General City Digital Twin  
User Manual

**Anderson Wong**

Urban Data Research Centre (urbandatacentre.ca)  
School of Cities (schoolofcities.ca)  
University of Toronto

22 April 2025

# Table of Contents

[Table of Contents 1](#_Toc197551645)

[Introduction 2](#_Toc197551646)

[Accessing the Dashboard 2](#_Toc197551647)

[Updating the Graph Database 3](#_Toc197551648)

[Representation of City Data 4](#_Toc197551649)

[Administrative Areas 4](#_Toc197551650)

[Address 5](#_Toc197551651)

[Geospatial Location 6](#_Toc197551652)

[Census Profiles and Characteristics 7](#_Toc197551653)

[Toronto Police Service Crime Data 8](#_Toc197551654)

[OpenStreetMap Tags Mapping 12](#_Toc197551655)

[Complete Community Walking Distance Indicator 15](#_Toc197551656)

[Parks 16](#_Toc197551657)

[Schools 17](#_Toc197551658)

[Kindergartens 18](#_Toc197551659)

[Universities 19](#_Toc197551660)

[Colleges 20](#_Toc197551661)

[Hospitals 21](#_Toc197551662)

[Clinics 23](#_Toc197551663)

[Doctor’s Office 24](#_Toc197551664)

[Pharmacy 25](#_Toc197551665)

[Supermarket 26](#_Toc197551666)

[Greengrocer 27](#_Toc197551667)

[Restaurant 28](#_Toc197551668)

[FastFood 29](#_Toc197551669)

[Representation of Road Data 30](#_Toc197551670)

[Junction 31](#_Toc197551671)

[Road 32](#_Toc197551672)

[RoadLink 32](#_Toc197551673)

[Ontario Road Network Tags Mapping 33](#_Toc197551674)

[Integration of ORN Data in TTL Generation 39](#_Toc197551675)

# Introduction

The City Digital Twin project at the University of Toronto’s Urban Data Research Centre aims to develop a digital twin of the City of Toronto. It builds on existing city data standards such as ISO/IEC 21972:2020, ISO/IEC 5087-1:2024, and ISO/IEC 5087-2:2024, and conforms to Linked Data requirements. The goal of this project is to demonstrate standards-based semantic interoperability of city data. These standard ontologies enable the integration of data from multiple sources, opening up new possibilities for the development of data analysis and visualization tools. Our City Digital Twin is represented as a knowledge graph containing a growing variety of city data (e.g. Canadian Census, Toronto Police Crime Data, City transportation infrastructure, parks, food stores, schools, …). Our City Digital Twin also supports a dashboard that can generate data visualizations using the city data from the graph database.

Our City Digital Twin is general in the sense the ontologies and standards being used to represent city data are city independent. Therefore, the CDT can be applied to any city with modification.

# Accessing the Dashboard

The City Digital Twin Dashboard can be directly accessed using the following link: <http://ec2-3-97-59-180.ca-central-1.compute.amazonaws.com:3001/.>

Alternatively, it can be downloaded from GitHub (<https://github.com/csse-uoft/city-digital-twin/tree/develop>) and run locally. You can run it locally on your computer by completing the following steps:

1. Download the “develop” branch of the City Digital Twin from our GitHub: <https://github.com/csse-uoft/city-digital-twin/tree/develop.> If you downloaded the code as a .zip file, unzip the file before proceeding to the next step.
2. Open the Command Prompt. On Windows 10 or 11, you can do this by pressing the Windows key and typing “cmd”. For Mac OS users, you can open Terminal instead by clicking the Launchpad icon and typing Terminal in the search field.
3. In the Command Prompt, navigate to the unzipped folder that you downloaded from GitHub using the “cd” command (e.g., “cd folder\_name”).
4. Enter and run “cd frontend” to navigate to the “frontend” folder
5. Enter and run “npm install” in the Command Prompt to install the dependencies.
6. Repeat steps 4 and 5 for the “backend” folder (you can return to the parent file directory by using “cd..”)
7. While in the “backend” folder, enter and run “npx nodemon index.js” to start up the backend of the dashboard
8. Open a new Command Prompt window by repeating step 2
9. In this new Command Prompt window, navigate to the unzipped folder by repeating step 3
10. Enter and run “cd frontend” to navigate to the “frontend” folder
11. Enter and run “npm start” to start up the frontend of the dashboard
12. Once the above steps are completed, the dashboard should be up and running at a localhost address specified in the second Command Prompt window. You can enter this localhost address into a web browser to access the dashboard.

# Updating the Graph Database

The graph database for the City Digital Twin project can be accessed here: <http://ec2-3-97-59-180.ca-central-1.compute.amazonaws.com:7200/>. To update the contents of the graph database, please refer to the steps below as a guide and more detailed information about how to use SPARQL to update graphs along with examples can be found here: <https://docs.progress.com/bundle/marklogic-server-develop-with-semantic-graphs-11/page/topics/sparql-update.html.> Official SPARQL Update documentation can also be found here: <https://www.w3.org/TR/sparql12-update/>.

1. If triples need to be removed from the graph database:
   1. If an entire graph needs to be deleted, you can run a “drop graph” query in SPARQL to remove the specified graph: DROP GRAPH <insert\_graph\_IRI>
   2. If specific triples need to be deleted from the graph database:
      1. You can run a “delete data” query if you are deleting individual triples explicitly or:
      2. A “delete where” query if you are deleting patterns of triples
2. If triples need to be added to the graph database:
   1. If an entire RDF file (e.g. a TTL or OWL file) needs to be uploaded to the graph database, you can upload and import the file in GraphDB Workbench using its user interface:
      1. Click on the “Import” button on the left sidebar in GraphDB Workbench
      2. Click on “Upload RDF files” button and upload your RDF file(s)
      3. After the file(s) have been uploaded, click on the “Import” button next to the uploaded file and a popup window will appear
      4. It is highly recommended that you specify the name of the target graph by clicking on “Named graph” under the “Target graphs” section and entering a name for the graph in the textbox
      5. Click on the “Import” button in the bottom right corner of the popup window when you are ready to import the data
      6. An “Imported successfully” message should show if the update was successful
   2. If specific triples need to be added to the graph database:
      1. You can run a “insert data” query if you are adding explicitly defined triples or:
      2. A “insert where” query if you are adding patterns of triples

# Representation of City Data

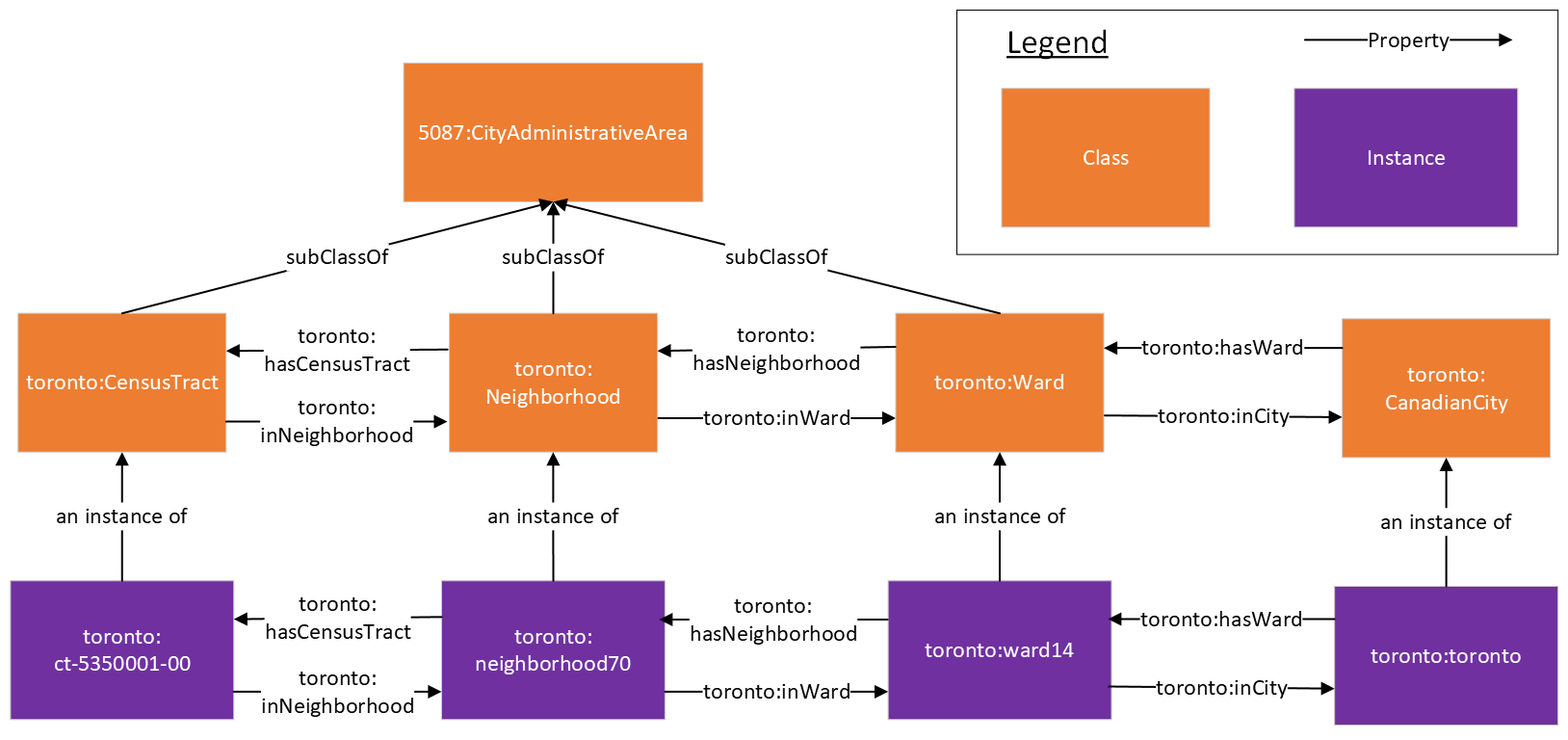
The following sections will provide an overview of how the city data in the graph database is ontologically represented.

The following is a list of namespace prefixes:

* cdt: http://ontology.eil.utoronto.ca/CDT#
* code: https://standards.iso.org/iso-iec/5087/-2/ed-1/en/ontology/Code/
* contact: https://standards.iso.org/iso-iec/5087/-2/ed-1/en/ontology/Contact/
* crime: http://ontology.eil.utoronto.ca/CKGN/Crime#
* foaf: http://xmlns.com/foaf/0.1/
* gcie: http://ontology.eil.utoronto.ca/GCI/Education/GCI-Education.owl#
* gcir: http://ontology.eil.utoronto.ca/GCI/Recreation/GCIRecreation.owl#
* genprop: https://standards.iso.org/iso-iec/5087/-1/ed-1/en/ontology/GenericProperties/
* geo: http://www.opengis.net/ont/geosparql#
* geof: http://www.opengis.net/def/function/geosparql/
* iso21972: http://ontology.eil.utoronto.ca/ISO21972/iso21972#
* iso50871: http://ontology.eil.utoronto.ca/5087/1/SpatialLoc/
* iso5087m: http://ontology.eil.utoronto.ca/5087/1/Mereology/
* loc: https://standards.iso.org/iso-iec/5087/-1/ed-1/en/ontology/SpatialLoc/
* rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#
* rdfs: http://www.w3.org/2000/01/rdf-schema#
* time: http://www.w3.org/2006/time#
* toronto: http://ontology.eil.utoronto.ca/Toronto/Toronto#
* uoft: http://ontology.eil.utoronto.ca/tove/cacensus#
* xsd: http://www.w3.org/2001/XMLSchema#

## Administrative Areas

Administrative areas are geographic regions that are defined for governance and administrative purposes. Examples include areas like neighborhoods or city wards as seen in the below diagram and are expressed as a subclass of the CityAdministrativeArea class from ISO 5087-2. These administrative areas also have properties such as ”hasNeighborhood” or ”inWard” to describe the relationships between each other. These properties are useful for quickly finding, for example, a list of neighborhoods that can be found inside a given ward.

  
An example SPARQL query for finding a list of neighborhoods in ward 1 (Etobicoke North) can be found below:

PREFIX toronto: <http://ontology.eil.utoronto.ca/Toronto/Toronto#>

SELECT ?neighborhood

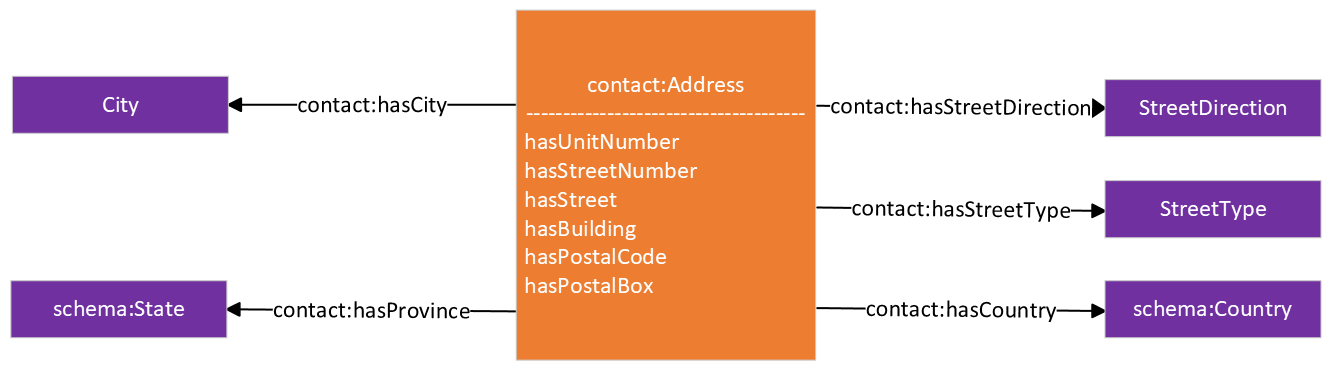
WHERE{

toronto:ward1 toronto:hasNeighborhood ?neighborhood

}

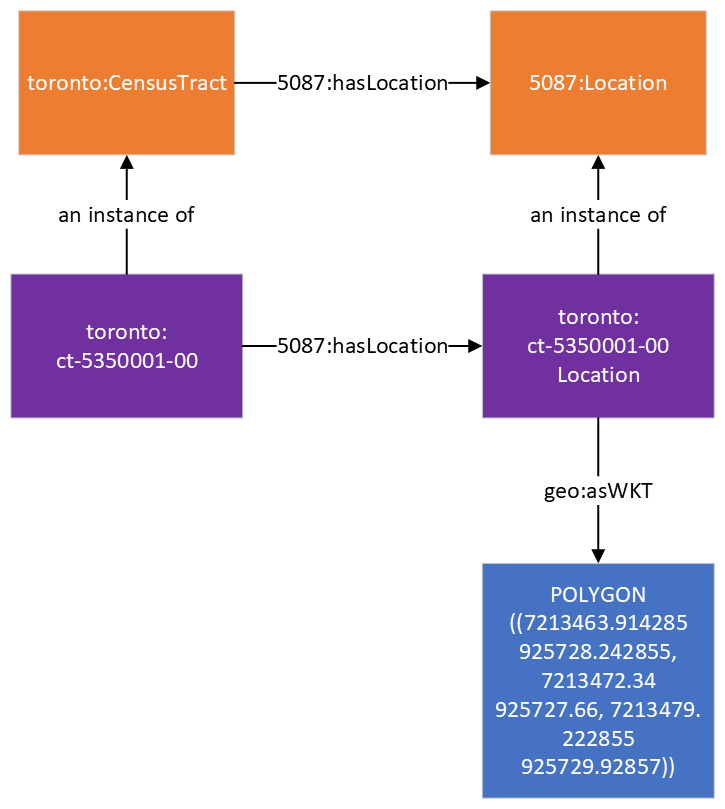
## Address

The address of an entity is represented using the Address representation from the Contact pattern from ISO 5087-2 where an entity is linked to an Address instance using the “hasAddress” property. This Address instance is then linked to individual components of an address using properties such as hasStreetNumber, hasStreet, hasPostalCode, etc.



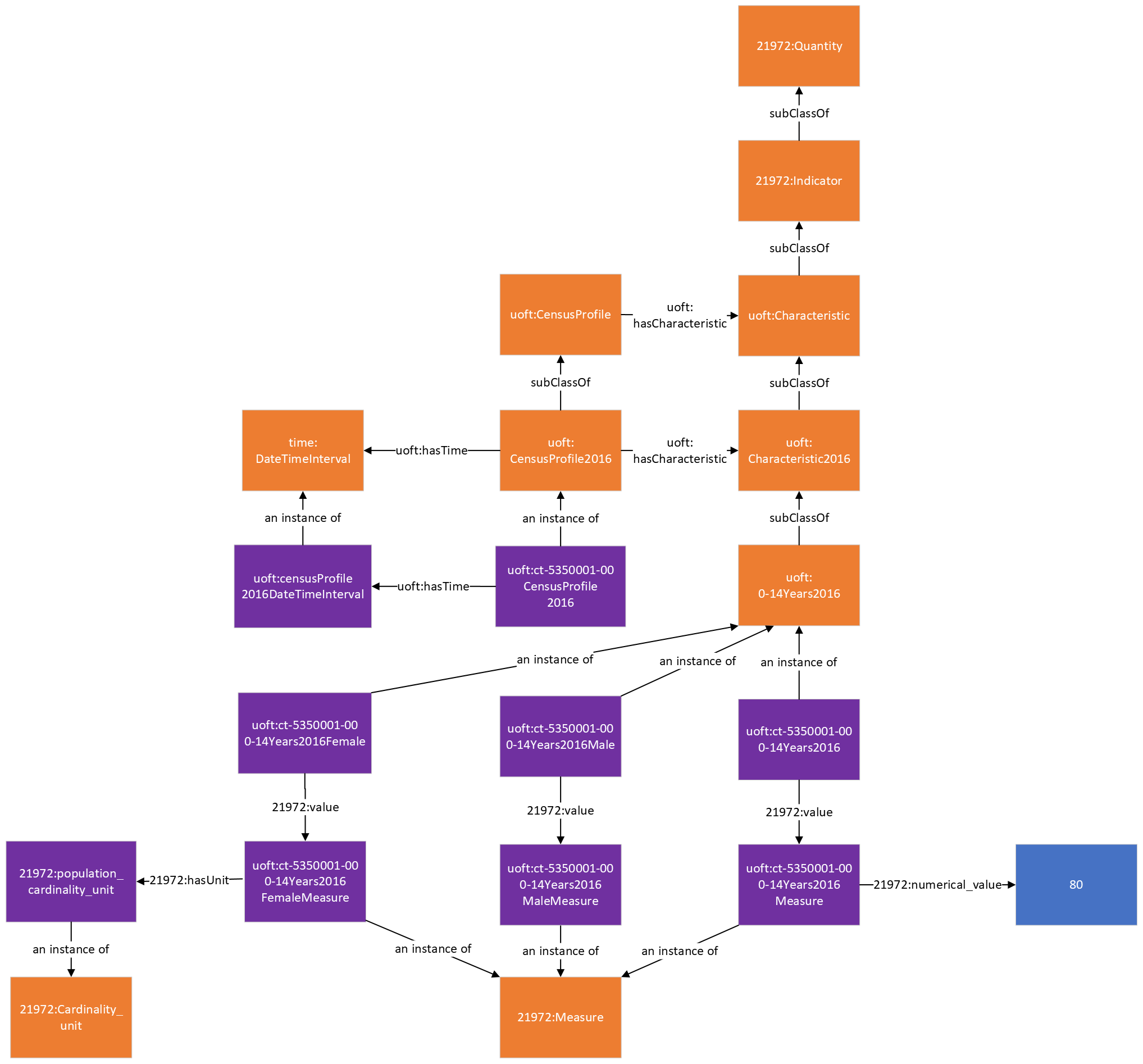
## Geospatial Location

The geospatial location of an entity is represented using the Location pattern from ISO 5087-1 where an entity is linked to a Location class instance using the “hasLocation” property. The location instance is then linked to the WKT polygon coordinates using the “asWKT” property.



## Census Profiles and Characteristics

Census data for geographic areas are organized as datasets called census profiles. Census profiles contain a collection of individual census characteristics, which are a subclass of the Indicator class from ISO 21972 and linked using the hasCharacteristic property. Each characteristic also has a class that is used to represent it, and this class is also a subclass of the Characteristic class. For example, the 0-14Years2016 class in the diagram above represents the characteristic for the number of people who were between the ages 0 to 14 during the 2016 Canadian Census. An instance of this class represents the characteristic for a given geographic area. For example, the “ct-5350001-000-14Years2016” instance in the diagram below represents the characteristic for the number of people in the census tract 5350001.00 who were between the ages 0 to 14 in 2016. These characteristic instances are linked to an instance of the Measure class from ISO 21972 using the ”value” property and this measure instance is linked to the value of the characteristic using the ”numerical\_value” property.



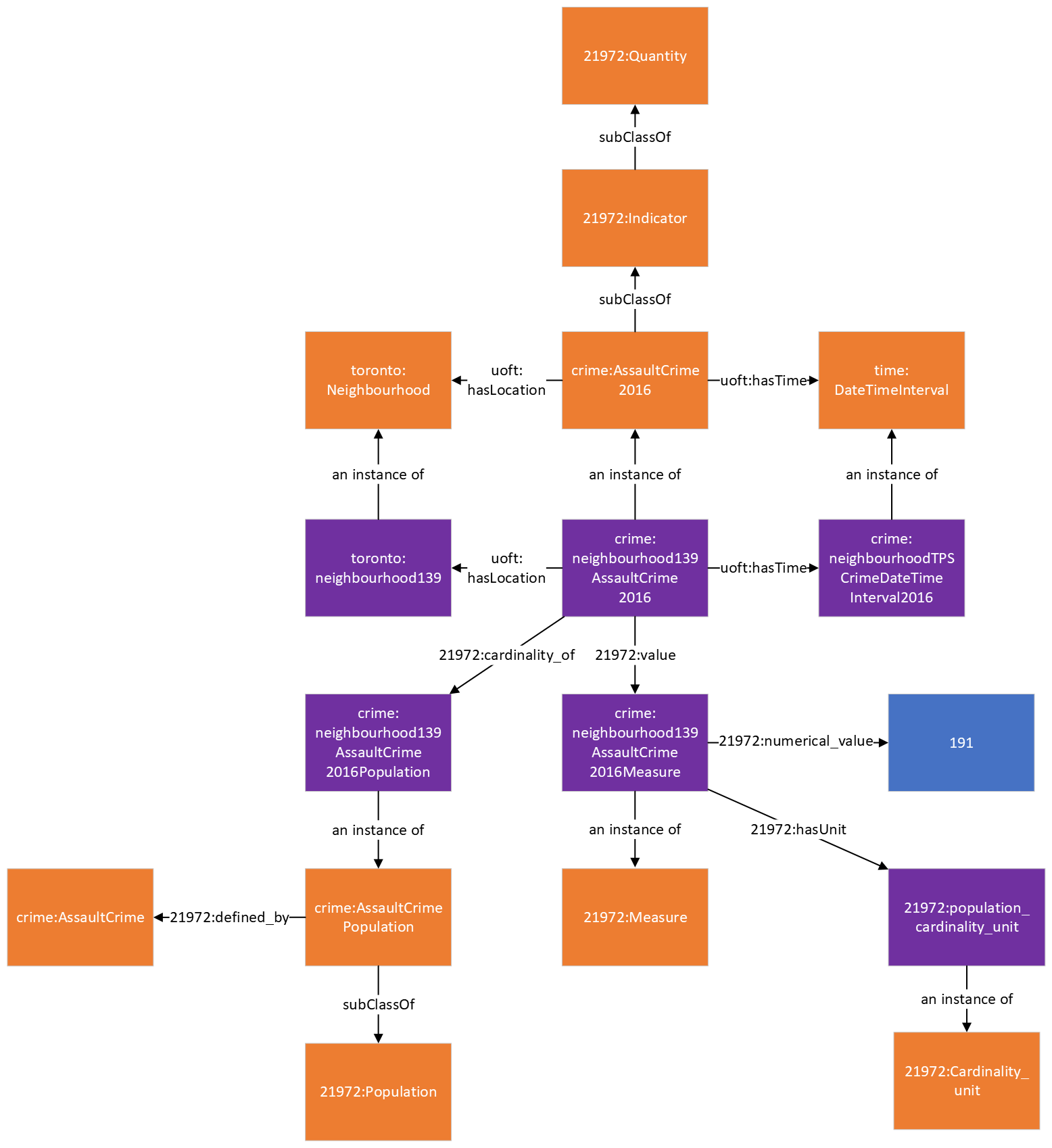
A more in-depth explanation of the Canadian Census ontology can be found in our paper titled “Semantically Interoperable Census Data: Unlocking the Semantics of Census Data Using Ontologies and Linked Data” which can be found here: <https://ijpds.org/article/view/2378#:~:text=Using%20census%20data%20as%20an,data%20points%20from%20large%20datasets>.

## Toronto Police Service Crime Data

This is a basic ontology for representing neighborhood crime indicators from the Toronto Police Service using the Indicator pattern from ISO/IEC 21972. An instance of a crime indicator (e.g. AssualtCrime2016 which measures the number of assault crimes that occurred during the year 2016 in a given location) is defined as a subclass of the Indicator class from ISO/IEC 21972. Crime indicators can be linked to their location using the hasLocation property and its corresponding time period using the hasTime property. The value of a crime indicator is defined as an instance of the Measure class from ISO/IEC 21972 and can be linked to its numerical value using the numerical\_value property. Additionally, crime indicators can be linked to its population instance using the cardinality\_of property from ISO/IEC 21972.

This section provides a brief description of the column labels found in the data spreadsheet published by the Toronto Police Service.

|  |  |
| --- | --- |
| Toronto Police Service Label | Description |
| HoodName,C,35 | Name of the neighborhood |
| HoodID,C,3 | The neighborhood number |
| Assault\_20,N,3,0 | Number of assault crime incidents in 2014 |
| Assault\_1,N,3,0 | Number of assault crime incidents in 2015 |
| Assault\_2,N,3,0 | Number of assault crime incidents in 2016 |
| Assault\_3,N,3,0 | Number of assault crime incidents in 2017 |
| Assault\_4,N,3,0 | Number of assault crime incidents in 2018 |
| Assault\_5,N,3,0 | Number of assault crime incidents in 2019 |
| Assault\_6,N,3,0 | Number of assault crime incidents in 2020 |
| Assault\_7,N,3,0 | Number of assault crime incidents in 2021 |
| AutoTheft\_,N,3,0 | Number of auto theft crime incidents in 2014 |
| AutoThef\_1,N,3,0 | Number of auto theft crime incidents in 2015 |
| AutoThef\_2,N,3,0 | Number of auto theft crime incidents in 2016 |
| AutoThef\_3,N,3,0 | Number of auto theft crime incidents in 2017 |
| AutoThef\_4,N,3,0 | Number of auto theft crime incidents in 2018 |
| AutoThef\_5,N,3,0 | Number of auto theft crime incidents in 2019 |
| AutoThef\_6,N,3,0 | Number of auto theft crime incidents in 2020 |
| AutoThef\_7,N,3,0 | Number of auto theft crime incidents in 2021 |
| BreakAndEn,N,3,0 | Number of break and enter crime incidents in 2014 |
| BreakAnd\_1,N,3,0 | Number of break and enter crime incidents in 2015 |
| BreakAnd\_2,N,3,0 | Number of break and enter crime incidents in 2016 |
| BreakAnd\_3,N,3,0 | Number of break and enter crime incidents in 2017 |
| BreakAnd\_4,N,3,0 | Number of break and enter crime incidents in 2018 |
| BreakAnd\_5,N,3,0 | Number of break and enter crime incidents in 2019 |
| BreakAnd\_6,N,3,0 | Number of break and enter crime incidents in 2020 |
| BreakAnd\_7,N,3,0 | Number of break and enter crime incidents in 2021 |
| Robbery\_20,N,3,0 | Number of robbery crime incidents in 2014 |
| Robbery\_\_1,N,3,0 | Number of robbery crime incidents in 2015 |
| Robbery\_\_2,N,3,0 | Number of robbery crime incidents in 2016 |
| Robbery\_\_3,N,3,0 | Number of robbery crime incidents in 2017 |
| Robbery\_\_4,N,3,0 | Number of robbery crime incidents in 2018 |
| Robbery\_\_5,N,3,0 | Number of robbery crime incidents in 2019 |
| Robbery\_\_6,N,3,0 | Number of robbery crime incidents in 2020 |
| Robbery\_\_7,N,3,0 | Number of robbery crime incidents in 2021 |
| TheftOver\_,N,2,0 | Number of thefts over $5000 crime incidents in 2014 |
| TheftOve\_1,N,2,0 | Number of thefts over $5000 crime incidents in 2015 |
| TheftOve\_2,N,2,0 | Number of thefts over $5000 crime incidents in 2016 |
| TheftOve\_3,N,2,0 | Number of thefts over $5000 crime incidents in 2017 |
| TheftOve\_4,N,2,0 | Number of thefts over $5000 crime incidents in 2018 |
| TheftOve\_5,N,2,0 | Number of thefts over $5000 crime incidents in 2019 |
| TheftOve\_6,N,2,0 | Number of thefts over $5000 crime incidents in 2020 |
| TheftOve\_7,N,2,0 | Number of thefts over $5000 crime incidents in 2021 |
| Homicide\_2,N,1,0 | Number of homicide crime incidents in 2014 |
| Homicide\_1,N,1,0 | Number of homicide crime incidents in 2015 |
| Homicide\_3,N,1,0 | Number of homicide crime incidents in 2016 |
| Homicide\_4,N,1,0 | Number of homicide crime incidents in 2017 |
| Homicide\_5,N,1,0 | Number of homicide crime incidents in 2018 |
| Homicide\_6,N,1,0 | Number of homicide crime incidents in 2019 |
| Homicide\_7,N,1,0 | Number of homicide crime incidents in 2020 |
| Homicide\_8,N,1,0 | Number of homicide crime incidents in 2021 |
| Shootings\_,N,2,0 | Number of shooting crime incidents in 2014 |
| Shooting\_1,N,2,0 | Number of shooting crime incidents in 2015 |
| Shooting\_2,N,2,0 | Number of shooting crime incidents in 2016 |
| Shooting\_3,N,2,0 | Number of shooting crime incidents in 2017 |
| Shooting\_4,N,2,0 | Number of shooting crime incidents in 2018 |
| Shooting\_5,N,2,0 | Number of shooting crime incidents in 2019 |
| Shooting\_6,N,2,0 | Number of shooting crime incidents in 2020 |
| Shooting\_7,N,2,0 | Number of shooting crime incidents in 2021 |
| TheftfromM,N,3,0 | Number of thefts from motor vehicle crime incidents in 2014 |
| Theftfro\_1,N,3,0 | Number of thefts from motor vehicle crime incidents in 2015 |
| Theftfro\_2,N,3,0 | Number of thefts from motor vehicle crime incidents in 2016 |
| Theftfro\_3,N,3,0 | Number of thefts from motor vehicle crime incidents in 2017 |
| Theftfro\_4,N,3,0 | Number of thefts from motor vehicle crime incidents in 2018 |
| Theftfro\_5,N,3,0 | Number of thefts from motor vehicle crime incidents in 2019 |
| Theftfro\_6,N,3,0 | Number of thefts from motor vehicle crime incidents in 2020 |
| Theftfro\_7,N,3,0 | Number of thefts from motor vehicle crime incidents in 2021 |



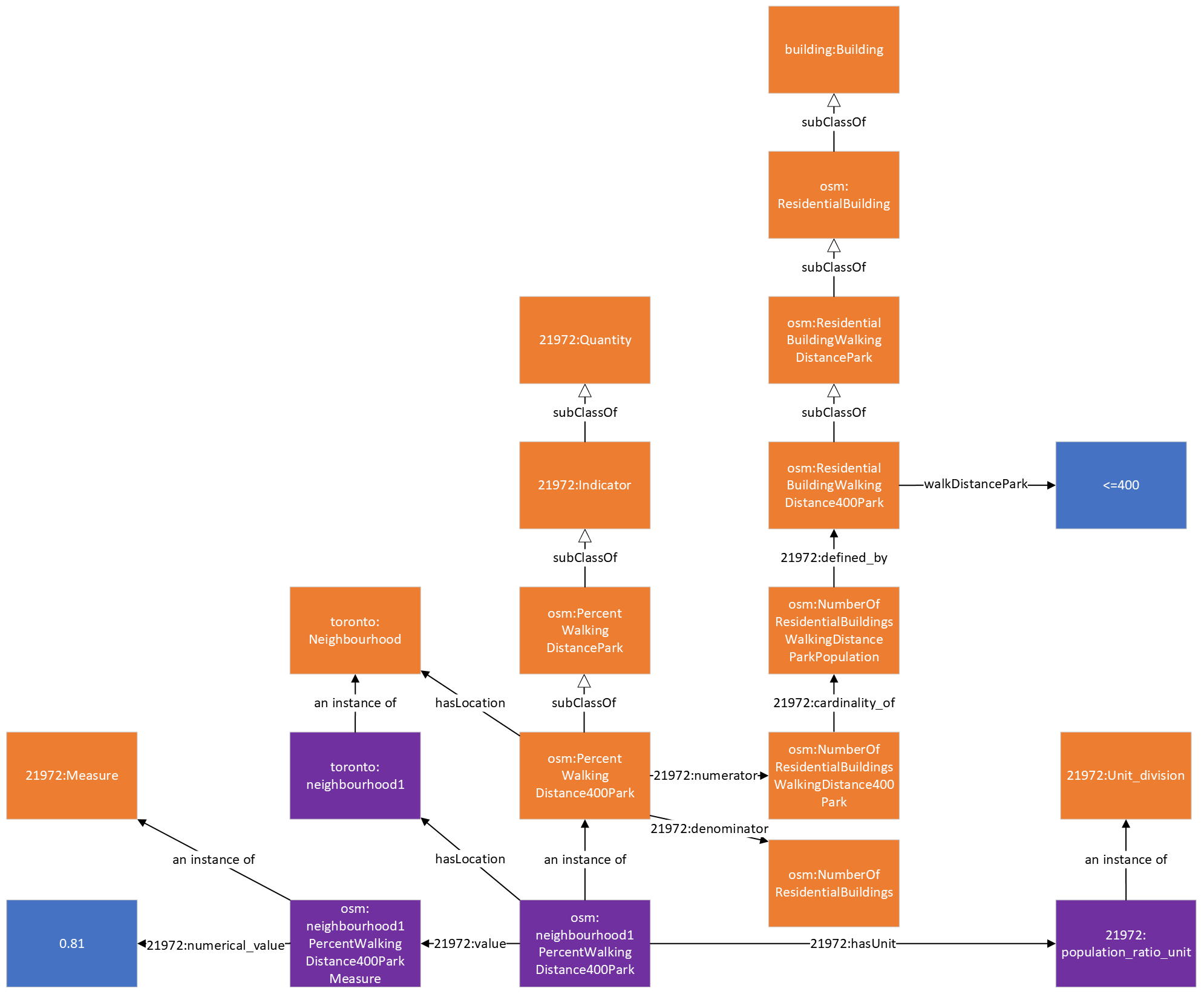
## OpenStreetMap Tags Mapping

This section provides a brief summary of how the tags used in OpenStreetMap were mapped into the ontology used in the City Digital Twin.

|  |  |  |
| --- | --- | --- |
| OpenStreetMap Tag | OpenStreetMap Description | CDT Equivalent |
| name | The primary name: in general, the most prominent signposted name or the most common name in the local language(s). | genprop:hasName |
| id | OSM ID is a numerical identifier that is assigned to every element in the OpenStreetMap (OSM) database. | cdt:osmID |
| addr:street | The street name that this address is (and any others in this location are) grouped by. This street name should match that of a nearby road, track or path. | contact:hasStreet (street name)  contact:hasStreetType (street type)  contact:hasStreetDirection(street direction |
| addr:housenumber | The house (or building) number that is included in the address. The number may contain non-digits and if recording multiple house numbers separate them by "," (e.g. "12b,12c"). | contact:hasStreetNumber |
| addr:postcode | The postal code / zip code that is included in the address. | contact:hasPostalCode |
| operator | Сompany, corporation, person or any other entity who is directly in charge of the current operation of a map object | cdt:operator |
| operator:type | Defines the type of operator, e.g. "public", "private", "government" | cdt:operatorType  gcie:hasOwnership (for schools) |
| website  contact:website | Specifies the link to the official website for a feature. | cdt:website |
| surface | The surface key is used to provide additional information about the physical surface of roads/footpaths and some other features, particularly regarding material composition and/or structure. | cdt:surface |
| opening\_hours | Describes when something is open or closed in a standard format | cdt:openingHours |
| phone | A telephone number associated with the object. Use +CC XXX XXX XXX format, where CC is a country code. | contact:hasTelephone |
| email  contact:email | An email address associated with the object | cdt:email |
| lit | The key lit=\* indicates the presence of lighting and can be used on nodes, ways, areas, or relations. | cdt:lit |
| wheelchair | Indicate if a special place can be used with wheelchairs | cdt:wheelchairAccess |
| school:language | The main language of teaching and the administration of a school | cdt:languageOfInstruction |
| isced:level | Indicates a level of education on a numeric scale inspired by the International Standard Classification of Education (ISCED) | code:hasCode  Note: the ISCED level of an educational facility is represented as an instance of the code:Code class with a name and description linked using the genprop:hasName and genprop:hasDescription properties respectively |
| religion | Defines the specific religion of a facility | cdt:religion |
| emergency | Used to indicate whether a hospital is equipped to deal with emergencies | cdt:emergencyServices |
| geometry | "Geometry" in this case means the shape and position of a node or way | geo:asWKT  Note: the geometry is converted to WKT format using Shapely’s to\_wkt function if the geometry is not initially in WKT format |
| dispensing | Whether a pharmacy dispenses prescription drugs or not | cdt:dispensing |
| cuisine | Describes the type of food served at a place | cdt:cuisine |
| delivery | Indicates whether a restaurant or shop offers delivery of meals or goods | cdt:delivery |
| takeaway | Indicates whether the restaurant offers meals for pick-up that can be consumed elsewhere | cdt:takeaway |
| smoking | Used to describe whether smoking is allowed or not in a facility or on a property | cdt:smoking |
| outdoor\_seating | To indicate if a feature offers outdoor seating | cdt:outdoorSeating |
| capacity | Describes the capacity a facility is suitable for | cdt:capacity |
| drive\_through | To indicate whether an amenity offers a drive-through service | cdt:driveThrough |

## Complete Community Walking Distance Indicator

This is an ontological representation of a complete community walking distance indicator that shows the percentage of residential buildings in a geographic area that are within a specified distance of a complete community amenity. The diagram below shows the representation of a walking distance indicator for parks where individual instances of an indicator represent the walking distance indicator for a geographic area specified by the “hasLocation” property. This indicator is linked to a measure instance using the “value” property and the measure instance is linked to the value of the indicator using the “numerical\_value” property. In the diagram below, we can see that the numerical value for the “neighborhood1PercentWalkingDistance400Park” indicator is 0.81 which means that 81% of the residential buildings in neighborhood1 are within a 400m distance of a park. This indicator is calculated by dividing the number of residential buildings that are within a 400m distance of a park by the total number of residential buildings in the area and this relationship is represented using the “numerator” and “denominator” properties, respectively. Similar indicators are also available for other Complete Community amenities.

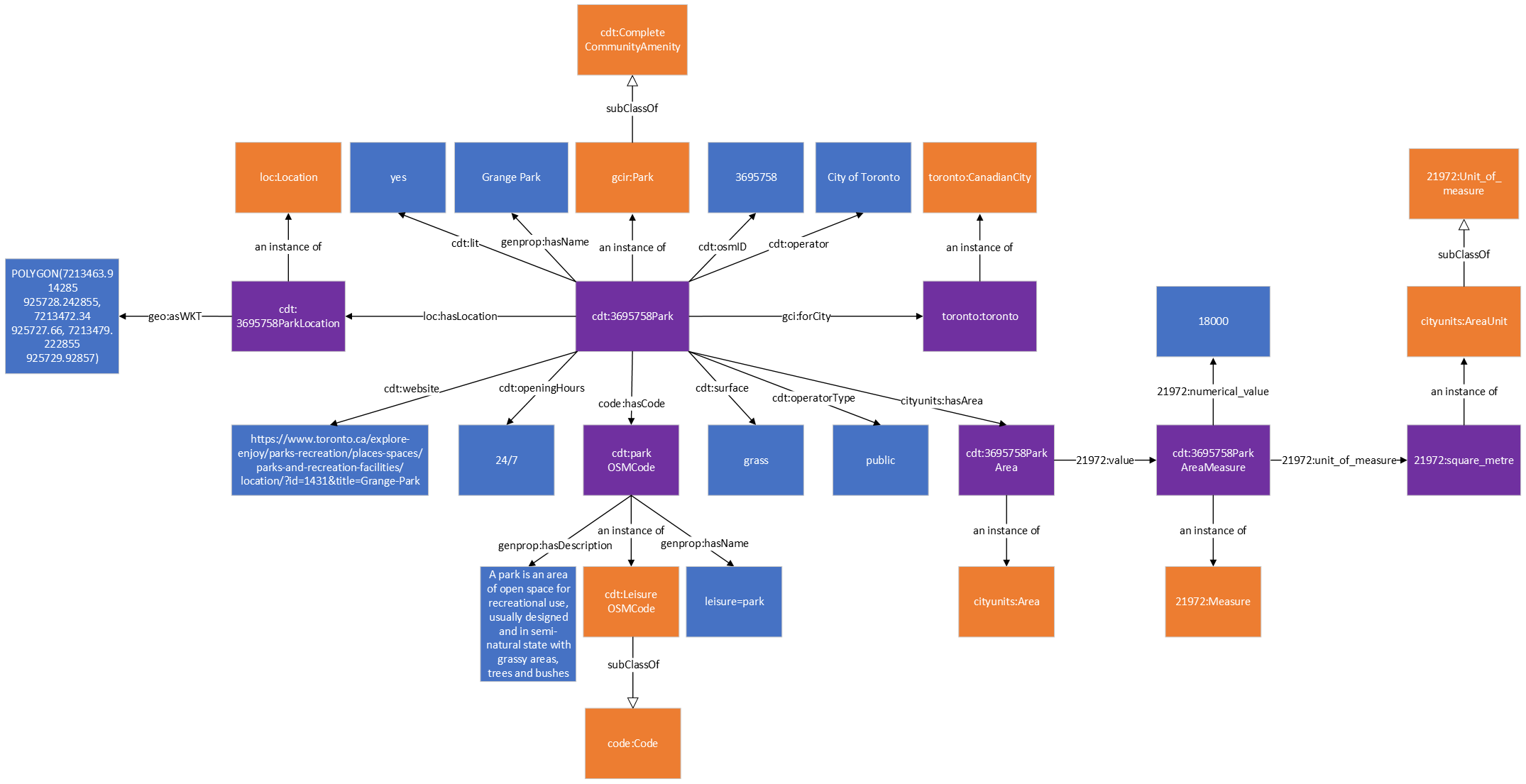


## Parks

This is an ontological representation of park data. Parks are represented as instances of the Park class from the Recreation Ontology for Global City Indicators. Parks can have a name which is linked via the hasName property, an operator which is linked via the operator property, an operator type (e.g. public, private) which is linked via the operatorType property, a value describing the physical surface of the park which is linked via the surface property, opening hours which is linked via the openingHours property, a website which is linked via the website property, an indicator for whether the park is lit or not via the lit property, a hasArea property that links the park to its area instance and area value using the numerical\_value property, a Location instance and a set of geospatial coordinates linked using the hasLocation and asWKT property respectively, and a code instance that can be used to link the park to a specific set list of values (e.g., classification systems).

Additionally, for OpenStreetMap park data, the instance can also have an osmID property that links the park to its OSM ID (a unique identifier that is used in OpenStreetMap). It can also be linked to OpenStreetMap’s “leisure=park” tag that is used to identify parks in OpenStreetMap using the hasCode property and Code class.

The value of a park’s surface area is calculated using the Pyproj (a cartographic projections and coordinate transformations library for Python) and its polygon coordinates. The Shapely Python package is also used to convert the coordinates into a Shapely shape for use with Pyproj functions.

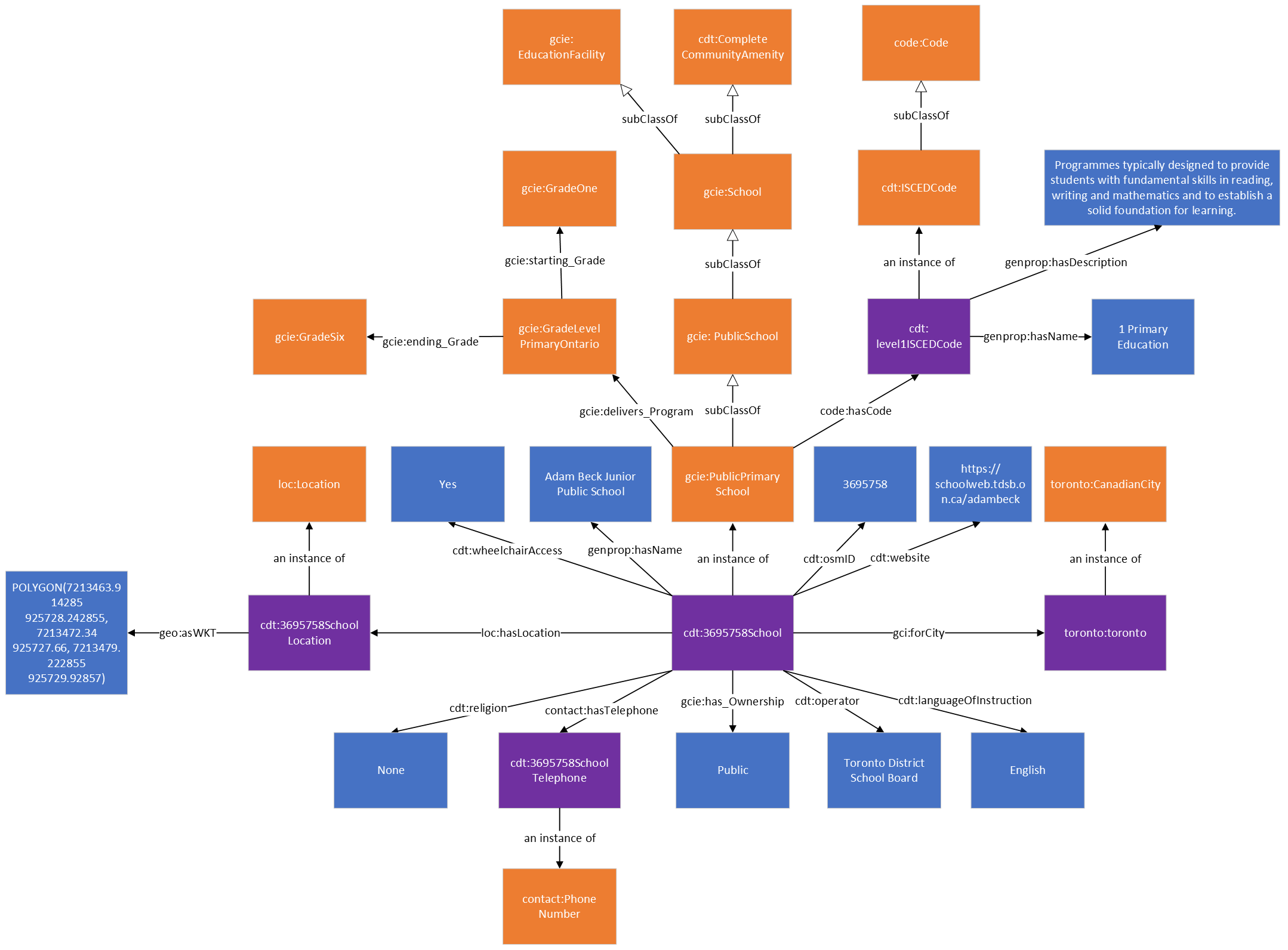


## Schools

This is an ontological representation of school data. Schools are represented as instances of the School class from the GCI Education Ontology (GCIE) which is also a subclass of the EducationFacility class from GCIE and the CompleteCommunityAmenity class. Schools can have a name which is linked via the hasName property, a website which is linked via the website property, a religion which is linked via the religion property, a phone number which is linked via the hasPhoneNumber property, an ownership type (e.g. public, private) which is linked via the has\_Ownership property, an operator which is linked via the operator property, a language of instruction (i.e. the language used to teach in a school) which is linked via the languageOfInstruction property, its wheelchair accessibility which is linked via the wheelchairAccess property, a Location instance and a set of geospatial coordinates linked using the hasLocation and asWKT property respectively. School classes are linked to an ISCED code instance which describes the level of education that type of school provides. For example, in the diagram below, the PublicPrimarySchool class has an ISCED code of level 1 which represents Primary Education (i.e., programmes typically designed to provide students with fundamental skills in reading, writing and mathematics and to establish a solid foundation for learning.)

Additionally, for OpenStreetMap school data, the instance can also have an osmID property that links the school to its OSM ID (a unique identifier that is used in OpenStreetMap).

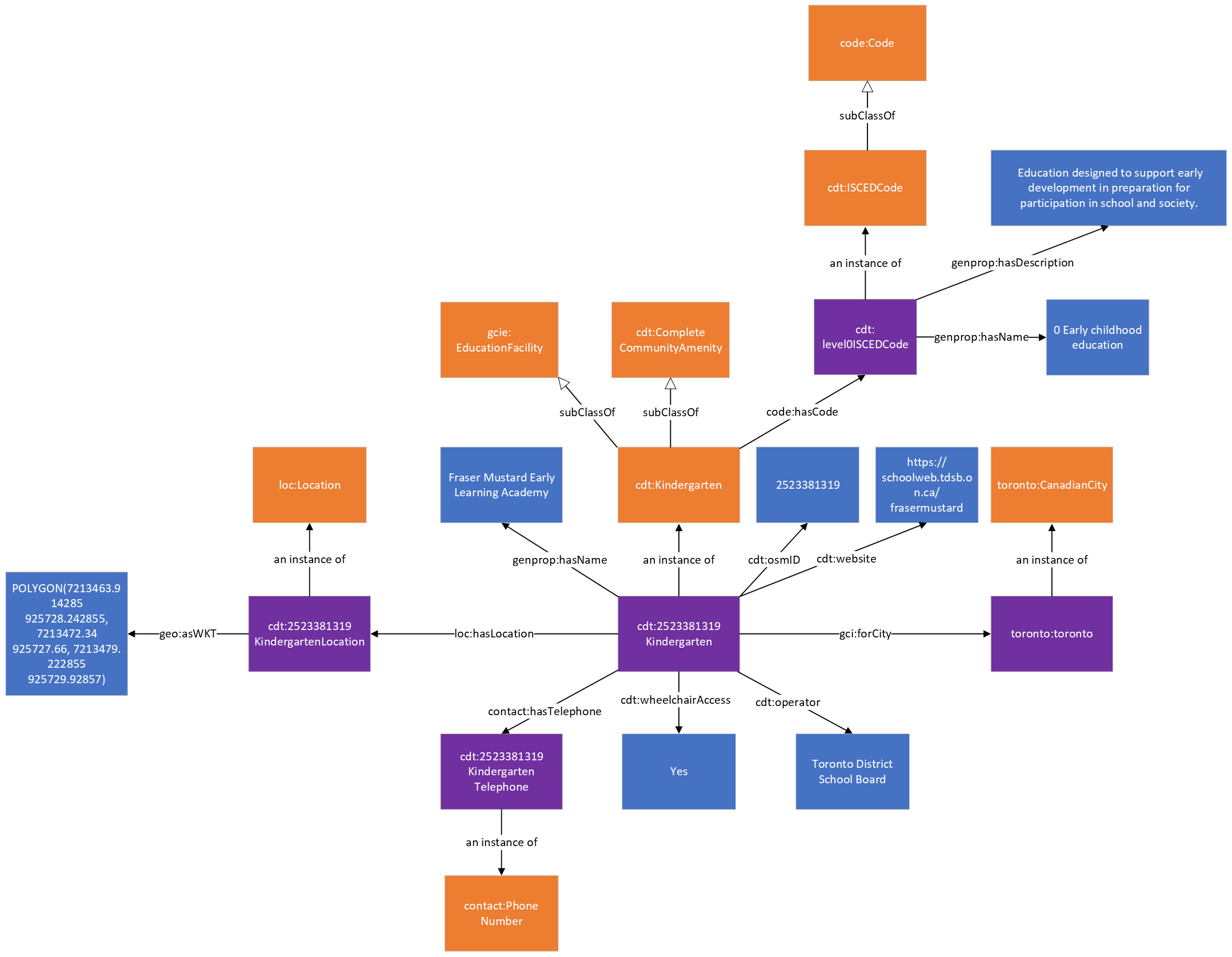
Schools with different operator types (e.g. public, private) and level of education are placed in the appropriate school subclass from the GCIE ontology (e.g. a primary school that is publicly operated would be an instance of the gcie:PublicPrimarySchool class).



## Kindergartens

This is an ontological representation of kindergarten data. Kindergartens are represented as instances of the Kindergarten class which is a subclass of the EducationFacility class from the GCI Education Ontology and the CompleteCommunityAmenity class. Kindergartens can have a name which is linked via the hasName property, a website which is linked via the website property, an operator which is linked via the operator property, a phone number which is linked via the hasTelephone property, its wheelchair accessibility which is linked via the wheelchairAccess property, a Location instance and a set of geospatial coordinates linked using the hasLocation and asWKT property respectively. The Kindergarten class is linked to an ISCED code instance which describes the level of education that a kindergarten provides using the hasCode property. Kindergartens provide ISCED level 0 education (i.e., education designed to support early development in preparation for participation in school and society).

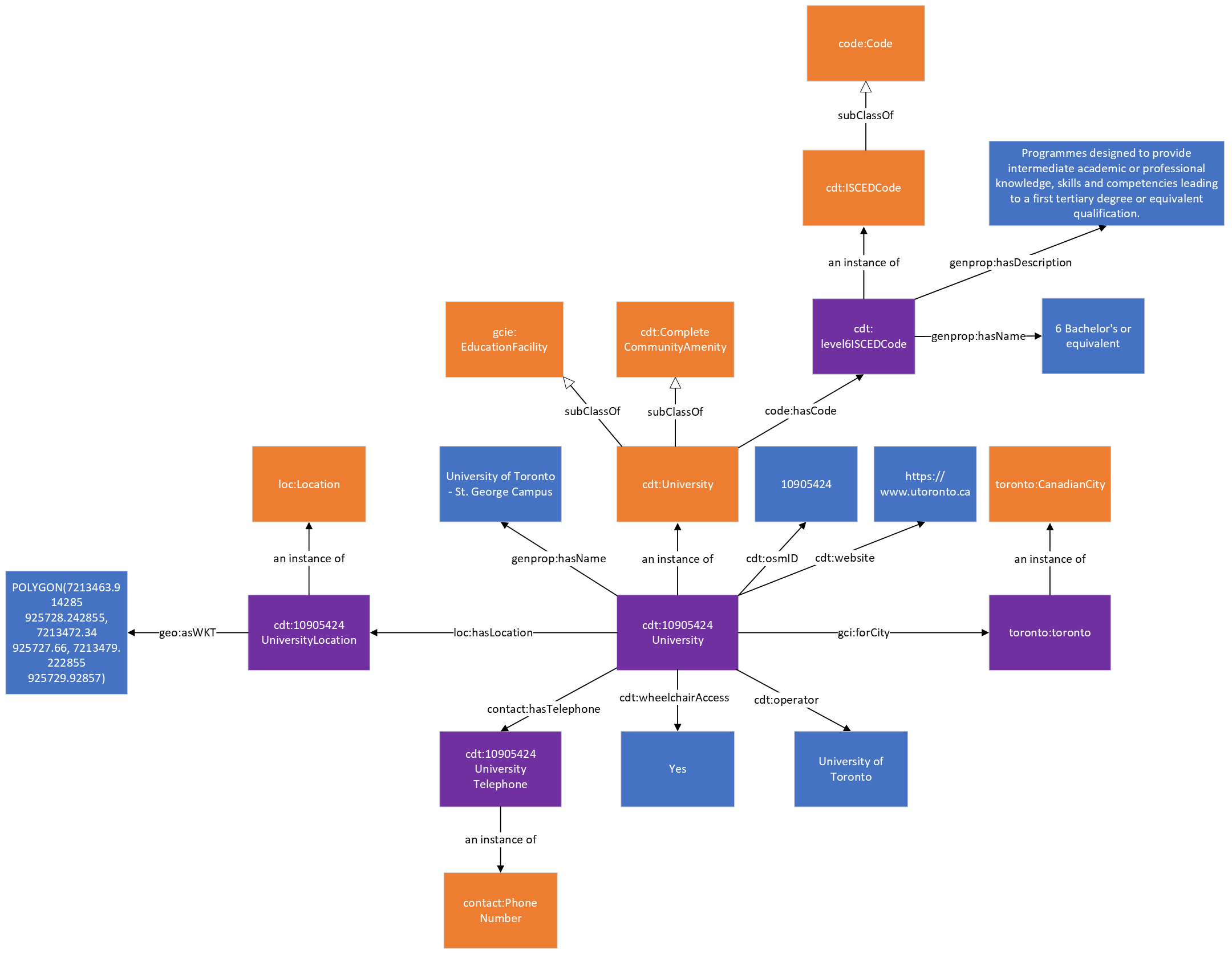
Additionally, for OpenStreetMap kindergarten data, the instance can also have an osmID property that links the kindergarten to its OSM ID (a unique identifier that is used in OpenStreetMap).



## Universities

This is an ontological representation of university data. Universities are represented as instances of the University class which is a subclass of the EducationFacility class from the GCI Education Ontology and the CompleteCommunityAmenity class. Universities can have a name which is linked via the hasName property, a website which is linked via the website property, an operator which is linked via the operator property, a phone number which is linked via the hasTelephone property, its wheelchair accessibility which is linked via the wheelchairAccess property, a Location instance and a set of geospatial coordinates linked using the hasLocation and asWKT property respectively. The University class is linked to an ISCED code instance which describes the level of education that a university provides using the hasCode property. Universities provide ISCED level 6 education (i.e., programmes designed to provide intermediate academic or professional knowledge, skills and competencies leading to a first tertiary degree or equivalent qualification) and higher.

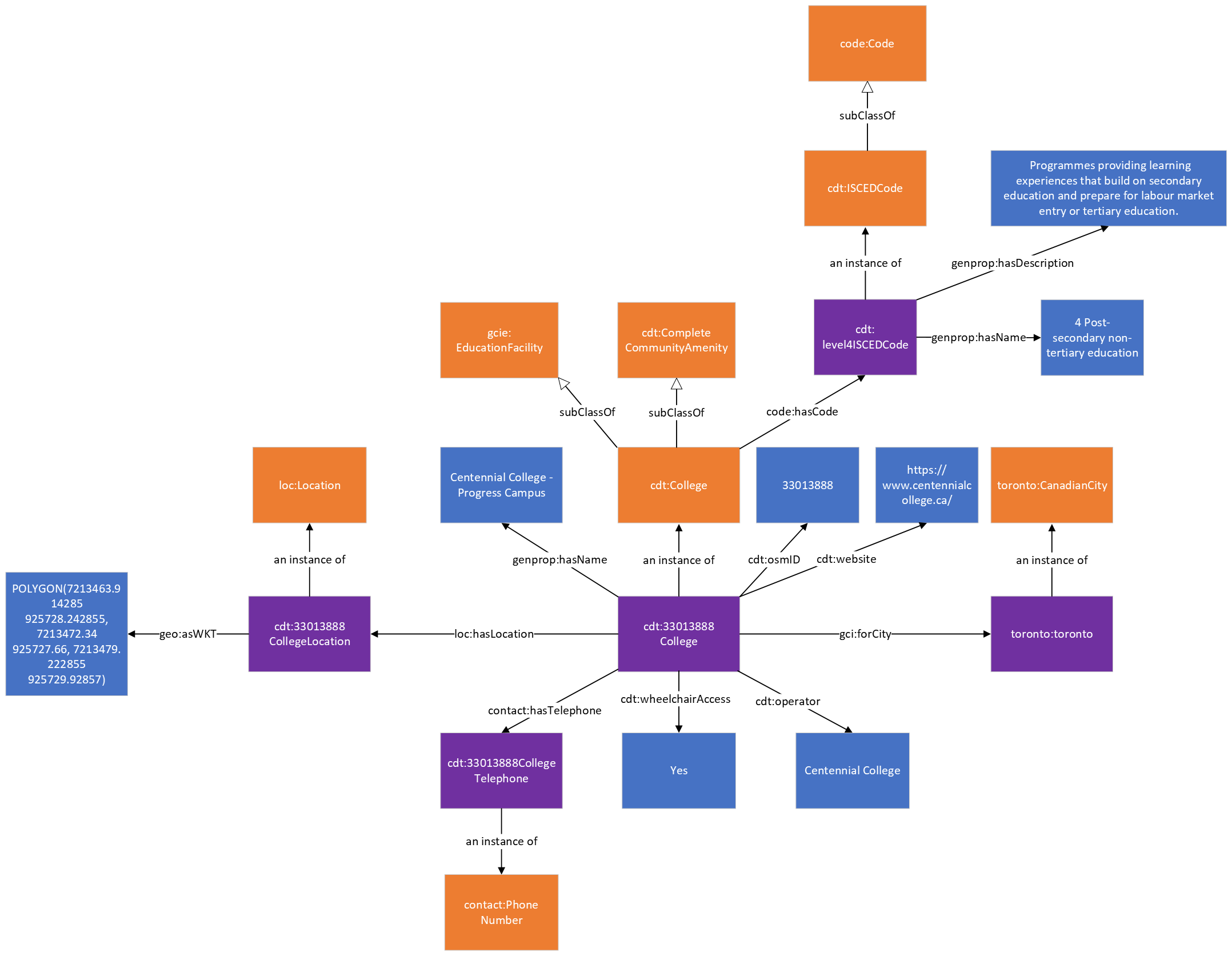
Additionally, for OpenStreetMap university data, the instance can also have an osmID property that links the university to its OSM ID (a unique identifier that is used in OpenStreetMap).



## Colleges

This is an ontological representation of college data. Colleges are represented as instances of the College class which is a subclass of the EducationFacility class from the GCI Education Ontology and the CompleteCommunityAmenity class. Colleges can have a name which is linked via the hasName property, a website which is linked via the website property, an operator which is linked via the operator property, a phone number which is linked via the hasTelephone property, its wheelchair accessibility which is linked via the wheelchairAccess property, a Location instance and a set of geospatial coordinates linked using the hasLocation and asWKT property respectively. The College class is linked to an ISCED code instance which describes the level of education that a college provides using the hasCode property. Colleges provide ISCED level 4 education (i.e., programmes providing learning experiences that build on secondary education and prepare for labour market entry or tertiary education.) and higher.

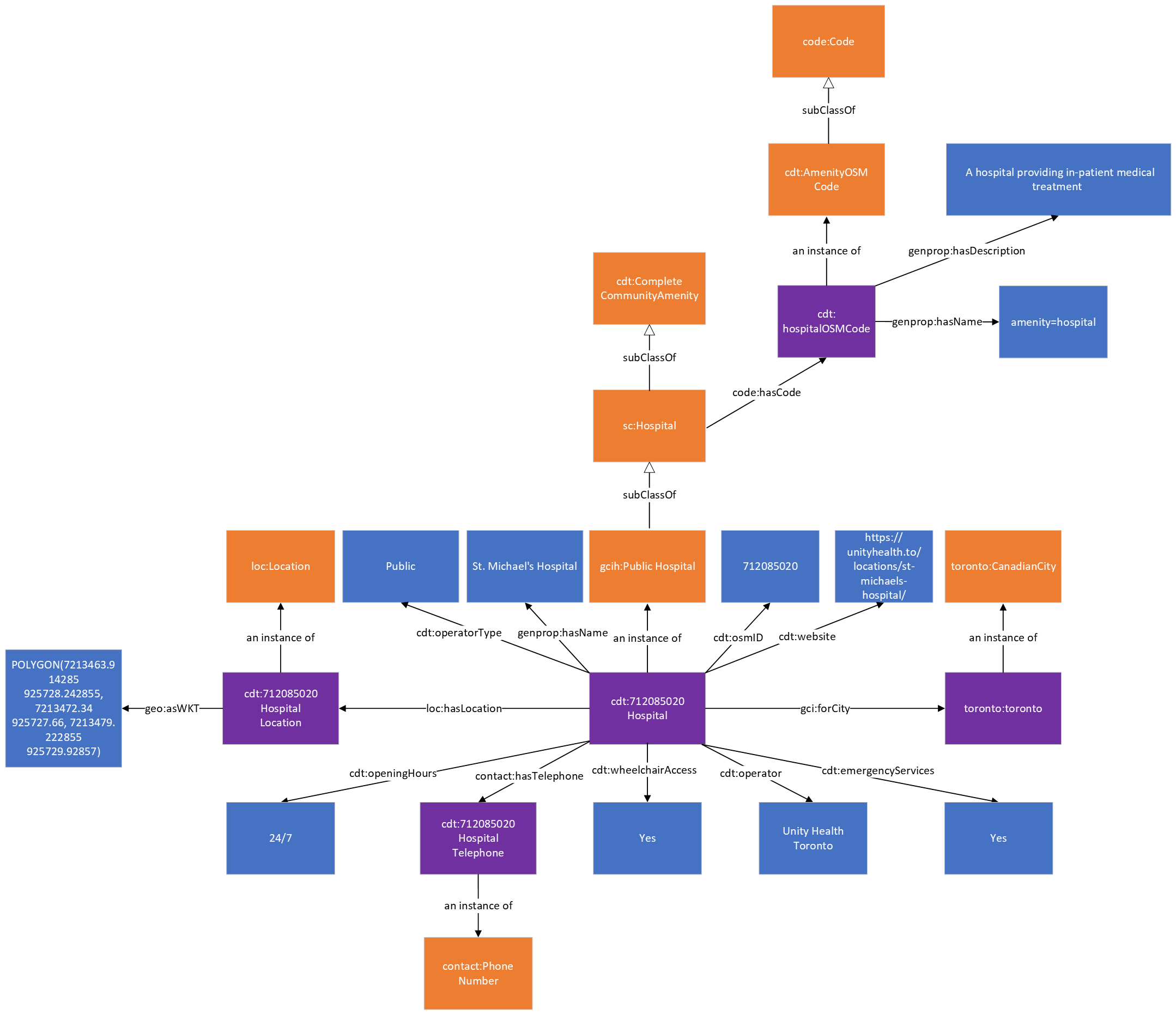
Additionally, for OpenStreetMap college data, the instance can also have an osmID property that links the college to its OSM ID (a unique identifier that is used in OpenStreetMap).



## Hospitals

This is an ontological representation of hospital data. Hospitals are represented as instances of the Hospital class from Schema.org which has the subclasses PublicHospital and PrivateHospital from the GCI Healthcare Ontology to indicate whether they are publicly or privately operated and the superclass CompleteCommunityAmenity class. Hospitals can have a name which is linked via the hasName property, whether it provides emergency services via the emergencyServices property, a website which is linked via the website property, an operator which is linked via the operator property, the type of the operator which is linked via the operatorType property, a phone number which is linked via the hasTelephone property, its wheelchair accessibility which is linked via the wheelchairAccess property, its opening hours via the openingHours property, a Location instance and a set of geospatial coordinates linked using the hasLocation and asWKT property respectively. The Hospital class can also use the hasCode property in order to link it to a specific set list of values (e.g., classification systems).

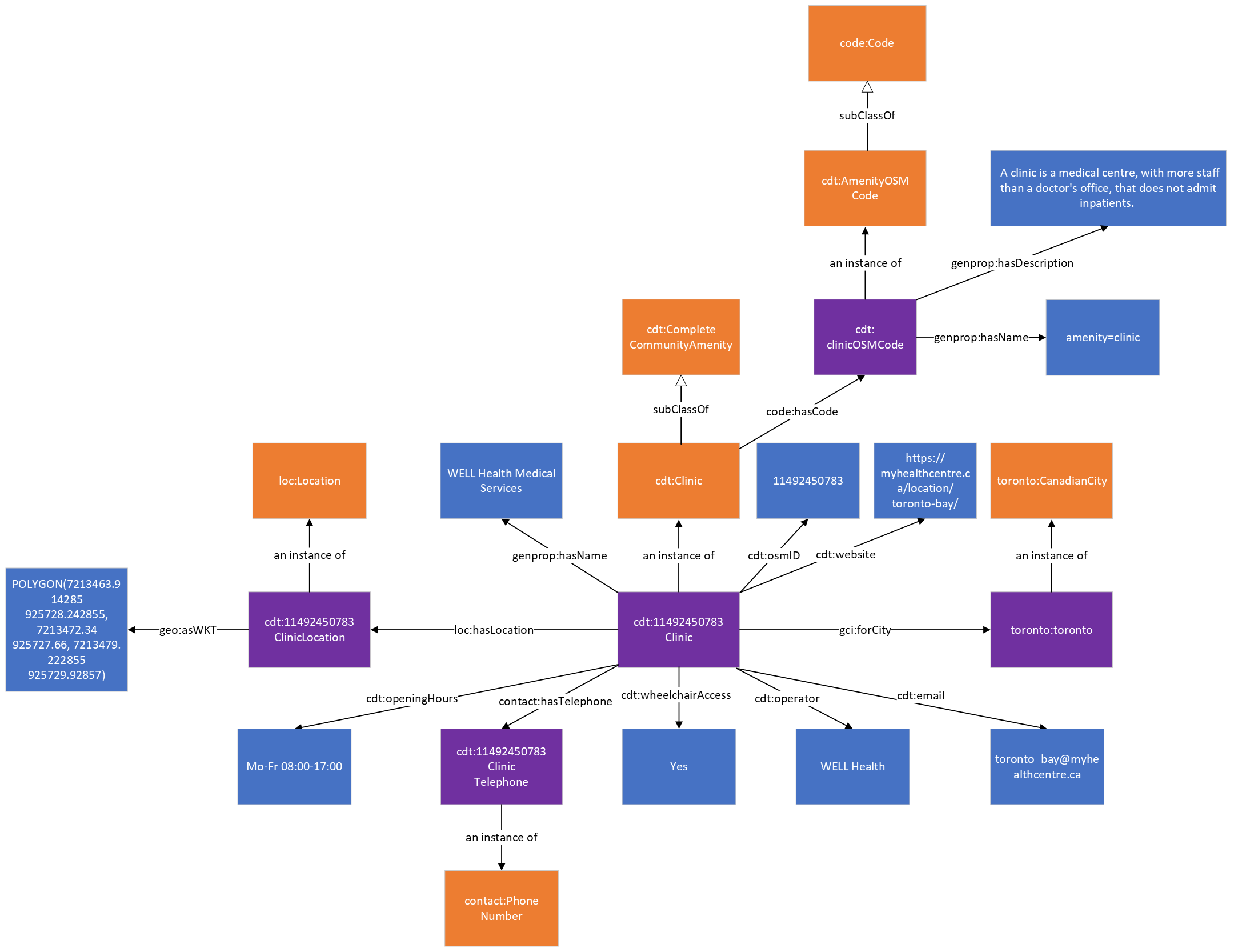
Additionally, for OpenStreetMap hospital data, the instance can also have an osmID property that links the hospital to its OSM ID (a unique identifier that is used in OpenStreetMap).



## Clinics

This is an ontological representation of clinic (i.e. medical centre) data. Clinics are represented as instances of the Clinic class which is a subclass of the CompleteCommunityAmenity class. Clinics can have a name which is linked via the hasName property, a website which is linked via the website property, an operator which is linked via the operator property, a contact email which is linked via the email property, a phone number which is linked via the hasTelephone property, its wheelchair accessibility which is linked via the wheelchairAccess property, its opening hours via the openingHours property, a Location instance and a set of geospatial coordinates linked using the hasLocation and asWKT property respectively. The Clinic class can also use the hasCode property in order to link it to a specific set list of values (e.g., classification systems).

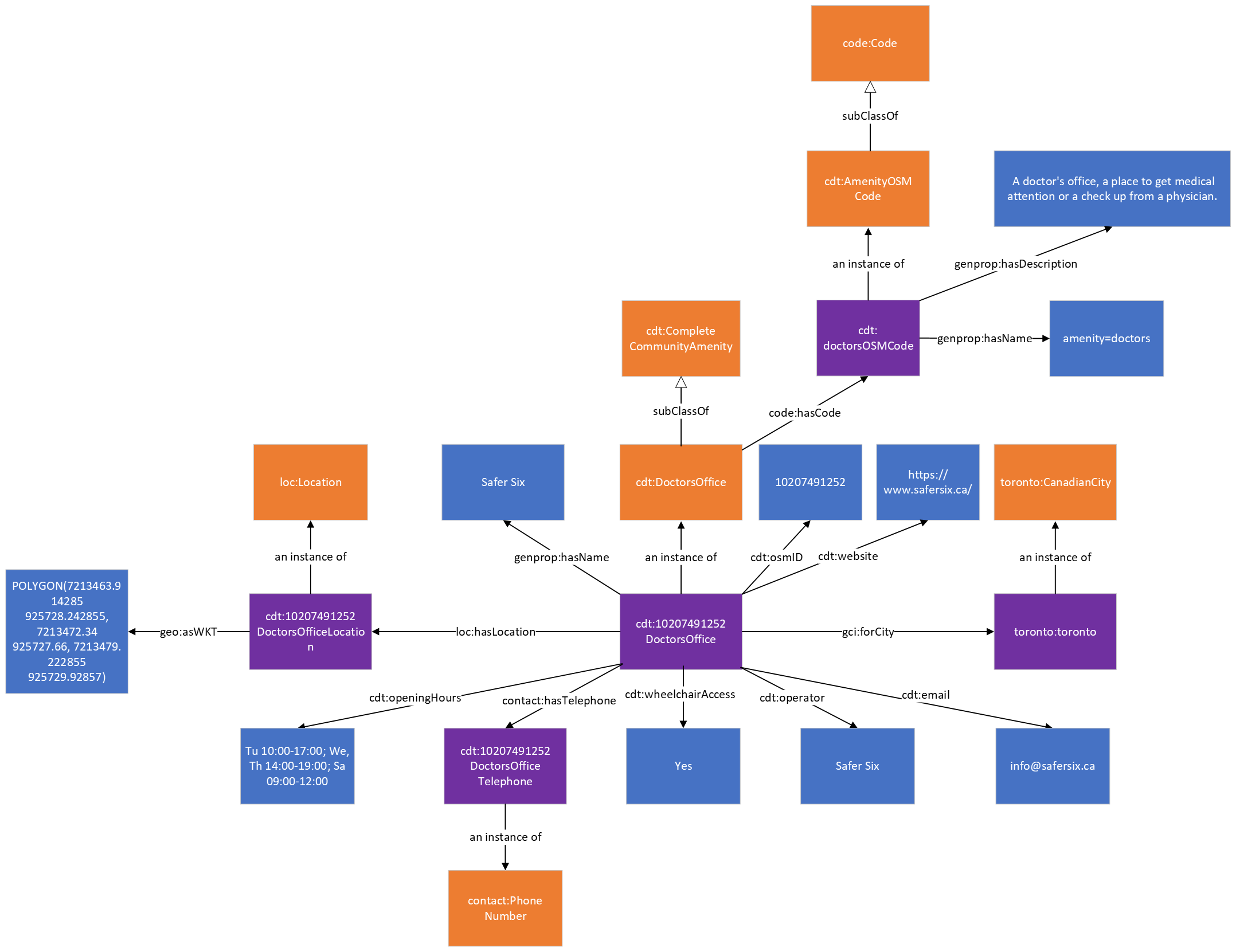
Additionally, for OpenStreetMap clinic data, the instance can also have an osmID property that links the clinic to its OSM ID (a unique identifier that is used in OpenStreetMap).



## Doctor’s Office

This is an ontological representation of doctor’s office data. Doctor’s offices are represented as instances of the DoctorsOffice class which is a subclass of the CompleteCommunityAmenity class. Doctor’s offices can have a name which is linked via the hasName property, a website which is linked via the website property, an operator which is linked via the operator property, a contact email which is linked via the email property, a phone number which is linked via the hasTelephone property, its wheelchair accessibility which is linked via the wheelchairAccess property, its opening hours via the openingHours property, a Location instance and a set of geospatial coordinates linked using the hasLocation and asWKT property, respectively. The DoctorsOffice class can also use the hasCode property in order to link it to a specific set list of values (e.g., classification systems).

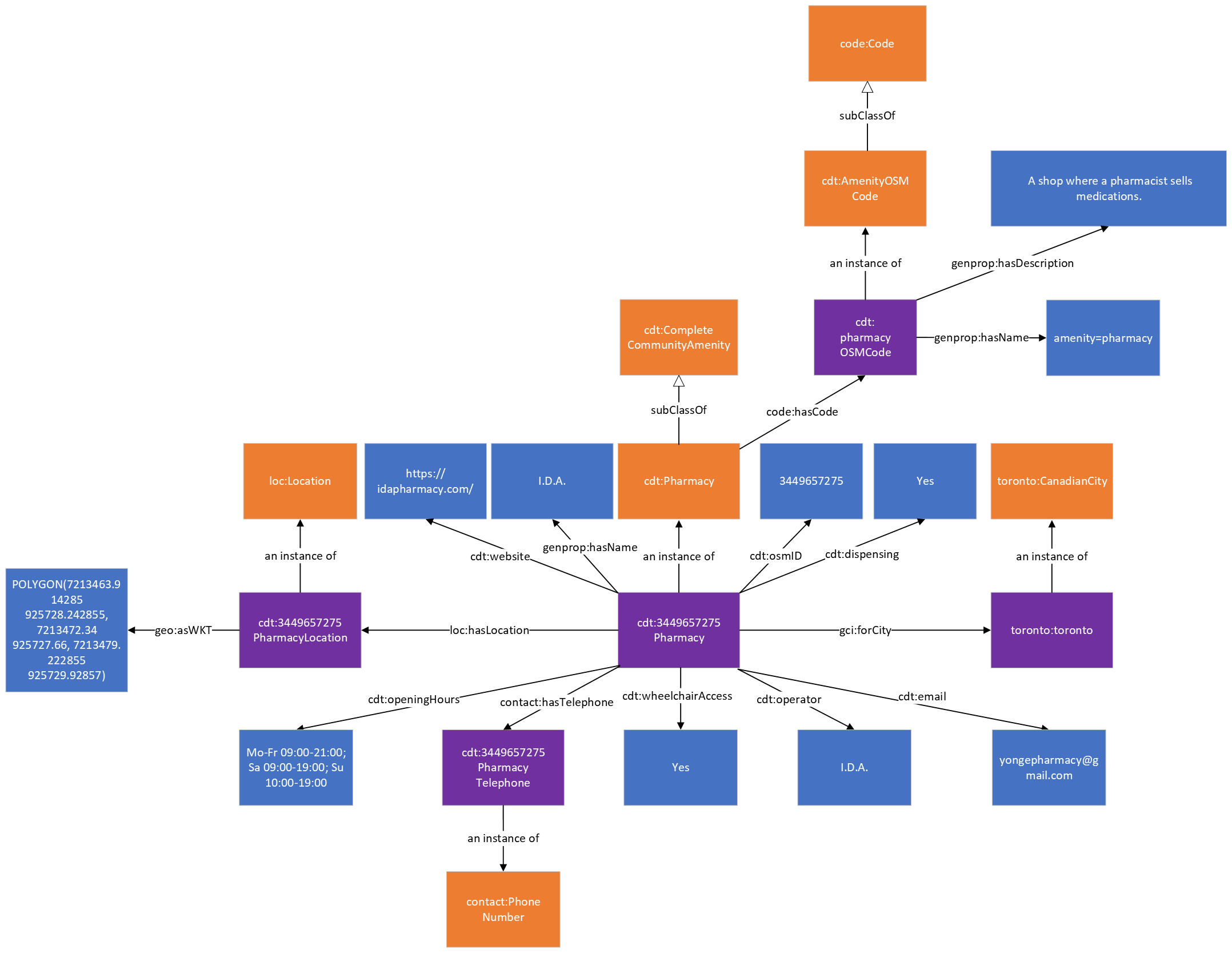
Additionally, for OpenStreetMap doctor’s office data, the instance can also have an osmID property that links the doctor’s office to its OSM ID (a unique identifier that is used in OpenStreetMap).



## Pharmacy

This is an ontological representation of pharmacy data. Pharmacies are represented as instances of the Pharmacy class which is a subclass of the CompleteCommunityAmenity class. Pharmacies can have a name which is linked via the hasName property, a website which is linked via the website property, an operator which is linked via the operator property, a contact email which is linked via the email property, a phone number which is linked via the hasTelephone property, its wheelchair accessibility which is linked via the wheelchairAccess property, its opening hours via the openingHours property, whether they provide drug dispensing services via the dispensing property, a Location instance and a set of geospatial coordinates linked using the hasLocation and asWKT property, respectively. The Pharmacy class can also use the hasCode property in order to link it to a specific set list of values (e.g., classification systems).

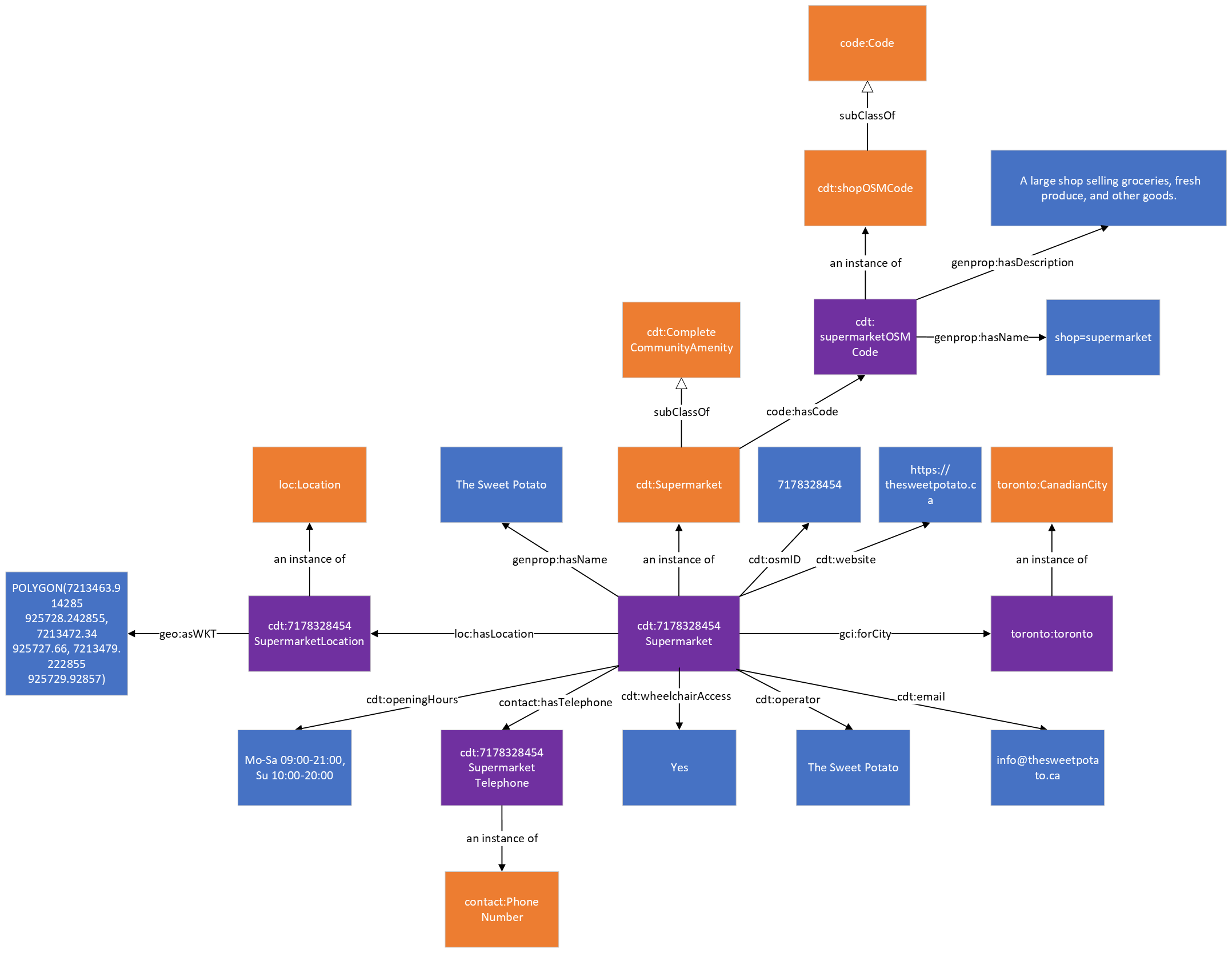
Additionally, for OpenStreetMap pharmacy data, the instance can also have an osmID property that links the pharmacy to its OSM ID (a unique identifier that is used in OpenStreetMap).



## Supermarket

This is an ontological representation of supermarket data. Supermarkets are represented as instances of the Supermarket class which is a subclass of the CompleteCommunityAmenity class. Supermarkets can have a name which is linked via the hasName property, a website which is linked via the website property, an operator which is linked via the operator property, a contact email which is linked via the email property, a phone number which is linked via the hasTelephone property, its wheelchair accessibility which is linked via the wheelchairAccess property, its opening hours via the openingHours property, a Location instance and a set of geospatial coordinates linked using the hasLocation and asWKT property, respectively. The Supermarket class can also use the hasCode property in order to link it to a specific set list of values (e.g., classification systems).

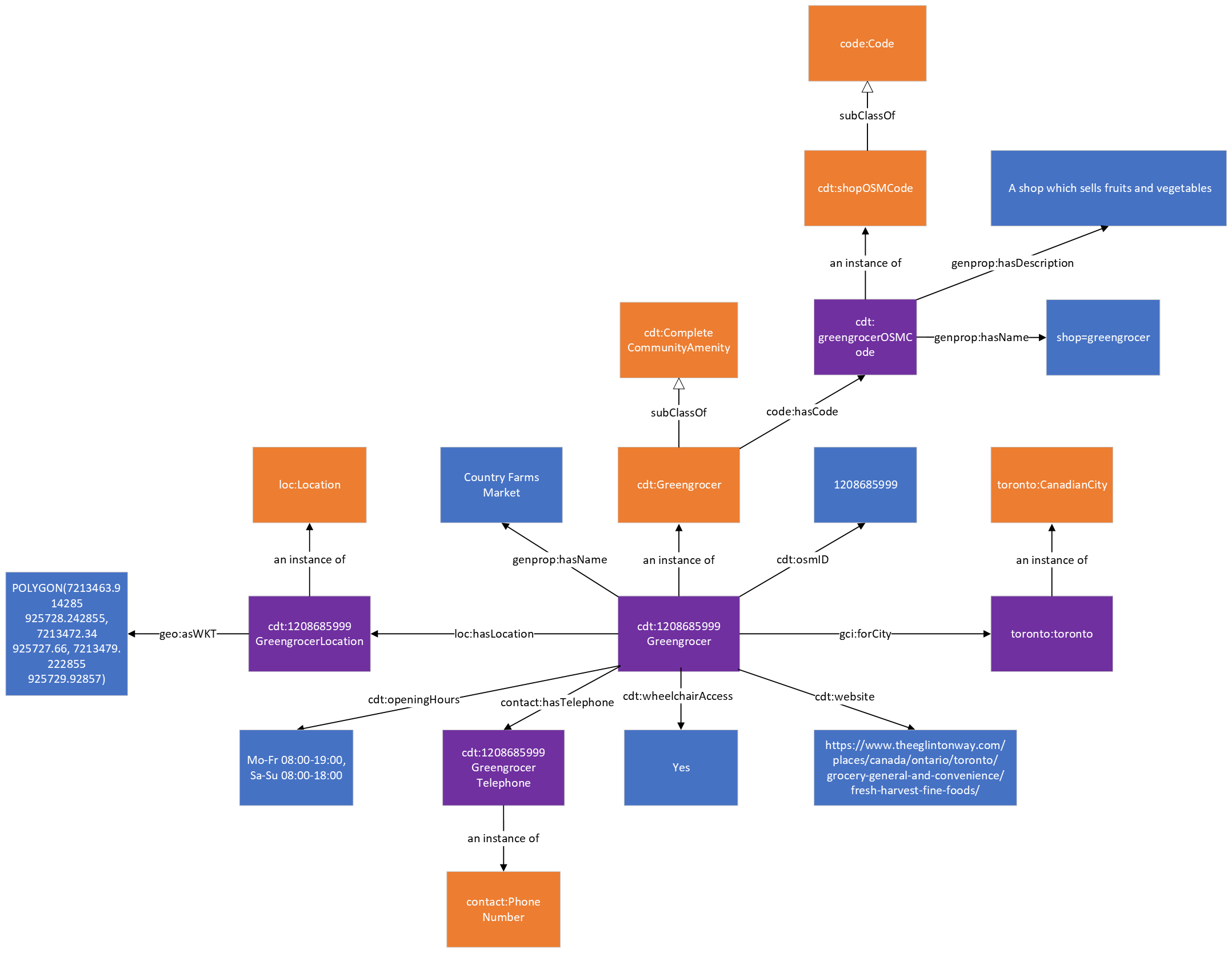
Additionally, for OpenStreetMap supermarket data, the instance can also have an osmID property that links the supermarket to its OSM ID (a unique identifier that is used in OpenStreetMap).



## Greengrocer

This is an ontological representation of greengrocer data. Greengrocers are represented as instances of the Greengrocer class which is a subclass of the CompleteCommunityAmenity class. Greengrocers can have a name which is linked via the hasName property, a website which is linked via the website property, a phone number which is linked via the hasTelephone property, its wheelchair accessibility which is linked via the wheelchairAccess property, its opening hours via the openingHours property, a Location instance and a set of geospatial coordinates linked using the hasLocation and asWKT property, respectively. The Greengrocer class can also use the hasCode property in order to link it to a specific set list of values (e.g., classification systems).

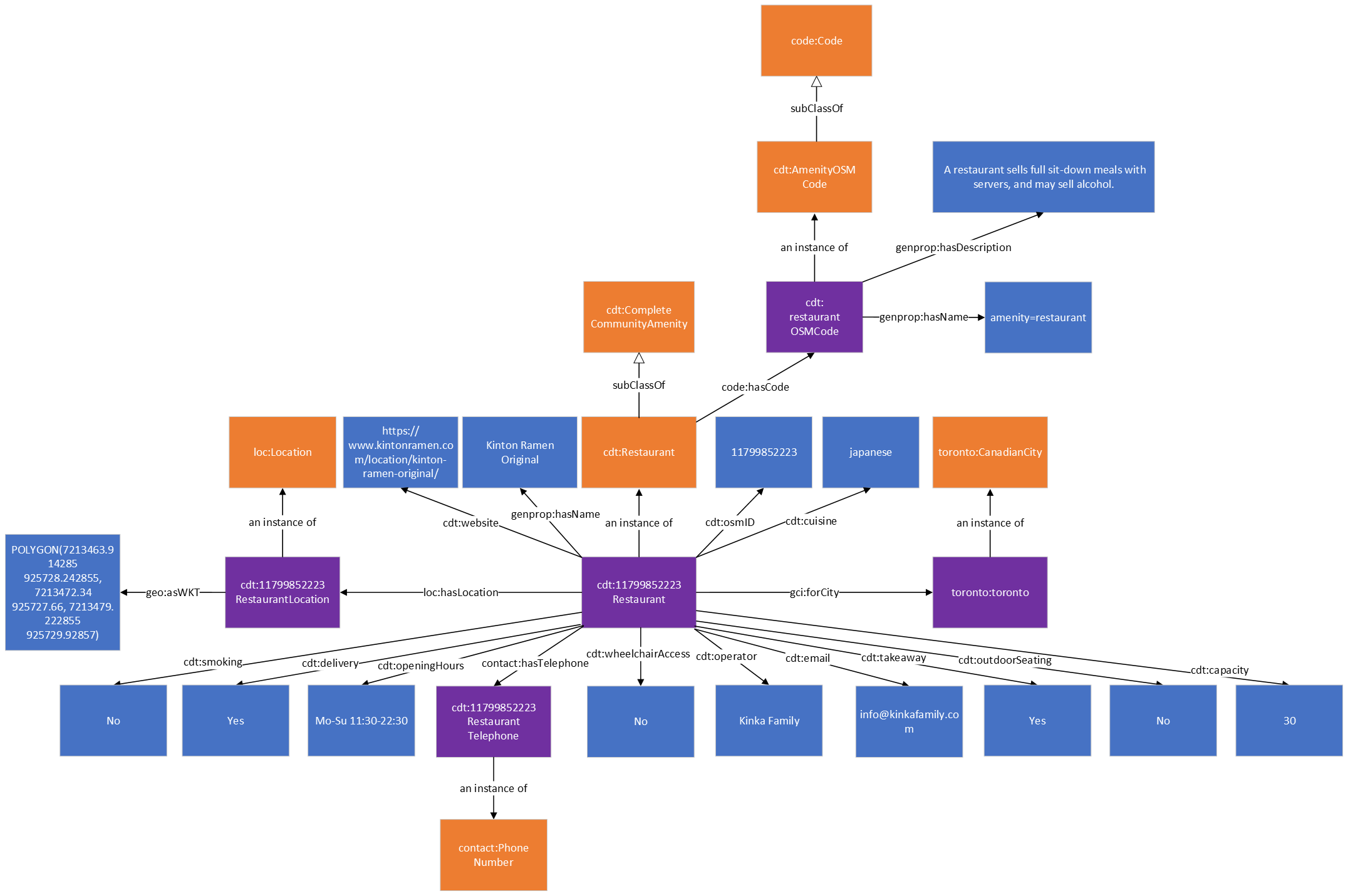
Additionally, for OpenStreetMap greengrocer data, the instance can also have an osmID property that links the greengrocer to its OSM ID (a unique identifier that is used in OpenStreetMap).



## Restaurant

This is an ontological representation of restaurant data. Restaurants are represented as instances of the Restaurant class which is a subclass of the CompleteCommunityAmenity class. Restaurants can have a name which is linked via the hasName property, a website which is linked via the website property, an email which is linked via the email property, a phone number which is linked via the hasTelephone property, its wheelchair accessibility which is linked via the wheelchairAccess property, its opening hours via the openingHours property, the type of cuisine that it serves via the cuisine property, whether delivery is available via the delivery property, whether takeaway is offered via the takeaway property, whether smoking is allowed via the smoking property, whether outdoor seating is available via the outdoorSeating property, its capacity (i.e. number of people it can accomodate) via the capacity property, a Location instance and a set of geospatial coordinates linked using the hasLocation and asWKT property, respectively. The Restaurant class can also use the hasCode property in order to link it to a specific set list of values (e.g., classification systems).

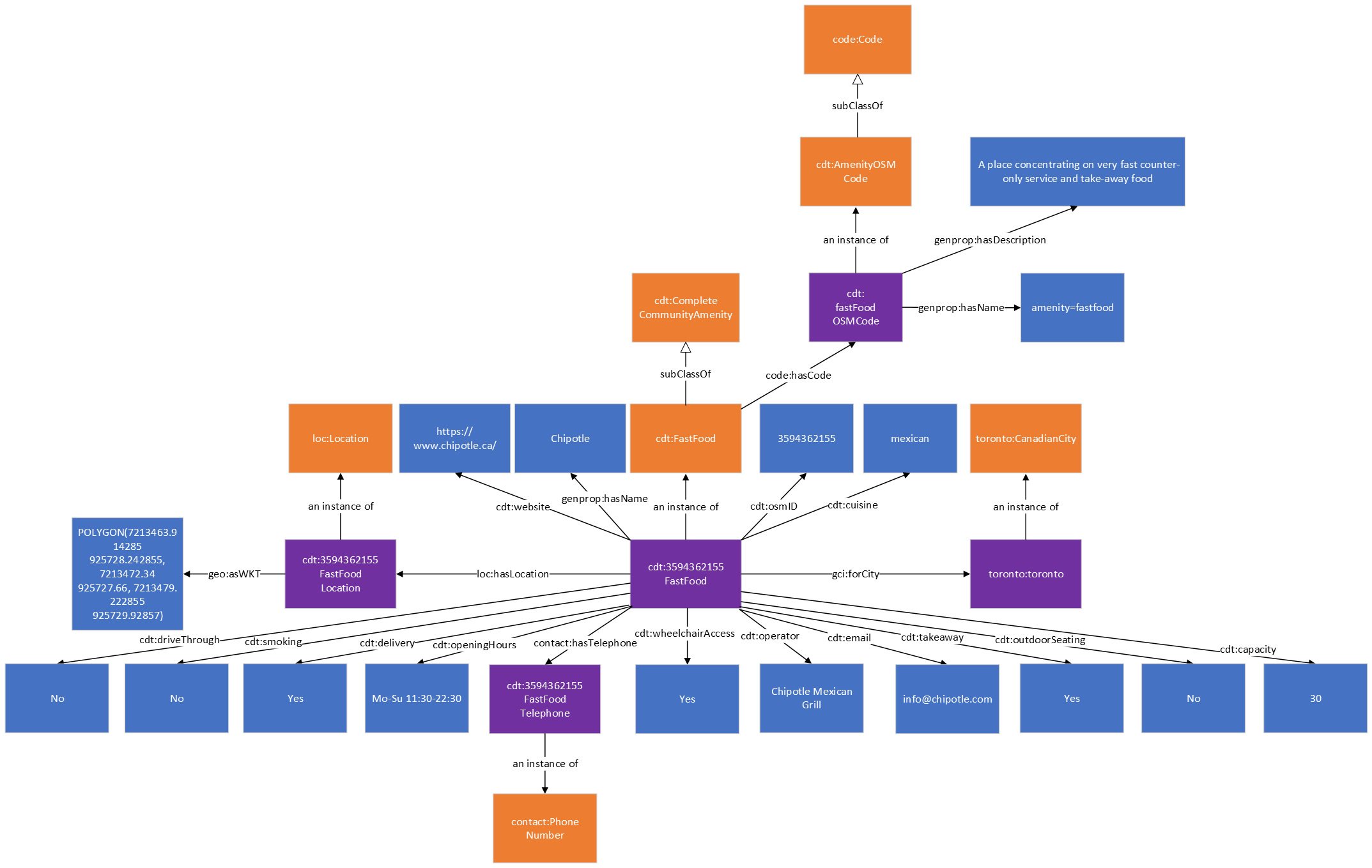
Additionally, for OpenStreetMap restaurant data, the instance can also have an osmID property that links the restaurant to its OSM ID (a unique identifier that is used in OpenStreetMap).



## FastFood

This is an ontological representation of fast food vendor data. Fast food vendorsare represented as instances of the FastFood class which is a subclass of the CompleteCommunityAmenity class. Fast food vendors can have a name which is linked via the hasName property, a website which is linked via the website property, an email which is linked via the email property, a phone number which is linked via the hasTelephone property, its wheelchair accessibility which is linked via the wheelchairAccess property, its opening hours via the openingHours property, the type of cuisine that it serves via the cuisine property, whether delivery is available via the delivery property, whether takeaway is offered via the takeaway property, whether smoking is allowed via the smoking property, whether outdoor seating is available via the outdoorSeating property, its capacity (i.e. number of people it can accomodate) via the capacity property, whether it offers drive-through services via the driveThrough property, a Location instance and a set of geospatial coordinates linked using the hasLocation and asWKT property, respectively. The FastFood class can also use the hasCode property in order to link it to a specific set list of values (e.g., classification systems).

Additionally, for OpenStreetMap fast food vendor data, the instance can also have an osmID property that links the fast food vendorto its OSM ID (a unique identifier that is used in OpenStreetMap).



# Representation of Road Data

The Ontario Road Network (ORN) dataset represents the provincial road infrastructure and is segmented at real-world intersections or junctions. Each road network element is bounded by a FROM\_JCT and a TO\_JCT, making it semantically aligned with the RoadLink class in the ISO/IEC 5087-3 ontology. Each RoadLink begins and ends at a TransportNode, which corresponds to a Junction in the ontology

Thus, in our City Digital Twin:

* **ORN road net elements** → transnet:RoadLink (subclass of TravelledWayLink)
* **ORN junctions** → transnet:Junction (subclass of TransportNode)
* A group of RoadLinks sharing the same street name → transnet:Road

The data was also filtered around only Toronto roads using the Toronto Bound Filter:

lat\_min: 43.5810, lat\_max: 43.8555, lon\_min: -79.6393, lon\_max: -79.1152

The following is a list of namespace prefixes:

* geo: http://www.opengis.net/ont/geosparql#
* transnet: https://standards.iso.org/isoiec/5087/3/ed1/en/ontology/TransportaionNetwork/
* infras: https://standards.iso.org/isoiec/5087/2/ed1/en/ontology/TransportationInfrastructure/
* loc: https://standards.iso.org/iso-iec/5087/-1/ed-1/en/ontology/SpatialLoc/>
* partwhole: https://standards.iso.org/iso-iec/5087/-1/ed-1/en/ontology/Mereology/
* cityunits: https://standards.iso.org/iso-iec/5087/-1/ed-1/en/ontology/CityUnits/
* cdt: http://ontology.eil.utoronto.ca/CDT#
* rdfs: http://www.w3.org/2000/01/rdf-schema#
* i72: http://ontology.eil.utoronto.ca/5087/2/iso21972/
* genprop: https://standards.iso.org/iso-iec/5087/-1/ed-1/en/ontology/GenericProperties/
* rdf: [http://www.w3.org/1999/02/22-rdf-syntax-ns#](http://www.w3.org/1999/02/22-rdf-syntax-ns)
* transalert: <https://standards.iso.org/iso-iec/5087/-3/ed-1/en/ontology/TransportAlert/>
* contact: https://standards.iso.org/iso-iec/5087/-2/ed-1/en/ontology/Contact/

## Junction

Junctions are an instance of transnet:Junctions and are a subclass of TransportNode. These entities connect travellers from one TravelledWayLink to another, in particular serving a connection between one or more RoadLinks and is uniquely identified using ORN-provided IDs. Geospatial coordinates are linked using geo:Geometry pointing to a loc:asWKT. Each Junction participates in one or more ingress and egress relationships with RoadLinks, ensuring accurate topological representation of the network.

A diagram of a network

Description automatically generated

## Road

A Road is modeled as a transinfras:Road, a subclass of TravelledWay, and is defined as a continuous sequence of RoadLinks that share a common entity (e.g., Highway 401, Dundas Street). A single RoadLink may be part of multiple Roads to accommodate overlapping entities. Each Road is linked to a unique designator (road name or number), and its extent is defined by the collective geometry of its constituent RoadLinks.

A diagram of a road

Description automatically generated

## RoadLink

A RoadLink is the fundamental linear segment between two TransportNodes (to and from junctions) and is represented using the transinfras:RoadLink class. Roadlinks are grouped together to form a Road.

Each RoadLink is characterized by:

* A unique identifier (ORN\_ROAD\_NET\_ELEMENT\_ID), mapped to the internal OGF\_ID in the ontology to ensure consistent referencing.
* Its capacity and the maximum number of lanes defined in the link
* A RoadLinkUser characteristics that can define characteristics such as the speed limit or average travel time for the link based on the type of user.

A diagram of a roadlink

Description automatically generated

## Ontario Road Network Tags Mapping

**Data Provided by ORN Shapefile:**

|  |  |  |  |
| --- | --- | --- | --- |
| ORN Data Tag | ORN Data Description | Type | Equivalence |
| OGF\_ID | A unique numeric provincial identifier assigned to each object. | Int | cdt:ogfID |
| FROM\_JCT | The beginning junction for a road element or ferry connection. | Int | transnet:from |
| TO\_JCT | A statement that identifies the positional accuracy of the ORN road geometry in metres. | Int | transnet:to |
| ACCURACY | A statement that identifies the positional accuracy of the ORN road geometry, in metres. | Int | cdt:roadAbsoluteAccuracy |
| NID | A unique national identifier assigned to a road net element, junction and selected event data such as Toll Point, Blocked Passage and Structure which are required to support the National Road Network (NRN). | String | cdt:nationUUID |
| LENGTH | The measured planimetric length of a road net element in meters. | Int | cityunits:length |
| DIRECTION | The direction(s) of vehicular or motor traffic flow. All road elements must have a direction of traffic flow assigned. | String | cdt:trafficFlow |
| EXIT\_NUM | The number of an exit on or off a freeway, expressway or highway, assigned by an administrating body and is represented by a valid number or character. | String | cdt:exitNumber |
| ELEM\_TYPE | An attribute describing the type of road net element. | String | cdt:roadElementType |
| TOLL\_ROAD | Indicates if the road net element is a toll road. | String | cdt:tollRoadIND |
| ACQTECH | The type of data source or technique used to create or revise the road net element. | String | cdt:aquisitionTechnique |
| CREDATE | The date the road net element was originally created. | Date | transalert:creationTime |
| REVDATE | The date the road net element was last revised or updated. | Date | transalert:lastUpdateTime |
| GEO\_UPD\_DT | Date/time the geometry was created or last modified in the source database. | Date | cdt:geoUpdateDate |
| EFF\_DATE | Date/time the record was created or last modified in the source database. | Date | cdt:effectiveDate |

**Data Provided by ORN csv Files:**

**Note:** All the following CSV files had a data tag ORN\_ROAD\_NET\_ELEMENT\_ID, an Integer representing a system-generated identifier unique at the application level.

|  |  |  |  |
| --- | --- | --- | --- |
| Data Provided by ORN\_SPEED\_LIMIT.csv: | | | |
| ORN Data Tag | ORN Data Description | Type | Equivalence |
| SPEED\_LIMIT | The maximum speed limit assigned to a road element in kilometres per hour in accordance with Municipal By-Laws or Provincial Law. | Int | infras:speedLimit |

ORN\_SPEED\_LIMIT.csv: The maximum speed limit assigned to a road element in kilometres per hour in accordance with Municipal By-Laws or Provincial Law. In cases where a road element has more than one speed limit value, the speed limit of the longest portion of the road element is supplied.

|  |  |  |  |
| --- | --- | --- | --- |
| Data Provided by ORN\_ROAD\_CLASS.csv: | | | |
| ORN Data Tag | ORN Data Description | Type | Equivalence |
| ROAD\_CLASS | The classification of a road. | String | contact:hasStreetType |

ORN\_ROAD\_CLASS.csv: A linear event identifying the class of road based on a functional classification schema.

|  |  |  |  |
| --- | --- | --- | --- |
| Data Provided by ORN\_OFFICIAL\_STREET\_NAME.csv: | | | |
| ORN Data Tag | ORN Data Description | Type | Equivalence |
| FULL\_STREET\_NAME | This attribute is derived from the individual street name components where present, namely directional prefix, street type prefix, street name body, street type suffix and directional suffix and is stored in upper case text. | String | contact:hasStreet |

ORN\_OFFICIAL\_STREET\_NAME.csv: An event identifying an official street name and may be associated with a bilingual name. A language code shall be used to specify the applied language which resides in the Street\_Name\_Parsed Table.

|  |  |  |  |
| --- | --- | --- | --- |
| Data Provided by ORN\_BLOCKED\_PASSAGE.csv: | | | |
| ORN Data Tag | ORN Data Description | Type | Equivalence |
| BLOCKED\_PASSAGE\_TYPE | A man-made or natural barrier or access restriction placed on a road net element to control or limit access to a road net element. | String | cdt:blockedPassage |

ORN\_BLOCKED\_PASSAGE.csv: A point event on a road element identifying the existence of an access barrier or an obstruction, either man-made or natural, which controls or limits access to a road element

|  |  |  |  |
| --- | --- | --- | --- |
| Data Provided by ORN\_JURISDICTION.csv: | | | |
| ORN Data Tag | ORN Data Description | Type | Equivalence |
| JURISDICTION | An indication of who has the jurisdictional, or custodianship responsibility for a road net element. The custodian would have the responsibility to ensure maintenance occurs, but is not necessarily the one who undertakes the maintenance directly. | String | cdt:jurisdiction |
| STREET\_SIDE | The side of the street for which the addressing applies. | String | cdt:streetSide |

ORN\_JURISDICTION.csv: Identifies jurisdictional, or custodianship responsibility of the road

|  |  |  |  |
| --- | --- | --- | --- |
| Data Provided by ORN\_ROAD\_SURFACE.csv | | | |
| ORN Data Tag | ORN Data Description | Type | Equivalence |
| SURFACE\_TYPE | A linear event indicating the surface type of a road element. | String | cdt:surfaceType |
| PAVEMENT\_STATUS | The surface type of a road element. | String | cdt:pavementStatus |

ORN\_ROAD\_SURFACE.csv: The surface type of a road element.

|  |  |  |  |
| --- | --- | --- | --- |
| Data Provided by ORN\_NUMBER\_OF\_LANES.csv | | | |
| ORN Data Tag | ORN Data Description | Type | Equivalence |
| NUMBER\_OF\_LANES | The number of lanes of a road. | Int | infras:numLanes |

ORN\_NUMBER\_OF\_LANES.csv: A linear event indicating the number of lanes. 

|  |  |  |  |
| --- | --- | --- | --- |
| Data Provided by ORN\_ROUTE\_NAME.csv | | | |
| ORN Data Tag | ORN Data Description | Type | Equivalence |
| ROUTE\_NAME\_ENGLISH | The English name that is attached to a road net element as defined by a Municipality, Provincial Ministry, or Federal Agency and is associated to an established and/or maintained route. | String | cdt:routeName |

ORN\_ROUTE\_NAME.csv: The name attached to a road net element as defined by a Municipality, Provincial Ministry, or Federal Agency and is associated to an established and/or maintained route.

|  |  |  |  |
| --- | --- | --- | --- |
| Data Provided by ORN\_ROUTE\_NAME.csv | | | |
| ORN Data Tag | ORN Data Description | Type | Equivalence |
| ROUTE\_NUMBER | The route number assigned to a road typically associated with provincial highways, secondary highways, county roads and regional roads and is represented by a numeric and/or an alpha-numeric character. A road can be assigned multiple route numbers. | String | cdt:routeNumber |

ORN\_ROUTE\_NUMBER.csv: The route number attached to a road net element as defined by a Municipality, Provincial Ministry, or Federal Agency and is typically associated with provincial highways, secondary highways, county roads and regional roads

|  |  |  |  |
| --- | --- | --- | --- |
| Data Provided by ORN\_STRUCTURE.csv | | | |
| ORN Data Tag | ORN Data Description | Type | Equivalence |
| STRUCTURE\_TYPE | The classification of a structure, that exists on a road element and is managed as a linear event. | String | cdt:structureType |

ORN\_STRUCTURE.csv: The classification of a structure, that exists on a road element and is managed as a linear event. The types are mutually exclusive.

|  |  |  |  |
| --- | --- | --- | --- |
| Data Provided by ORN\_TOLL\_POINT.csv | | | |
| ORN Data Tag | ORN Data Description | Type | Equivalence |
| TOLL\_POINT\_TYPE | A point event on a road element identifying the existence of an underpass. An underpass occurs where the road element runs underneath a passage accommodating the movement of water, a building, road, rail, pedestrian or wildlife. options are: 1. Physical - A construction along or across a road element where toll can be paid to employees of the organization in charge of collecting the toll or to machines involving electronic methods of payment like credit cards or bank cards 2. Virtual - At a virtual point along a road element, toll will be charged via automatic registration of the passing vehicle by subscription or invoice 3. Hybrid - A tollbooth, which is both physical and virtual. Valid values: 'Hybrid', 'Physical', 'Virtual' | String | cdt:tollPointType |

ORN\_TOLL\_POINT.csv: A point event along a road element indicating the presence of a toll point.

|  |  |  |  |
| --- | --- | --- | --- |
| Data Provided by ORN\_UNDERPASS.csv | | | |
| ORN Data Tag | ORN Data Description | Type | Equivalence |
| UNDERPASS\_TYPE | Identifies the type of underpass present at this road location. | String | cdt:underpassType |

ORN\_UNDERPASS.csv: A point event on a road element identifying the existence of an underpass. An underpass occurs where the road element runs underneath a passage accommodating the movement of water, a building, road, rail, pedestrian or wildlife.

## Integration of ORN Data in TTL Generation

BASE\_URI = ...

GEO = Namespace(...)

TRANSPORT = Namespace(...)

…

* Defines RDF namespaces used to classify data according to ISO/IEC 5087 standards and custom extensions (CDT).

road\_network\_gdf = gpd.read\_file(shapefile\_path)

road\_network\_gdf["OGF\_ID"] = …

* Loads the ORN Road Net Element shapefile.
* Converts OGF\_ID to string so it can be merged with external attribute tables.

toronto\_bounds = {...}

* Defines the bounding box for filtering geometry and junctions within the Toronto area.

csv\_files = {...}

...

road\_network\_gdf = road\_network\_gdf.merge(...)

* Loads ORN supplementary data: road names, jurisdiction, pavement, structure, etc.
* Renames columns to avoid conflicts.
* Merges each CSV into the main GeoDataFrame using the road network element ID.
* Each feature in the ORN\_ROAD\_NET\_ELEMENT.shp file is uniquely identified by an OGF\_ID — this serves as the **primary key** for the shapefile.
* Meanwhile, the supplementary CSV files (e.g., speed limits, road names, number of lanes, etc.) use a corresponding column named ORN\_ROAD\_NET\_ELEMENT\_ID to reference the road network elements they describe. This acts as a **foreign key** pointing back to the shapefile data.

g = Graph()

g.bind(...)

* Sets up an RDFLib Graph and binds prefixes for Turtle serialization.

for \_, row in data\_frames["junctions"].iterrows():

    ...

* Processes junctions as instances of TransportNode.
* Associates them with geospatial coordinates (asWKT), filtered by Toronto bounds.
* Adds junctionType as a CDT extension.

def format\_date(value: str) -> str | None:

* Safely converts date strings from ORN format (YYYYMMDDHHMMSS) into standard ISO xsd:date.

road\_groups = road\_network\_gdf.groupby("FULL\_STREET\_NAME\_road\_names")

* Groups all ORN elements sharing the same official road name to build a single Road entity made of multiple RoadLinks.

for road\_name, group in road\_groups:

    …

* For each road:
  + Creates a infras:Road.
  + Adds hasProperPart links to each associated RoadLink.
  + Each use contact:hasStreet property to refer to the name of the road.
* For each ORN element (line geometry):
  + Associates them with geospatial coordinates (asWKT), filtered by Toronto bounds.
  + Creates a infras:RoadLink instance.
  + Adds metadata (speed, lanes, pavement, etc).
  + Links to source and destination junctions.
  + Attaches location as geo:asWKT.

g.serialize(destination=output\_file, format="turtle")

* Serializes the RDF graph to a Turtle file ready for semantic reasoning or ingestion into a triple store.