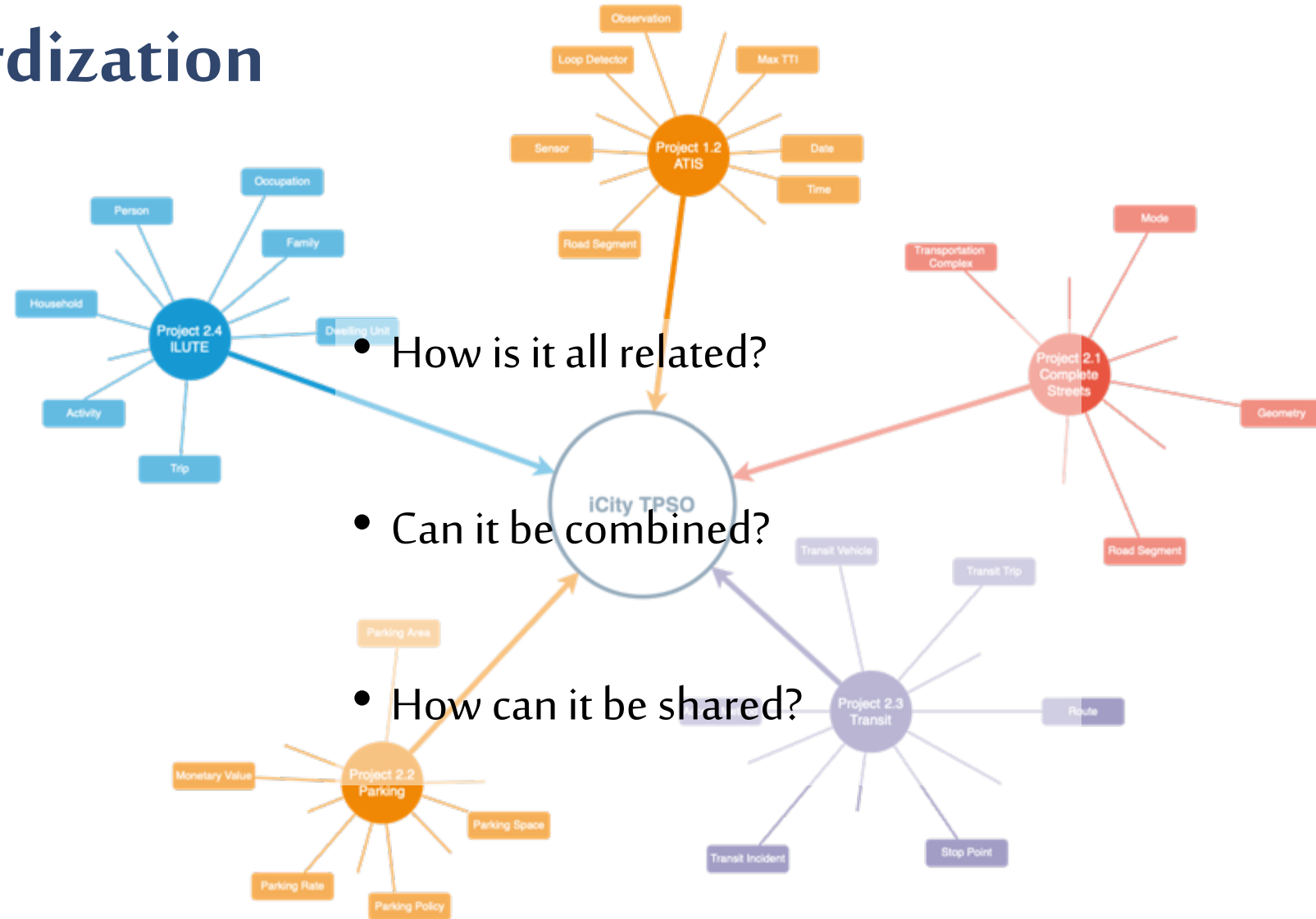
iCity TPSO

- iCity Transportation Planning Suite of Ontologies
- Developed & implemented for transportation planning activities in iCity-ORF project



iCity TPSO: Standardization

- **Observation:** this is a general challenge for transportation planning
- Led to the creation of multi-part ISO standards projects for a city data model



Transportation Planning: The Data Problem

- Data is siloed: acquired and generated data is expensive, but often not reused
- Multitude of transportation planning tools are in use by researchers and cities
- No easy way to compare results as each has their own unique data models

A standard for this data is needed!



UrbanSim



National Center
for Sustainable
Transportation



What about existing standards?

- Scope: existing standards overlap with, but **don't cover the domain of transportation planning**
- Encoding: traditional standards are **subject to ambiguity**, despite detailed definitions

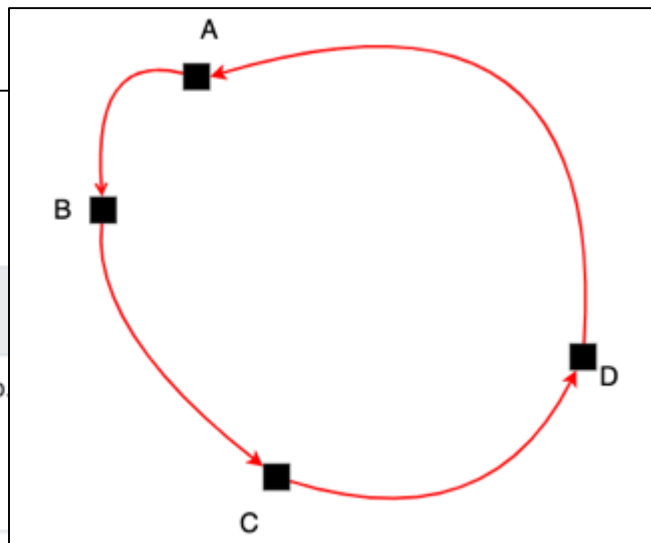


Example: GTFS

stop_times.txt

File: **Required**

Field Name	Type	Required	Description
trip_id	ID referencing trips. trip_id	Required	Identifies a trip
arrival	Time	Conditionally	Arrival time at a specific stop for a specific trip on a route. If there



What about loop routes?

stop_id	ID referencing stops. stop_id	Required	Identifies the serviced stop. All stops serviced during a trip must have a record in stop_times.txt . Referenced locations must be stops, not stations or station entrances. A stop may be serviced multiple times in the same trip, and multiple trips and routes may service the same stop.
stop_sequence	Non-negative integer	Required	Order of stops for a particular trip. The values must increase along the trip but do not need to be consecutive. <i>Example: The first location on the trip could have a stop_sequence=1, the second location on the trip could have a stop_sequence=23, the third location could have a stop_sequence=40, and so on.</i>

- GTFS: General Transit Feed Specification¹
- Common format for public transit data (schedules, locations,...)
- Highly successful, widely adopted

¹ <https://developers.google.com/transit/gtfs>

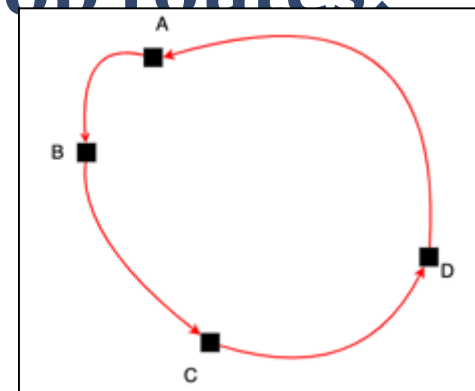
Example: How to define stop times for loop routes?

Option A:

Trip_id	arrival_time	departure_time	stop_id	stop_sequence	stop_headsign
trip_1	06:10:00	06:10:00	stop_A	1	"outbound"
trip_1	06:15:00	06:15:00	stop_B	2	"outbound"
trip_1	06:20:00	06:20:00	stop_C	3	"inbound"
trip_1	06:25:00	06:25:00	stop_D	4	"inbound"

Option B:

Trip_id	arrival_time	departure_time	stop_id	stop_sequence	stop_headsign
trip_1	06:10:00	06:10:00	stop_A	1	"outbound"
trip_1	06:15:00	06:15:00	stop_B	2	"outbound"
trip_1	06:20:00	06:20:00	stop_C	3	"inbound"
trip_1	06:25:00	06:25:00	stop_D	4	"inbound"
trip_1	06:30:00	06:30:00	Stop_A	5	""



- Not explicitly addressed in the reference
- Identified as a special case in the GTFS best practices document
- *Recommendation*: Option B

Ambiguity in Standards Specifications

Traditional approach

- Detailed documentation
- Modelling languages that focus on the data's structure but not its semantics
- Meaning is grounded in natural language



Challenges

- Natural language inherently ambiguous
- Need for supplementary material to resolve individual issues
 - Examples, best practices,...
 - But can't predict or detect everything
- Differences in interpretation lead to differences in adoption
 - Impacts the standard's effectiveness

An ontology-based standard for transportation planning

Our proposal: Ontologies (the iCity TPSO in particular) provide a way to address:

- a major challenge for transportation planning,
and
 - a limitation of traditional approaches to standards specification
-
- Has a **unique** interpretation:
 - Explicit, unambiguous encoding
 - Incorrect and correct interpretations may be automatically identified
 - Added benefits:
 - Works with different tools and data formats
 - Supports a dynamic domain: core concepts are easily extended
 - May be implemented for other applications (e.g. reasoning)

ISO/IEC JTC1 WG11 Smart Cities

City Data Model NP5087

City Data Model Standards Projects

(NP5087-3,

SERVICE
LEVEL

City Service-Level Ontologies

This level is comprised of multiple standards, including – the first such standard – a standard for transportation planning. Each standard at this level includes ontologies to cover data particular to a city service domain. The data modelled at this level may be consumed by a



This level is comprised of multiple standards, including – the first such standard – a standard for transportation planning. Each standard at this level includes ontologies to cover data

(NP5087-2)

CITY
LEVEL

City-Level Ontologies

These ontologies cover concepts that are specific to the city domain, but generic in the sense that they represent data that could be expected to be both generated and consumed by many city services.



(NP5087-1)

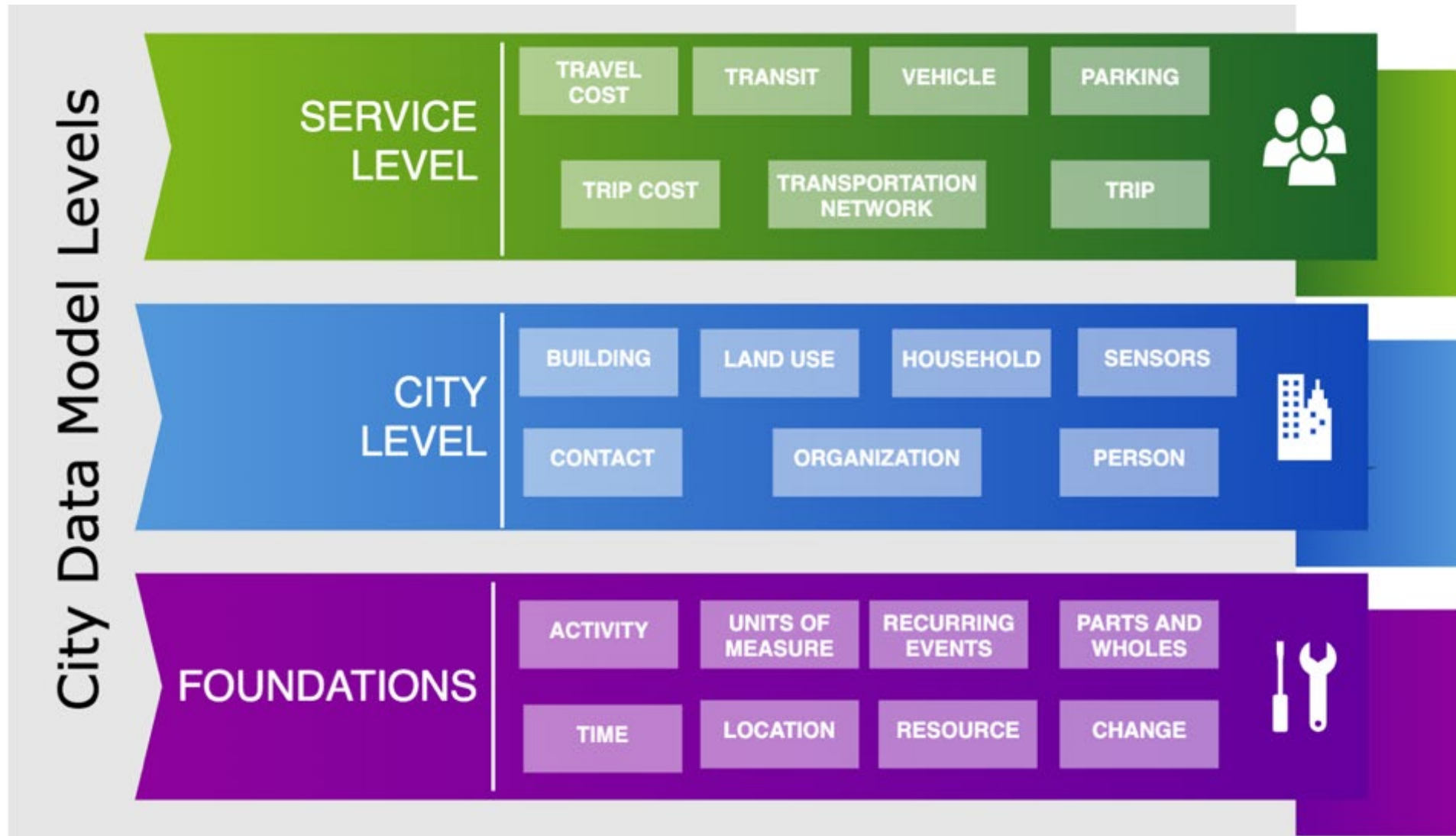
FOUNDATIONS

Foundational Ontologies

These ontologies define the fundamental, generic concepts that are necessary to formulate an accurate definition of the domain. They provide a reusable foundation for the development of other ontologies in the transportation domain, thus ensuring interoperability and consistency in the representation of key concepts such as time and location.



City Data Model: Transportation Planning (NP5087-3)



Standards Collaboration

- Another standard, another silo?
 - How can we avoid this?
- There is a need to collaborate with other groups in order to understand how definitions of overlapping concepts are related
- We created a **Global Collaboratory** to support the alignment of the city data model standards with other standards efforts

The City Data Model Global Collaboratory

<http://citydata.utoronto.ca>²

- Develop a global consensus on the City Data Model
 - Identify concepts and definitions to be included
 - Align related concepts across standards
- Tasks supported:
 - Browse and review content
 - Comment on existing content and suggest changes or revisions
 - Propose terms and definitions
 - Submit use cases to explain/justify terms and definitions

The screenshot shows the website for the City Data Model Project, a global collaboratory. The page has a light blue header with the title "CITY DATA MODEL PROJECT" and subtitle "A GLOBAL COLLABORATORY". On the left, there is a navigation menu with links for "Main page", "About", "Use Case Listing", "Class Listing", "Object Property Listing", "Data Property Listing", "Recent changes", and "Help about MediaWiki". Below this is a "Tools" section with links for "What links here", "Related changes", "Upload file", "Special pages", "Printable version", "Permanent link", and "Page information". On the right, there is a "Contents [hide]" section with a table of contents listing: "1 Welcome to the City Data Model Collaboration Wiki", "1.1 Introduction", "1.1.1 Beta Disclaimer", "1.2 Getting started", "1.3 How it works", and "1.4 The Review Process". Below the table of contents is the "Introduction [edit]" section, which begins with the text: "This website is intended to foster international collaboration... will feed into the various city data standards development... A common data model enables city software applications..."

² To be: citydatastandard.org


The Proposal Process: New terms

- Anyone can propose a new term.
- Once a term has been proposed, it becomes open for the specification of definitions from the community.
- Terms may be independent, or proposed as specializations of existing terms (i.e. sub-classes or sub-properties)

The screenshot shows a MediaWiki page for the term "Route". On the left is a sidebar with navigation links: Main page, About, Use Case Listing, Class Listing, Object Property Listing, Data Property Listing, Recent changes, Help about MediaWiki, Tools, What links here, Related changes, Upload file, Special pages, Printable version, Permanent link, Page information, Cite this page, and Browse properties. The main content area has a title "Route" with a "Contents [show]" button. Below the title is a "Class [edit]" section with the text "Route". This is followed by a "Class Description [edit]" section with the text "A Route represents a possible path of travel in one or more transp". Below that is a "Term status [edit]" section with the text "Pending Approval". The next section is "Definition [edit]". The final section is "Subclasses [edit]" with a bullet point for "Tpso:Route" and a "submit" button at the bottom.

The Proposal Process: New definitions

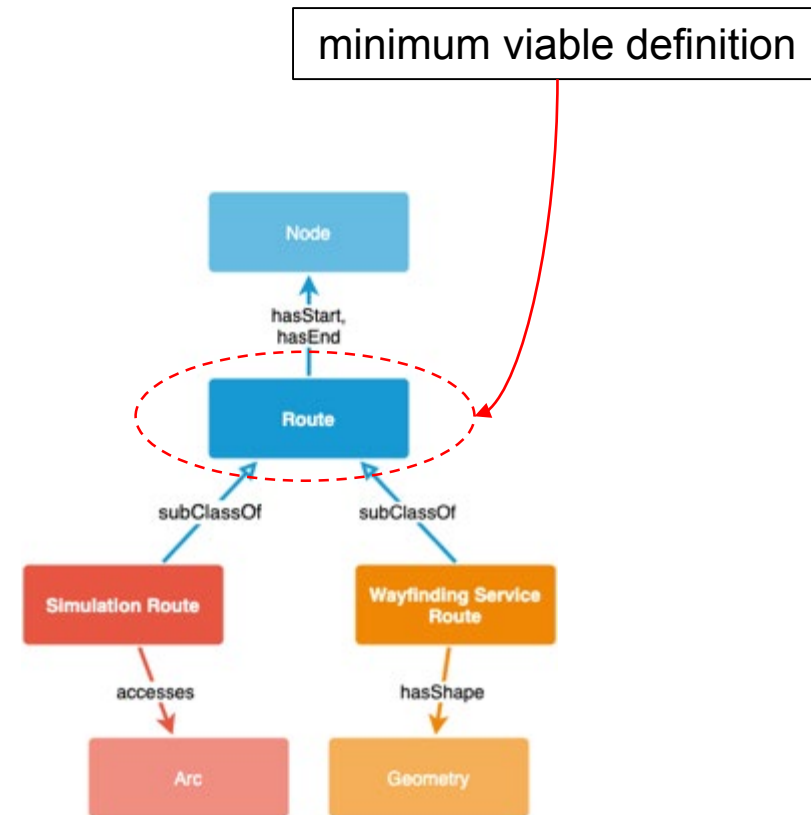
- Proposed definitions for a term must be specified in a *formal* language (Description Logic or UML)
- Proposed definitions must be accompanied by a use case
- Subject to community review
- **Multiple definitions allowed and expected**
 - Do not need to agree on one
 - Once discussion of definitions has settled, administrators will review all proposed definitions and identify the **minimum viable definition** for the term.



TpsO:Route	
Subclass Of	Route
Namespace (context for definition)	http://ontology.eil.utoronto.ca/cdm/Transportation
Description (what distinguishes this sense of the term?)	<p>A Route describes a possible path of travel through a transportation network that it accesses (i.e. travel from the start node of first arc that is accessed by the route to the end node of last arc that is accessed by the route).</p> <p>Routes may be decomposed into smaller sections. A location is associated with the arcs as opposed to the nodes.</p>
Required by Use Case(s) (why is this specialized definition needed?)	Routes for Transportation Planning Travel Demand
Formal Definition (UML and DL)	<p>subClassOf (accessesArc only TpsO:ArcPD) and</p> <p>subClassOf hasSubRoute only tpso:Route</p> <p>subClassOf routeBegins only TpsO:NodePD</p> <p>subClassOf routeEnds only TpsO:NodePD</p>
Status	Pending Approval

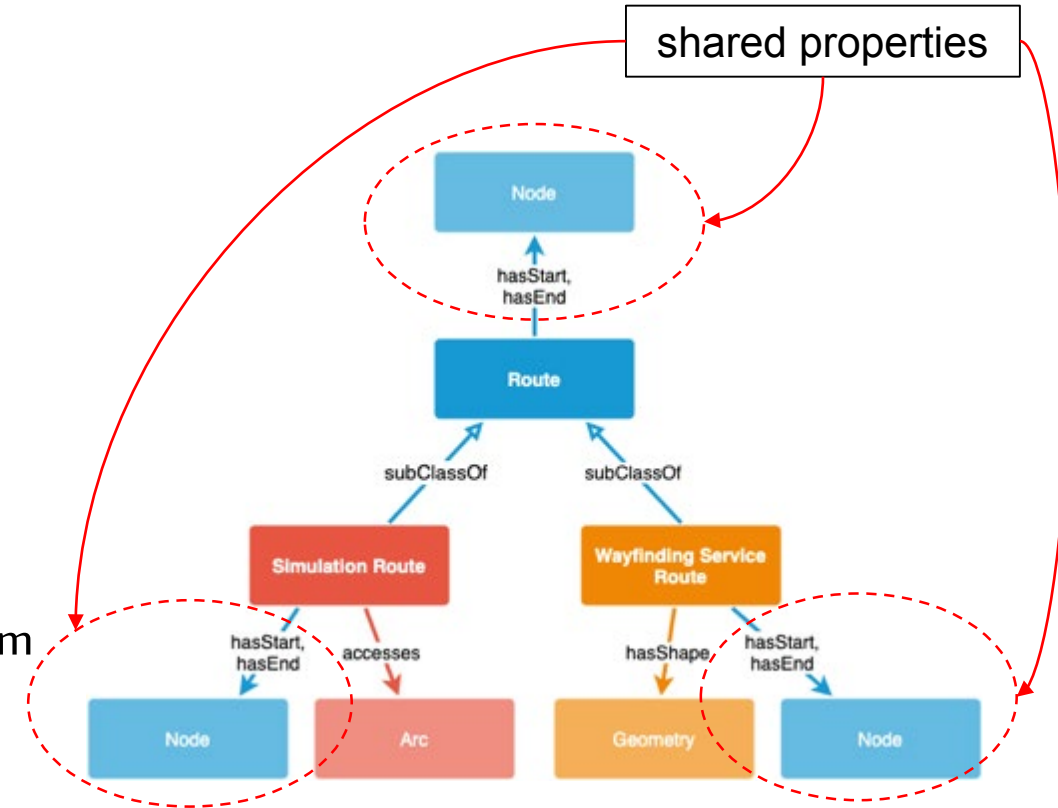
The Proposal Process: Minimum Viable Definitions

- Identified relative to a set of definitions for the same term
- The *minimum* semantics required for a term
 - Shown to be shared between each definition in the set



Minimum Viable Definitions

- Clearly identifies the shared meaning that all stakeholders have in common regarding a particular term
 - Definition identifies the set of properties shared amongst distinct definitions
 - Subset of shared terminology
 - Distinct definitions identified as subclasses of the minimum viable definition



Participants welcome!

<http://citydata.utoronto.ca>²

[Log in](#) [Request account](#)

CITY DATA MODEL PROJECT A GLOBAL COLLABORATORY

[Read](#)

[View source](#)

[View history](#)

Search City Data Model Project Collabora

[Main page](#)
[About](#)
[Use Case Listing](#)
[Class Listing](#)
[Object Property Listing](#)
[Data Property Listing](#)
[Recent changes](#)
[Help about MediaWiki](#)

[Tools](#)

[What links here](#)
[Related changes](#)
[Special pages](#)
[Printable version](#)
[Permanent link](#)
[Page information](#)
[Cite this page](#)
[Browse proper](#)

Contents [hide]

- 1 [Welcome to the City Data Model Collaboration Wiki](#)
 - 1.1 [Introduction](#)
 - 1.1.1 [Beta Disclaimer](#)
 - 1.2 [Getting started](#)
 - 1.3 [How it works](#)
 - 1.4 [The Review Process](#)

Introduction

This website is intended to foster international collaboration between stakeholders and related standards groups on a common [City Data Model](#). The results of this effort will feed into the various city data standards development projects being undertaken by various Standards Development Organizations.

A common data model enables city software applications to share information, plan, coordinate, and execute city tasks, and support decision making within and across cities. It is an unambiguous representation of information and knowledge commonly shared across city services. This requires a clear understanding of the terms used in defining the data, as well as how they relate to one another. Beyond syntactic integration (e.g. common data types and protocols), this

² To be: citydatastandard.org