An Education Ontology for Global City Indicators (ISO 37120)

Mark S. Fox
Enterprise Integration Laboratory and Global Cities Institute
University of Toronto
5 King's College Road, Toronto ON, M5S 3G8
University of Toronto

msf@eil.utoronto.ca

EIL Working Paper, First Published: 28 November 2014; Revised: 16 May 2015

1. Introduction

Global City Indicators[1] (GCI) refer to city metrics defined by the Global City Indicator Facility at the University of Toronto (McCarney 2012; McCarney 2013), and refined, extended and published as ISO 37120. ISO 37120 is a global standard comprised of over 100 city indicators each having a definition and methodology. The GCIs span areas such as Education, Energy, Health, Safety, Finance and Shelter. This paper defines an ontology for representing the Education indicators defined in ISO 37120. It builds on our prior research in foundation ontologies for representing city Indicators and their meta-data (Fox, 2013).

Why build an ontology? The development of ISO 37120 represents a significant step forward in the global standardization of how we measure city performance. Key to its development and success is the precise definition of what each indicator means. But the measurement of city performance does not end with an agreement on an English definition of each indicator, it is the beginning. We now have to take the next step of creating a computer-based representation of indicators and their supporting information so that the indicators are replicable, auditable, and truly comparable (Hoornweg et al., 2007).

Consider the ISO 37120 educational indicator 2.4: "Primary Education Student/Teacher Ratio." On the surface the definition is simple, being the ratio of the number of students to the number of teachers, but the process of representing the definition reveals greater complexity:

- The indicator is the ratio of two numbers whose units and scale must be the same (measurement theory).
- The number of students (numerator) and teachers (denominator) are cardinal measures of two different sets (measurement theory).

[1] "Global City Indicators©" is a term created by the Global City Indicators Facility in 2010 at the University of Toronto. All rights apply.

- The sets are based on a population defined within a geographic area (geolocation/placename).
- The populations being sampled are determined by a definition of a student or teacher (description logic).
- A student is defined as a full time student in primary school (description logic).
- Administrative staff are not to be included in the teachers counted (description logic).
- A Primary School has to be a public school that teaches primary grades (description logic).

The indicator "Student/Teacher Ratio" is the root of a dependency tree where the supporting definitions and data branch out below it. The tree is heterogeneous in that its nodes span various types of representations including analytical, statistical, spatial and logical. In addition the tree must represent meta-information such as the processes used to derive the data, its validity and trust. In order to do this, we need to design an ontology.

A second issue is automating performance analysis. Within the context of cities, a major interest is longitudinal analysis, i.e., analyzing the changing performance of a city over time, and transversal analysis, i.e., analyzing the differences in performance of two or more cities. Our interest lies in creating a theory of comparative analysis with heterogeneous models that can be used to identify the root causes of differences of a system (e.g., city) over time or between systems. In order to do this, we again need an ontology.

In the remainder of this paper we first reprint the Education indicators defined in ISO 37120. Adopting the ontology engineering methodology of (Gruninger & Fox, 1995), for each indicator we define a set of competency questions the ontology must be able to answer. We then review how existing vocabularies and ontologies represent education related concepts to determine whether they satisfy our competency requirements. The next section introduces our Education ontology, followed by a demonstration of how the ISO 37120 education indicators are represented using it. Finally, we evaluate the ontology from a competency perspective.

2. Indicators and their Competency Requirements

In this section we reprint the Education indicators as defined in ISO 37120. For each indicator we define a set of competency questions, motivated by each indicator that our education ontology must be able to answer. Note that questions that refer to measurement theory, provenance, validity and trust are not included as they are addressed in the GCI foundation ontology. Competency questions fall into the following categories:

- **Factual (F)**: Questions that ask what the value of some property is.
- **Consistency - Definitional (CD)**: Determine whether the instantiation of an indicator by a city is consistent with the ISO 37120 definition.
- **Consistency - Internal (CI)**: Determine whether different parts of the instantiation are consistent with each other.
- **Deduced (D)**: A value or relationship that can be deduced form the instantiation.

2.1. Percentage of female school-aged population enrolled in schools (ISO37120:6.1)

The first ISO 37120 education indicator focuses on measuring female education:

> "The percentage of female school-aged population enrolled in schools shall be calculated as the number of female school-aged population enrolled at primary and secondary levels in public and private schools (numerator) divided by the total number of female school-aged population (denominator). The result shall then be multiplied by 100 and expressed as a percentage.
>
> The definitions of primary and secondary school detailed in Clauses 3.5 and 3.6 shall apply.
>
> The proportion of enrolment in public and private schools should be reported, and cities shall note if private school data are included. In many cities, private schools are a significant component of education in the city. Private schools shall be recognized as providing real, bona fide education; many ministries or departments of education have a program that recognizes such schools. Enrolment in religious schools and home schools should be included if they are recognized.
>
> One part-time enrolment of a half-day or more shall be counted as a full-time enrolment.
>
> If the geographies of school districts and the city are different, best judgment should be used to related enrolment data to the city boundaries."

Competency Questions

1. (F) What city is the indicator for?
2. (CD) Are the students residents of the city?
3. (D) What is the age range for school age women?
4. (F) Is a school a private or public institution?
5. (F) Does a school teach Primary or Secondary courses?
6. (F) Is a school a home school? Religious school?
7. (D) Is the private school certified by the government?
8. (F) What grades comprise primary and secondary school?
9. (F) How many hours of school do you have to attend to be full time?
10. (D) What school did person X attend in year Y?
11. (D) What proportion of the students are in private schools for school year x?

2.2. Percentage of Students Completing Primary Education: Survival Rate (ISO37120:6.2)

Following is the ISO 37120 definition of Percentage of Students Completing Primary Education. We assume the definition of student and primary school as provided in section 2.1.

> "The percentage of students completing primary education or survival rate shall be calculated as the total number of students belonging to a school-cohort who complete the final grade of primary education (numerator) divided by the total number of students belonging to a school-cohort, i.e. those originally enrolled in the first grade of primary education (denominator). The result shall then be multiplied by 100 and

expressed as a percentage. The survival rate of primary education shall be expressed as the percentage of a cohort of students enrolled in the first grade of primary education who reached the final grade of primary education.

Survival rates for the private education sector should be reported, if known. The user of this International Standard should note if private school data are included."

"Example: If the city reporting year is 2012 and primary education last five years, report the percentage of students entered primary education in 2006 and reached the final grade of primary education in 2011."

Competency Questions

We extend the competency questions in section 2.1 to include the following:

1. (F) What grades are included in primary school?
2. (D) What students in final primary year X are cohorts?
3. (D) If a student was in their first grade of primary school in year X, what would be their final year in primary school?
4. (F) How many students started first grade of primary school in year X?
5. (D) How many students whose first grade in primary school was year X, were in the final grade of primary school?
6. (D) What percentage of students who survived were in private school?

2.3. Percentage of Students Completing Secondary Education: Survival rate (ISO37120_6.3)

Following is the ISO 37120 definition of Percentage of Students Completing Secondary Education. We assume the definitions of student and secondary school are as defined in section 2.1.

"The percentage of students completing secondary education or survival rate shall be calculated as the total number of students belonging to a school-cohort who complete the final grade of secondary education (numerator) divided by the total number of students belonging to a school-cohort, i.e. those originally enrolled in the first grade of secondary education (denominator). The result shall then be multiplied by 100 and expressed as a percentage. The survival rate of secondary education shall be expressed as the percentage of a cohort of students enrolled in the first grade of secondary education who reached the final grade of secondary education."

"Example: If the city reporting year is 2012 and secondary education lasts seven years, report the percentage of students that entered secondary education in 2004 and reached the final grade of secondary education in 2011."

Competency Questions

The competency questions for this indicator are the same as in section 2.2 with the exception of substituting secondary for primary.

2.4. Primary Education Student/Teacher Ratio (ISO37120:6.4)

Following is the ISO 37120 definition of Student Teacher Ratio:

"The student/teacher ratio shall be expressed as the number of enrolled primary school students (numerator) divided by the number of full-time equivalent primary school classroom teachers (denominator). The result shall be expressed as the number of students per teacher.

Private educational facilities shall not be included in the student/teacher ratio.

One part-time student enrolment shall be counted as one full-time enrolment; in other words a student who attends school for half a day should be counted as a full-time enrolment. If a city reports full-time equivalent (FTE) enrolment (where two half day students equal one full student enrolment), this shall be noted.

The number of classroom teachers and other instructional staff (e.g. teachers' aides, guidance counselors), shall not include administrators or other non-teaching staff. Kindergarten or pre-school teachers and staff shall not be included.

The number of teachers shall be counted in fifth time increments, for example, a teacher working one day per week should be counted as 0.2 teachers, and a teacher working three days per week should be counted as 0.6 teachers."

Competency Questions

1. (F) What city is the numerator/denominator for?
2. (CD) Are the numerator, denominator and indicator for the same city?
3. (F) Is the teacher administrative staff or teaching staff?
4. (F) Is the student part time or full time?
5. (F) Did the teacher work at a public school or private school in Year Y?
6. (F) Did the student attend a public school or private school in Year Y?
7. (D) What grades did teacher X teach in year Y?
8. (D) What schools did student X attend in year Y?
9. (D) What public schools are included in the indicator?
10. (D) Are there any private school included in the indicator?

2.5. Percentage of male school-aged population enrolled in schools (ISO37120:6.5)

Following is the ISO 37120 definition of percentage of male school-aged population enrolled in schools:

"The percentage of male school-aged population enrolled at primary and secondary levels in public and private schools (numerator) divided by the total number of male school-aged population (denominator). The result shall then be multiplied by 100 and expressed as a percentage.

The definitions of primary and secondary school detailed in 3.5 and 3.6 shall apply.

Enrolment in public and private schools should be reported, and cities shall note if private school data are included. In many cities, private schools are a significant component of education in the city. Private schools shall be recognized as providing real, bona fide education; many ministries or departments of education have a program

that recognizes such schools. Enrolment in religious schools and home schools should be included if they are recognized.

One part-time enrolment of a half-day or more shall be counted as a full-time enrolment.

If the geographies of school districts and the city are different, best judgment should be used to relate enrolment data to the city boundaries."

Competency Questions

The competency questions for this indicator are the same as in section 2.1 with the exception of substituting male for female.

2.6. Percentage of school-aged population enrolled in schools (ISO37120:6.6)

Following is the ISO 37120 definition of percentage of school-aged population enrolled in schools:

"The percentage of school-aged population enrolled in schools shall be calculated as the number of school- aged population enrolled in primary and secondary levels in public and private schools (numerator) divided by the total number of the school-aged population (denominator). The result shall then be multiplied by 100 and expressed as a percentage.

Enrolment in public and private schools should be reported, and cities shall note in the comment section if private school data are included. In many cities, private schools are a significant component of education in the city. Private schools shall be recognized as providing real, bona fide education; many ministries or departments of education have a program that recognizes such schools. Enrolment in religious schools and home schools should be included if they are recognized.

Part-time enrolment of a half-day or more shall be counted as a full-time enrolment.

If the geographies of school districts and the city are different, best judgement should be used to relate enrolment data to the city boundaries."

Competency Questions

The competency questions for this indicator are the same as in section 2.1 with the exception of substituting all people for female.

2.7. Number of higher education degrees per 100 000 population (ISO37120:6.7)

Following is the ISO 37120 definition of number of higher education degrees per 100,000 population:

"The number of higher education (tertiary education) degrees per 100 000 population shall be calculated as the number of people holding higher education degrees (numerator)

divided by one 100 000th of the city's total population. The result shall be expressed as the number of higher degrees per 100 000 population."

Competency Questions

1. (F) What are tertiary degrees?
2. (F) What degrees does a person have?
3. (D) How many people have a tertiary degree X?
4. (D) How many females/males have tertiary degree X?

3. Background

Our focus is the development of an education ontology that will represent the definition of each education indicator and answer their corresponding competency questions. We build on the Global City Indicator Foundation ontology (Fox, 2013)[2]. In that work, we integrated and extended existing ontologies depicted in Figure 1:

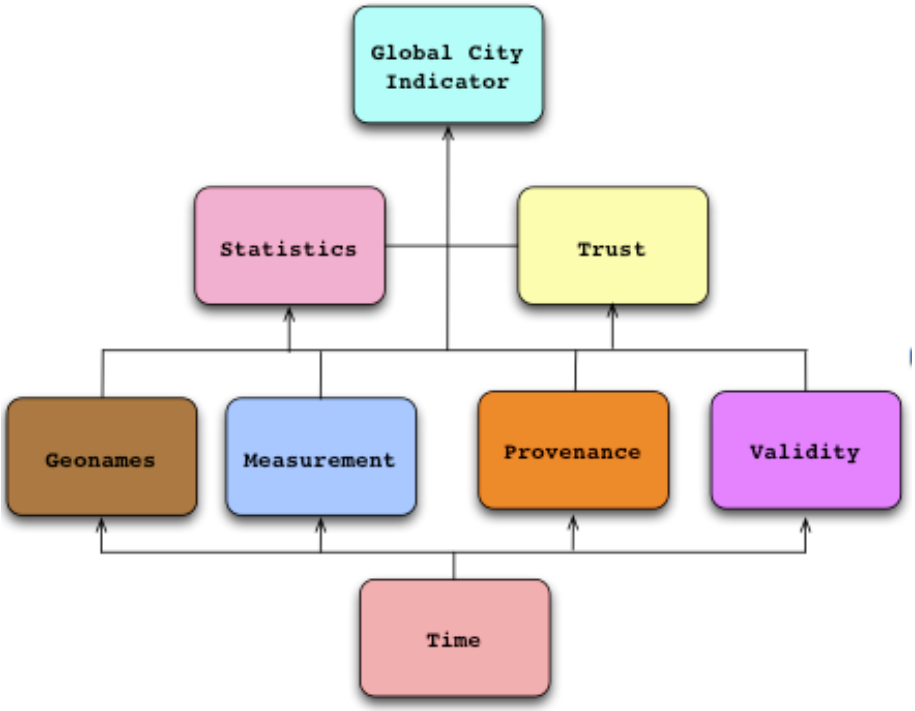


Figure 1: GCI Foundation Ontology Components

The ontologies included are:

- Time (Hobbs & Pan, 2006).
- Measurement (Rijgersberg et al., 2011)
- Statistics (Pattuelli, 2009).

[2] The GCI Foundation ontology can be found at http://ontology.eil.utoronto.ca/GCI/GCI-Foundation.owl along with its documentation at http://ontology.eil.utoronto.ca/GCI/GCI-Foundation.html. We will use the prefix "gci" where needed.

- Validity (Fox & Huang, 2005).
- Trust (Huang & Fox, 2006).
- Placenames (www.geonames.org).

In order to answer the competency questions for the Education indicators, we need additional concepts, properties and axioms that span:

- the types of educational institutions at the primary and secondary level, including whether they are public or private, and certified,
- the grade structure of schools including cohorts,
- the definition of students and what educational programs they have taken, and
- the definition of teachers, including where and what they have taught.

We reviewed a number of ontologies to determine their relevance. Our review was hindered by their lack of competency questions; in most cases the ontologies were published as RDF or OWL with little documentation. Hence a detailed examination of their axioms was required. In this section we identify some of the vocabularies and ontologies that we reviewed. A more detailed analysis of their relevant concepts is provided in the sections that follow.

Schema.org is an initiative primarily led by the major search engine vendors. Its goal is to enhance search results by providing a vocabulary of concepts and properties that web page creators can embed in their web pages using RDFa. Many of the classes defined in the OWL version of schema.org only have subclassof property specified.

SUMO (Niles & Pierce, 2001) is an upper level ontology[3]. It attempts to provide an overarching taxonomy of knowledge. In other words, its taxonomy of concepts is meant to span most of what we may want to represent.

OpenCYC (Matuszek et al., 2006) is a large ontology that is both very broad and very deep. It has been under development for over 15 years. The ontology is very rich in the areas of intelligence/defence.

An ontology has been created for describing the national curricula across the UK (Mohamed et al., 2013). The purpose of the ontology is to:

- "provide a model of the national curricula across the UK",
- "organise learning resources, e.g. video clips and revision content", and
- "allow users to discover content via the national curricula".

It is focused on course content as opposed to the organization and resourcing of its educational system.

Scribe (Uceda-Rosa et al., 2011) is an ontology designed specifically to represent city information. From an education perspective, it refers to Educational Service (a service) and School District (a local government area), but not to schools, grades, teachers nor students.

[3] The SUMO ontology can be found at http://ontologyportal.org/sumo.owl. We will use the prefix "sumo" where needed.

While some of these ontologies provide some basic classes, e.g., school, teacher, and student, they neither provide the full set of class and properties nor the axioms necessary to model educational indicators. It is clear that there is a need for a well engineered educational ontology with competency questions necessary to support them.

4. Architecture of the ISO 37120 Ontology

The following diagram (Figure 2) depicts the organization of files used to define the ISO 37120 ontology we are developing. At the highest level, i.e., ISO 37120 Ontology level, the ISO 37120 module[4] contains the globally unique identifier (IRI) for each ISO 37120 indicator. For example, the IRI for the Student/Teacher Ratio indicator is: "http://ontology.eil.utoronto.ca/ISO37120.owl#6.5".

For each category of indicators in the ISO 37120 specification, for example Education, there is a separate file that provides the definition of each indicator in that category. For example, ISO37120/Education.owl[5] provides a complete OWL definition for all seven of the indicators in the ISO 37120 specification.

The GCI Ontology level provides the category specific ontologies required to define each category's indicators. For example, to define the ISO 37120 Education indicators, we need an educational ontology covering concepts such as schools, teachers, students, cohorts, etc. GCI-Education.owl[6] provides the classes used by ISO37120/Education.owl.

All of the category specific indicator ontologies rely about the GCI Foundation ontology[7] for more generic concepts such as population counts and ratios, meta-information, etc.

[4] http://ontology.eil.utoronto.ca/GCI/37120.owl.
[5] http://ontology.eil.utoronto.ca/GCI/ISO37120/Education.owl.
[6] The GCI Education ontology can be found at http://ontology.eil.utoronto.ca/GCI/Education/GCI-Education.owl along with its documentation at http://ontology.eil.utoronto.ca/GCI/Education/index.html. We will use the prefix "gcie" where needed.
[7] The GCI Foundation ontology can be found at http://ontology.eil.utoronto.ca/GCI/Foundation/GCI-Foundation.owl along with its documentation at http://ontology.eil.utoronto.ca/GCI/Foundation/index.html. We will use the prefix "gci" where needed.

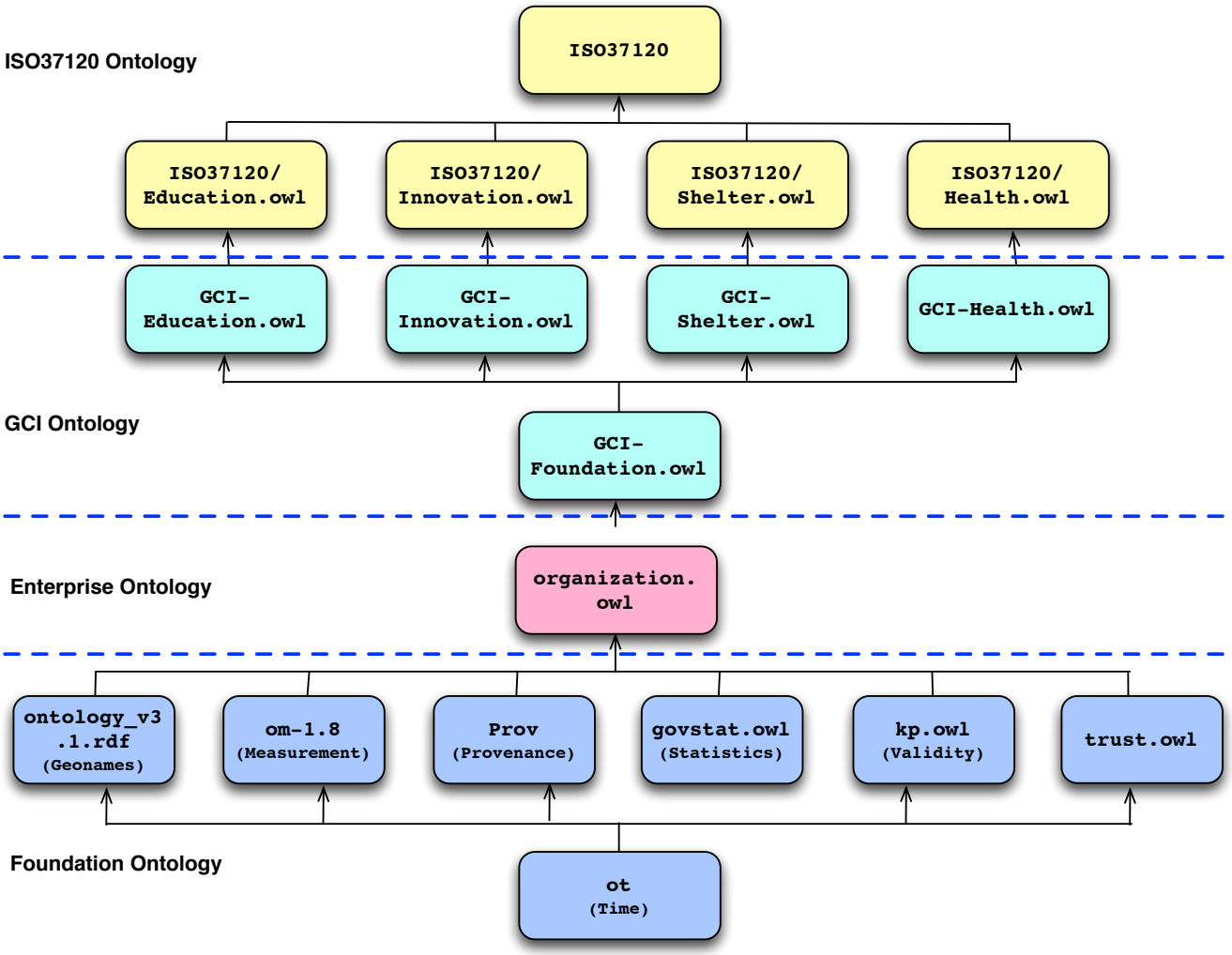


Figure 2: ISO 37120 Ontology Modules

The Enterprise Ontology level contains Enterprise Modelling ontologies. In this figure we only show the Organization Ontology file[8] (Fox et al., 1998), which is one of the TOVE Enterprise Modelling ontologies (Fox & Grüninger, 1998). In addition to the Organization ontology, TOVE has ontologies spanning:

- Activities and States (Gruninger & Fox, 1994)
- Resources (Fadel et al., 1994; Fadel, 1994).
- Quality Measurement (Kim & Fox, 1994).
- Activity-Based Costing (Tham et al., 1994).
- Product (Lin et al., 1997).
- Product Requirements (Lin et al., 1996).
- Human Resources (Fazel-Zarandi & Fox, 2012).

Finally, the Foundation Ontology level provides very basic ontologies that were selected as the foundation for the GCI-Foundation.owl ontology described in section 3.

[8] The Organization ontology can be found at http://ontology.eil.utoronto.ca/organization.owl along with its documentation at http://ontology.eil.utoronto.ca/organization.html. We will use the prefix "org" where needed.

5. GCI Education Ontology

As discussed in the previous section, in order to computationally represent the definitions of the ISO 37120 education indicators and answer their competency questions, we need to add educational concepts not included in the GCI Foundation ontology. This section defines the GCI Education ontology that can found at http://ontology.eil.utoronto.ca/GCI/Education/GCI-Education.owl.

5.1. Teacher and Student Classes

Core to the Education indicators is the description of students and teachers. The following competency questions, selected from section 2, focus on Teacher and Student knowledge.

1. (D) What is the age range for school age women?
2. (CD) Are the students residents of the city?
3. (F) Is the teacher administrative staff or teaching staff?
4. (F) Did the teacher work at a public school or private school in Year Y?
5. (F) Did the student attend a public school or private school in Year Y?
6. (D) What grades did teacher X teach in year Y?
7. (D) What schools did student X attend in year Y?
8. (F) Was the student part time or full time in year Y?

In reviewing existing ontologies, student and teacher definitions are often limited to taxonomic relations. In SUMO, classes do not exist for student nor teacher. OpenCYC (Figure 3) defines a 'teacher' to be a subclass of 'academic' and 'person'. It has a subclass 'schoolteacher' that is further specialized as 'government schoolteacher' that is 'affiliating with regional government'. A 'student' is a subclass of 'person' and has many specializations including 'elementary school student', 'full time student', and 'high school student'. These classes do not contain additional axioms other than a 'high school student' being a 'teenager'.

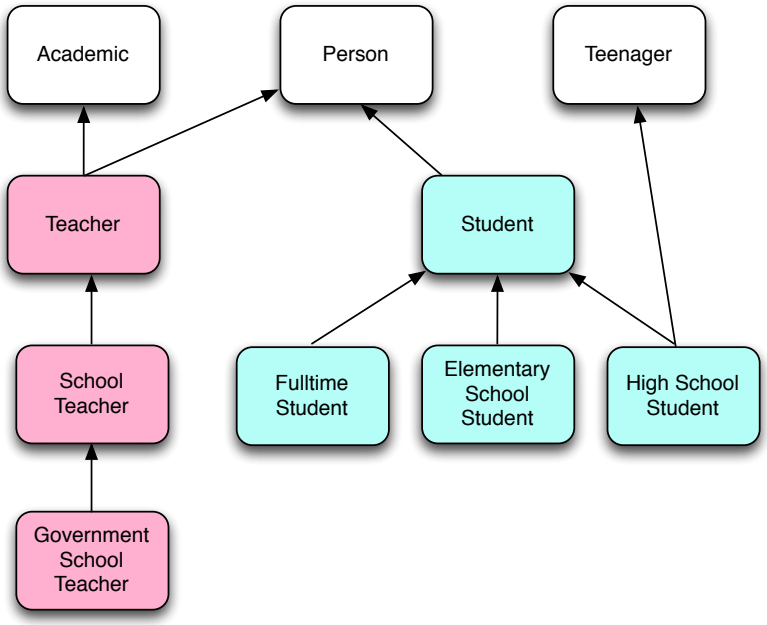


Figure 3: OpenCYC Student and Teacher Taxonomy

In the GCI Education ontology, 'Teacher' is part of an 'Education Staff' taxonomy (Figure 4). The top level class of all education employees is 'Education Staff' which is a subClassOf 'Person' and 'Organization Agent'. It has two subclasses: 'Education Staff Administrative' and 'Education Staff Instructional'. 'Teacher' is subClassOf 'Education Staff Instructional'.

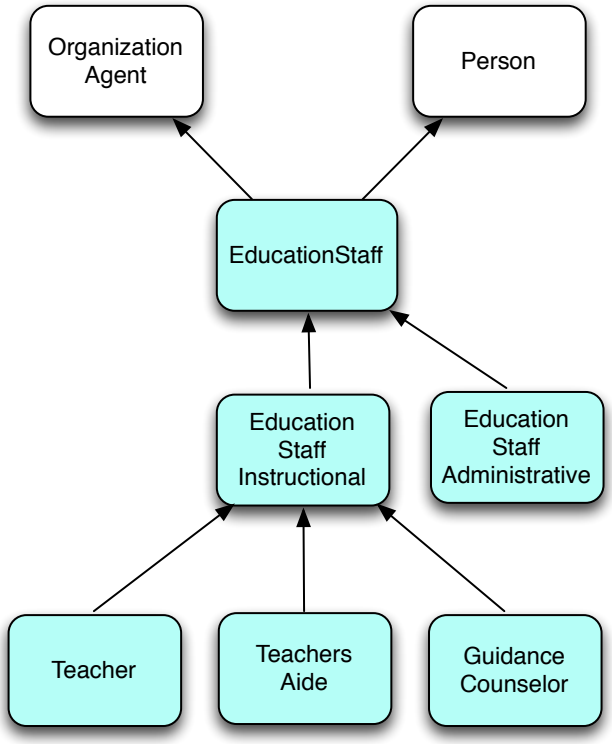


Figure 4: GCI Education Staff Taxonomy

In the following table[9], we define an 'Education Staff' member as 'Organization Agent' (as defined in the Organization ontology) and a 'Person' (as defined in Schema.org) that has at least one 'Placement'. A 'Teacher' is a subclass of 'Instructional Education Staff' that has a 'Placement' in which they teach a minimum of one 'Course'.

Class	Property	Value Restriction
EducationStaff	owl:subClassOf	org:OrganizationAgent
	owl:subClassOf	sc:Person
	has_Placement	min 1 Placement
EducationStaffAdministrative	owl:subClassOf	EducationStaff
EducationStaffInstructional	owl:subClassOf	EducationStaff
Teacher	owl:subClassOf	EducationStaffInstructional
	has_Placement	min 1 (Placement and teaches min 1 Course)

[9] The table defines an OWL 2 (Hitzler et al., 2012) class using the Manchester Syntax (Horridge & Patel-Schneider, 2012).

	owl:equivalentClass	cyc:Teacher

A ‘Placement’ provides the details of where an ‘Education Staff’ member worked, the ‘School Year’, how many days a week they worked, and ‘Course’s they taught, if any. An ‘Education Staff’ member may have many ‘Placements’, one for each year that they worked at a ‘School’, or more than one per year if they worked at multiple locations during the same year.

Class	**Property**	**Value Restriction**
Placement	educational_Staff_At	EducationFacility
	days_Worked	exactly 1 positiveInteger
	for_SchoolYear	exactly 1 SchoolYear
	teaches	min 1 Course
	Min_Days_Worked	value 1

A ‘Student’ is defined to be a ‘Person’ that has been enrolled in one or more ‘Educational Program’s. Each ‘Grade’ they attend is represented as a separate ‘Enrollment’ due to the information that is required to represent it. For example, a ‘Student’ may attend different ‘Grade’s at different ‘School’s, they may be part time in one grade and full time in another, etc.

Class	**Property**	**Value Restriction**
Student	owl:subClassOf	sc:Person
	has_Enrollment	min 1 Enrollment
	has_Birthdate	exactly 1 xsd:dateTime
	owl:equivalentClass	cyc:Student
	has_primary_residence	exactly 1 ic:HomeAddress

Attendance during a school year at a school corresponds to a separate ‘Enrollment’. An ‘Enrollment’ is composed of the ‘Program’ (defined in the next section) the student is enrolled in, an ‘Educational Facility’ they attend, ‘School Year’, ‘Course’s they took, ‘Grade’, and an ‘enrolled Status’ of full or part time.

Class	**Property**	**Value Restriction**
Enrollment	attends	exactly 1 EducationFacility
	enrolled_Program	exactly 1 Program
	for_SchoolYear	exactly 1 SchoolYear
	enrolled_Courses	min 1 Enrolled_Course
	enrolled_Grade	exactly 1 Grade
	enrolled_Status	exactly 1 Enrollment_Status
Enrolled_Course	for_Course	exactly 1 Course
	has_Result	exactly 1 xsd:string
	has_Comment	only xsd:string

‘Enrolled Course’ is defined by identifying the ‘Course’ that was enrolled in, having a result (i.e., ‘Grade’) and some comment.

5.2. Educational Facility and Grade Classes

In this section we define our ontology for educational facilities and grades based on the following competency questions introduced throughout Section 2.

1. (F) Is a school a private or public institution?
2. (F) Does a school teach Primary or Secondary courses?
3. (D) Is the private school certified by the government?
4. (F) What grades comprise primary and secondary school?
5. (D) What students in final primary year X are cohorts?
6. (D) If a student was in their first grade of primary school in year X, what would be their final year in primary school?
7. (F) How many students started first grade of primary school in year X?
8. (D) How many students whose first grade in primary school was year X, were in the final grade of primary school?
9. (D) What percentage of students who survived were in private school?
10. (F) Did the teacher work at a public school or private school in Year Y?
11. (F) Did the student attend a public school or private school in Year Y?
12. (D) What grades did teacher X teach in year Y?
13. (D) What schools did student X attend in year Y?
14. (D) What public schools are included in the indicator?
15. (D) Are there any private school included in the indicator?

Our competency questions require the distinction between private and public, secondary and primary schools. They also require identifying the education programs they provide, the grades that make up each level, and whether students are enrolled in them.

Schema.org's only relevant class is 'School', which is a subclass of 'EducationalOrganization', and inherits the following properties (http://www.schema.org/School) from 'Organization' which do not address the needs of the competency questions

- Address
- aggregateRating
- brand
- contactPoint
- department
- duns
- email
- employee
- employees
- event
- events
- faxNumber
- founder
- founders

SUMO has a class 'EducationalOrganization' (no axioms provided) that is a subclass of 'Organization' whose axioms are:

- members of the same 'Organization' share the same purpose, and
- that a member of an 'Organization' is an instance of 'Agent'.

SUMO defines 'School' as a subclass of 'EducationalOrganization'. Within 'School' it has 'HighSchool', 'PrivateSchool', 'PublicSchool' and 'SecondarySchool' as subclasses. They have the following axioms associated with them:

- 'PrivateSchool' is disjoint from a 'GovernmentOrganization'.

- ‘PublicSchool’ is a subclass of ‘GovernmentOrganization’.

OpenCYC has ‘school’ defined as a subclass of ‘educational organization’, that is a subclass of ‘institution’. It has a specialization ‘K-12 institution’ which in turn has specializations of ‘elementary school’, ‘middle school’ and ‘high school’. Finally, a ‘K-12 institution that is a publically funded thing’ is equivalent to the intersection of a ‘K-12 institution’ and a ‘publicly funding thing’. Note the similarity of Schema.org and SUMO to OpenCYC; Schema.org acknowledges portions of their taxonomy are based on CYC.

In the GCI Education ontology, we have imported the Organization ontology (Fox et al., 1998), which provides the concepts of Organization, Goal, Activity and Member. Organization is specialized into the following sub classes: “Non Government Organization”, “For Profit Organization” and “Government Organization”, the latter being used to define publically funded schools.

The basic taxonomy of ‘School’s is depicted in Figure 5.

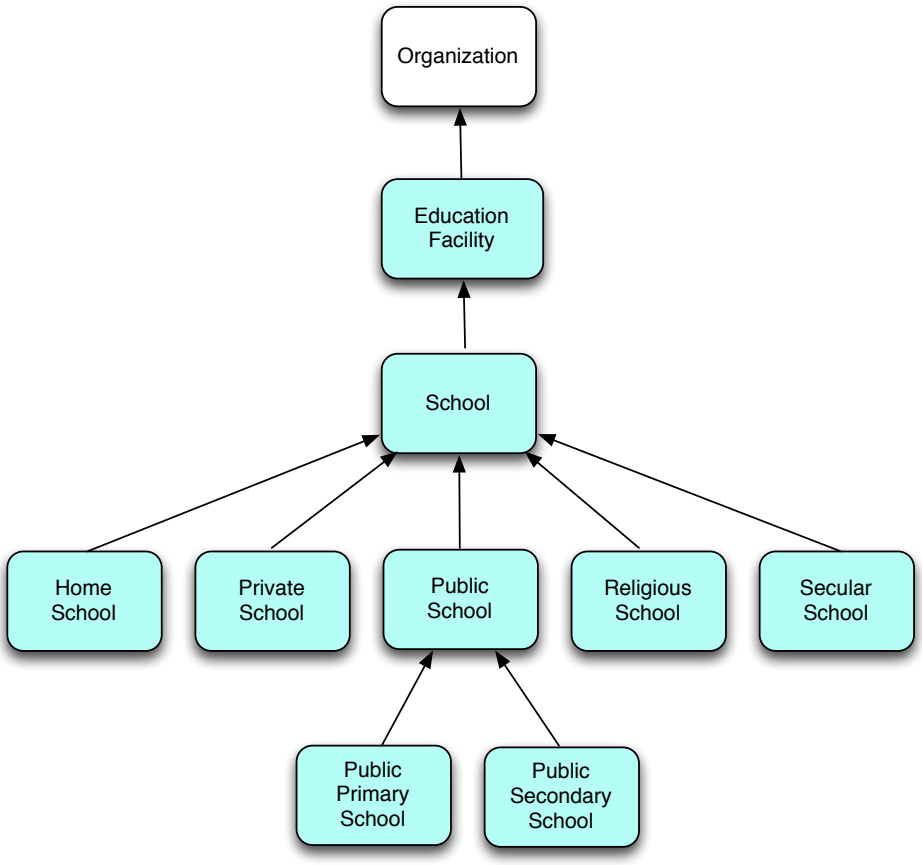


Figure 5: Education Facility Taxonomy

School has the following properties:

1. delivers_Program that identifies the type of ‘School Program’ that is taught,
2. org:has_Ownership that distinguishes among, public, private, government and charity ownership, and

3. has_SchoolType that distinguishes among religious, secular, home, French Immersion, etc.

Class	Property	Value Restriction
School	owl:subClassOf	EducationFacility
	delivers_Program	some SchoolProgram
	org:has_Ownership	exactly 1 Ownership
	has_SchoolType	min 1 SchoolType
	org:hasName	only xsd:string
	org:consistsOf	only org:Division
	org:hasLegalName	exactly 1 xsd:string
	org:hasGoal	only org:Goal

The following defines the 'Public Primary School' class that teaches 'Grade Level Primary'. This defines the grades being taught at the primary level for the corresponding city. The choice of "some" is to allow a school to teach other things than primary grades, e.g., swimming lessons.

Class	Property	Value Restriction
PublicPrimarySchool	owl:subClassOf	PublicSchool[10]
	delivers_Program	some GradeLevelPrimary
	Has_SchoolType	value secularSchoolType
	has_Ownership	value government_ownership
PrivateSchool	owl:subClassOf	School
	has_Ownership	value privately_owned
	has_Certification	some Certification
Certification	certified_By	some GovernmentOrganization
	certification_Date	exactly 1 dateTime

A 'Program' is anything that requires 'Certification'. It also defines what it means to be "Fulltime' in terms of the number of hours required over a designated period of time, such as a 'day', 'week', 'month' or 'year'. A 'School Program' defines the 'Course's that are taught and whether the program is primary, secondary, etc.. 'Grade Level' is a subclass of 'School Program'. The 'Grade Level' class allows each city to define the grades that correspond to primary and secondary school. 'Grade Level' has a starting_Grade and ending_Grade that define the first and last grades of the level. Each city defines its own version of 'Grade Level Primary' that is appropriate for their school system. In the case of Toronto, the starting and ending grades are constrained by the definitions provided by the Province of Ontario. A 'Grade Level' also has a starting and ending age to represent the range of ages that can attend this level of school.

Class	Property	Value Restriction
Program	has_Certification	some Certification

[10] Note that any semantic distinction between private and public organizations would be inherited from existing foundation ontologies and are not defined in the Education ontology.

	has_Fulltime_Hours	exactly 1 positiveInteger
	has_Fulltime_Period	exactly 1 TimePeriod
SchoolProgram	owl:subclassOf	Program
	has_Course	min 1 Course
	has_SP_Type	all SP_Type
GradeLevel	owl:subClassOf	SchoolProgram
	starting_Grade	exactly 1 Grade
	ending_Grade	exactly 1 Grade
	gci:for_City	exactly 1 City
	starting_age	exactly 1 positiveInteger
	ending_age	exactly 1 positiveInteger

The 'Grade' class has subclasses covering all possible grades, e.g., Grade One, Grade Two. Each grade is connected to another via the next_Grade property to define the ordering. It also has sub classes 'Primary Grade' and 'Secondary Grade'.

We link a 'City' to a 'Grade Level' by the following:

Class	**Property**	**Value Restriction**
City	owl:subClassOf	geo:Feature
	has_Primary_Grade_Level	exactly 1 GradeLevelPrimary
	has_Secondary_Grade_Level	exactly 1 GradeLevelSecondary

We also define 'School Age Person' by associating them with a 'City' and 'School Year'. Determining whether someone is of school age is defined by a constraint that uses this information along with their birthdate which is a property inherited from 'Person'.

Class	**Property**	**Value Restriction**
School_Age_Person	owl:subClassOf	Person
	gci:for_City	exactly 1 City
	for_SchoolYear	exactly 1 SchoolYear

The starting grade for all schools in the Province of Ontario at the primary level is 'GradeOne' and the ending grade is 'GradeSix', hence we define:

Class	**Property**	**Value Restriction**
GradeLevelPrimaryCanadaOntario	owl:subClassOf	GradeLevelPrimaryCanada
	starting_Grade	exactly 1 GradeOne
	ending_Grade	exactly 1 GradeSix

We introduce the concept of a 'Cohort', i.e., the students who started primary or secondary school together and entered the final year of each together. For example, for any given year, e.g., 2014, the cohort is defined to be the subset of students who entered the final year of the grade level in 2014, who also were in the starting grade of the grade level together. If primary school covers grades one through six, then the starting year for the 2014 cohort is 2009.

Class	Property	Value Restriction
Cohort	owl:subClassOf	EducationThing
	starting_SchoolYear	exactly 1 SchoolYear
	ending_SchoolYear	exactly 1 SchoolYear
	for_GradeLevel	Exactly 1 GradeLevel

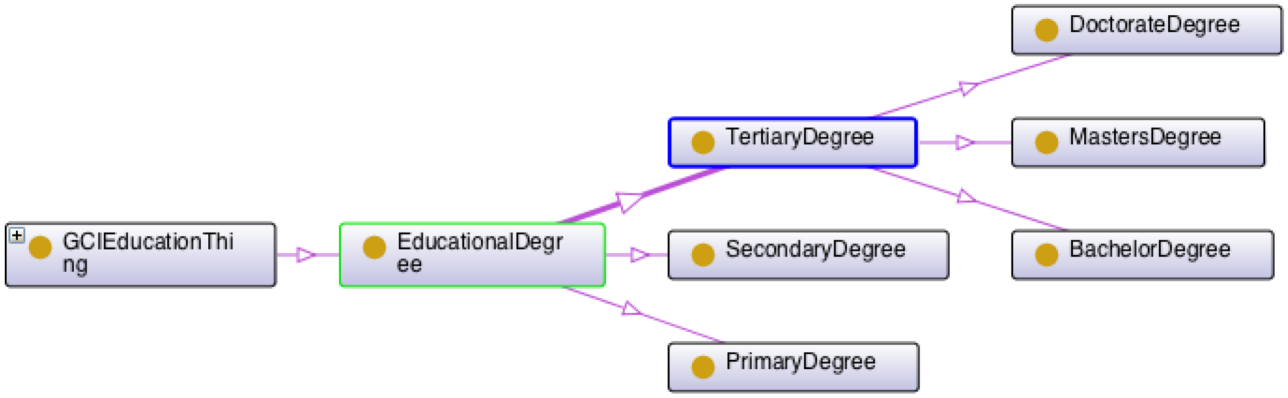
In order to guarantee that the school year of the Education Program matches the school year of the Indicator, we will need to add a consistency axiom in the next section.

5.3. Higher Education Degrees

For the 7th indicator, we need to extend the Education ontology to allow a city to identify what educational degrees count as Tertiary.

1. (F) What are tertiary degrees?
2. (F) What degrees does a person have?
3. (D) How many people have a tertiary degree X?
4. (D) How many females/males have tertiary degree X?

The set of admissible degrees may differ from city to city, but is assumed to be post-secondary. We introduce the concept of Education Degree as follows:



The various types of tertiary degrees can be refined, such as arts, science, engineering, etc. We extend the definition of a Person to include the property: has_EducationDegree, and the 'Education Degree' has the properties:

Class	Property	Value Restriction
EducationDegree	owl:subClassOf	EducationThing
	awarded_Year	exactly 1 Year
	awarding_EducationFacility	exactly 1 EducationFacility
	enrolled_Program	only 1 EducationProgram
	degree_Name	exactly 1 string

The enrolled_Program property allows for the specification of the courses taken in each year of the program.

5.4. Consistency Axioms

The following are additional axioms that cannot be formulated in OWL, but in our implementation are implemented in Prolog:

1. The ending grade of a grade level must be after the starting grade of the same grade level.
2. The starting grade of secondary school is the next grade after the final grade of primary school.
3. All Students who attend a primary school must attend primary grades.
4. All Teachers in a primary school must teach at least one primary course.
5. For students to be a cohort, if they are counted in the final year population then they must be a subset of the students in the first year cohort population.
6. A Grade that is a member of Primary Grade must be contained within the Primary Grade Level (same for Secondary Grade).
7. A student's age must be within the age range of the grade level they are associated with.
8. The difference in years in the start and ending year of a cohort is equal to the difference in years in the starting and ending grade of a grade level.
9. The value of the Educational Program for_School_Year has to be the same as the value for an indicator's for_School_Year.

6. Foundation Ontology Infrastructure

Before we present the education indicators' definitions, we review the basic structure of a ratio indicator, as defined in the GCI Foundation ontology (Fox, 2013), and upon which the education indicators are based.

At the core of the Foundation ontology is the OM measurement ontology (Rijgersberg et al., 2011). The purpose of a measurement ontology is to provide the underlying semantics of a number, such as what is being measured and the unit of measurement. The importance of grounding an indicator in a measurement ontology is to assure that the numbers are comparable, not that they are measuring the same thing, but the actual measures are of the same type, e.g., the counts of the student and teacher populations, that comprise the ratio of student and teacher population sizes, are of the same scale (i.e., thousands vs millions).

Figure 6 depicts the basic classes of the OM ontology used to represent an indicator. There are three main classes in OM: a 'Quantity' that denotes what is being measured, e.g., diameter of a ball; a 'Unit of Measure' that denotes how the quantity is measured, e.g., centimeters; and a 'Measure' that denotes the value of the measurement which is linked to the both 'Quantity' and 'Unit of Measure'. For example, Student Teacher Ratio is a subclass of 'Quantity' that has a value that is a subclass of 'Measure' whose units are a 'Population ratio unit' that is a subclass of 'Unit of Measure'. The actual value measured is a property of the 'Measure' subclass 'Student teacher ratio measure'.

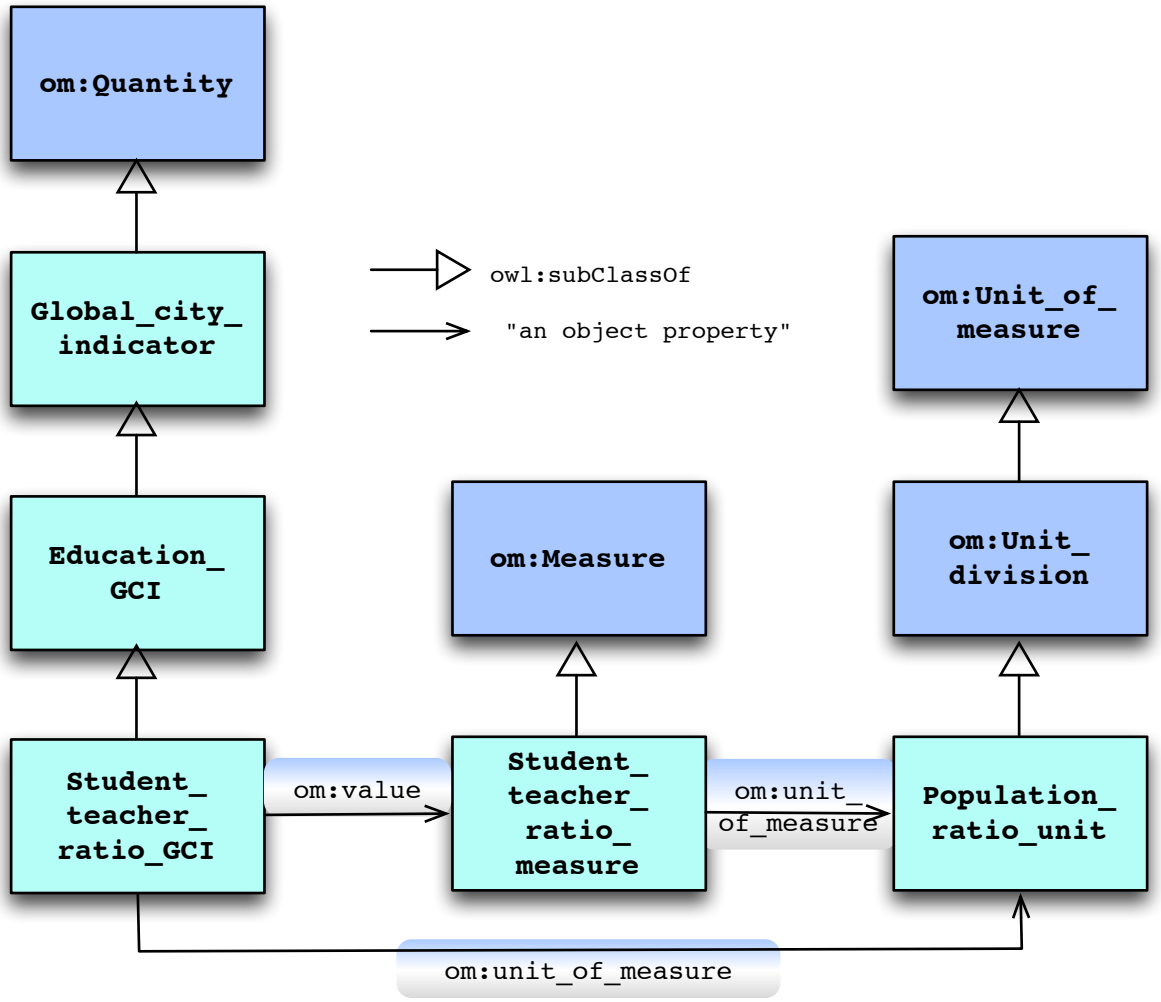


Figure 6: Measurement Ontology

The Student Teacher ratio indicator is based on a measure of the number of students and teachers (that satisfy the indicators' definition of each) within a city's population. One can view both as a statistical measurement in the sense that there is a population that we want to perform a measurement of, the measurement being a count of the number of members that satisfy a description of a Student and a Teacher, respectively. While the indicators require a count of members of the population, other measures may require statistics such as mean, standard deviation, etc. We have included in our core the GovStat[11] general statistics ontology (Pattuelli, 2009). The core class is the 'Population' to be measured. 'Population' is linked to a parameter (e.g., mean, standard deviation) by the is_described_by property, and the parameter is a subclass of 'Parameter'. In order to define the what portion of a city we are determining the size of, we extended the GovStat ontology with a property to located_in, that identifies the area (i.e., city) that the Population is drawn from, and the property defined_by, that identifies the class that all members of the Population are subsumed by.

[11] The GovStat Ontology is not available online, but a version with the GCI extensions can be found at: http://ontology.eil.utoronto.ca/govstat#.

All of the education indicators are ratios. A ratio indicator (Figure 7) has a unit of measure defined to be a 'Population Ratio Unit' that specifies that the indicator is the ratio of the sizes (cardinalities) of two populations. One population size is the numerator and the other the denominator. A 'Population Size' is defined as the cardinality of a 'Population', and 'Population' is defined by a 'City' that the population is located in, and by a description of a 'Person' within the 'City'. For example, the 'Person' could be 'Female Student'. Hence the 'Population Size' could be the number of 'Female Student's in a particular 'City'. This structure is used in the indicator definitions that follow.

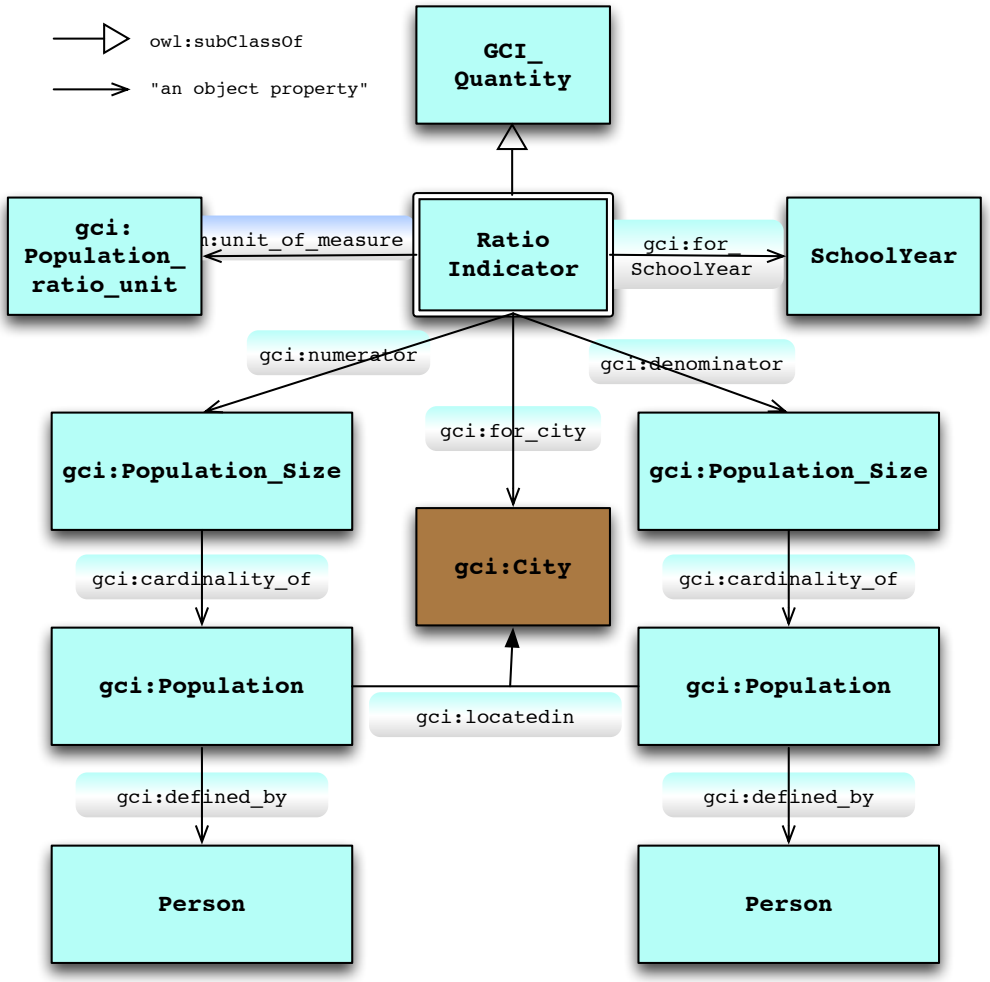


Figure 7: Foundation Ontology Ratio definition

7. ISO 37120 Education Indicators Ontology

With the GCI Education and Foundation ontologies defined, we now have the classes and properties necessary to represent the definitions of the ISO 37120 Education indicators. In this section we represent the seven ISO 37120 Educational indicators. The OWL 2 definitions can be found in http://ontology.eil.utoronto.ca/GCI/ISO37120/Education.owl.

7.1. Percentage of female school-aged population enrolled in schools (ISO37120:6.1)

The following diagram shows a partial definition of ISO37120:6.1. Some of the subClassOf links have been omitted but can be found in the OWL definition file.

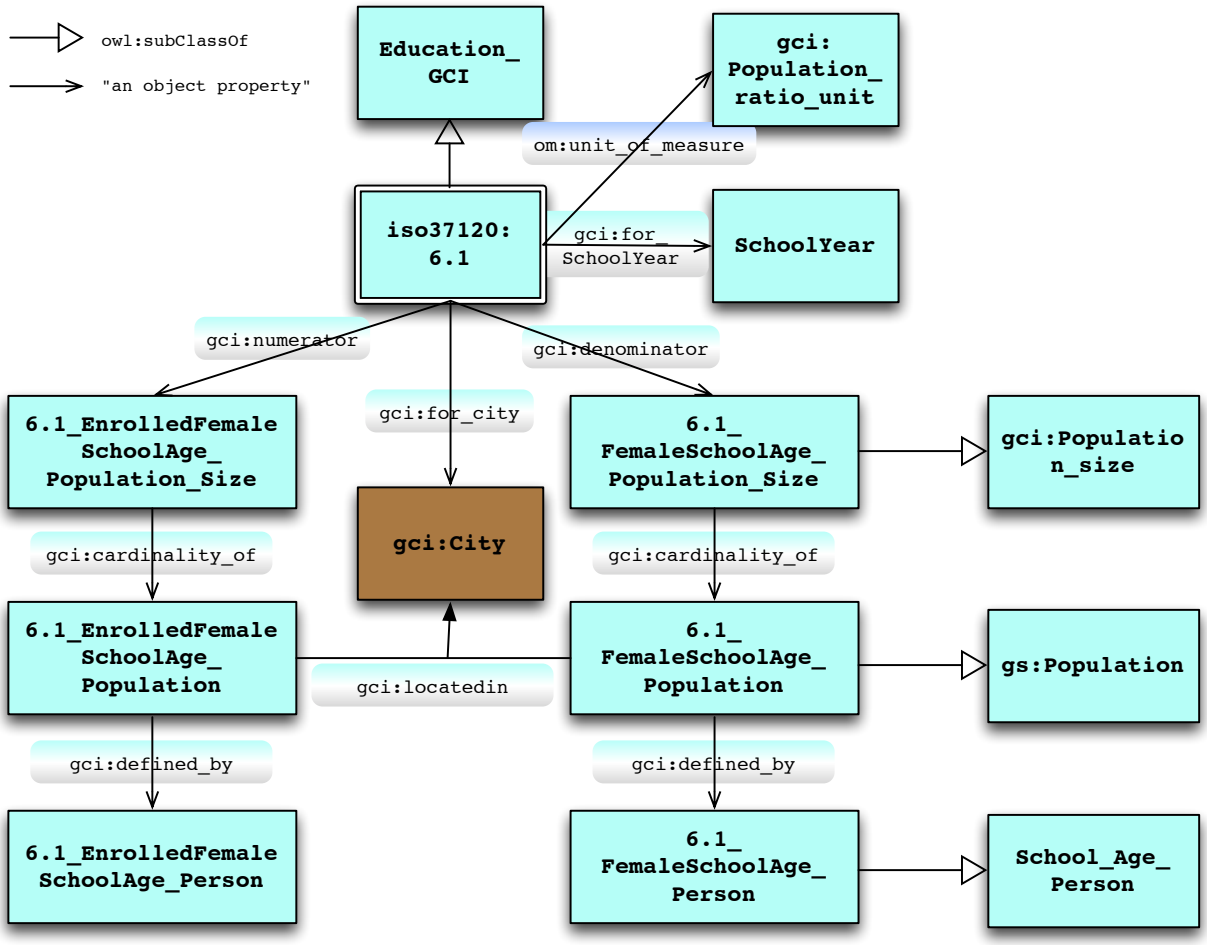


Figure 8: 6.1 Definition

Figure 8 uses the Foundation ontology to provide the "scaffolding" for indicator 6.1. It is an Education Global City Indicator. It is a ratio ('Popultation_ratio_unit') that has a numerator of the size of the population of enrolled school age women. The denominator is the size of the population of all school age women.

What is unique to this indicator is the definition of the people making up the populations (linked using defined_by), namely 'Enrolled Female School Age' and 'Female School Age'. The following defines 'Enrolled Female School Age':

Class	Property	Value Restriction
6.1_EnrolledFemaleSchoolAge_Person	owl:subClassOf	Female_Person
	Owl:subClassOf	School_Age_Person
	has_Enrollment	6.1_Enrollment

There are two issues we have to address in this definition:

1. We have to make sure that for the year the metric is being reported that the student is of school age in that year, and
2. They are enrolled on a full or part time basis, in a public or private school, in a primary or secondary grade.

In the previous section, as part of the 'Grade Level' class, we introduced a starting and ending age. This allows us to determine the age range for both primary and secondary school. To determine whether a 'Person' is in the range, we have to compute their age using their birthdate defined in the 'Student' class and the year for the metric defined by the 'for SchoolYear' property of the '6.1' class. This calculation is performed by an axiom.

The '6.1 Enrollment' class defines the properties of an enrolled 'Student'. Namely, it is for 'School Year' that is the same as 6.1, they attend some 'School', the grade is primary or secondary, they are full or part time and they are enrolled.

Class	Property	Value Restriction
6.1_Enrollment	owl:subClassOf	Enrollment
	for_SchoolYear	exactly 1 (schoolYear_For 6.1)
	attends	exactly 1 School
	enrolled_Program	exactly 1 (GradeLevelPrimary or GradeLevelSecondary)
	enrolled_Grade	exactly 1 (PrimaryGrade or SecondaryGrade)
	enrolled_Status	exactly 1 (Full_Time or Part_Time)
	enrolled_Courses	some Course

To complete this definition we need the following axioms:

1. For the specified year, the age of the 'Student' is within the range defined by the grade levels.
2. The school year of the '6.1_SchoolProgram' is the same as the school year for the '6.1' indicator.
3. The 'School' teaches the enrolled 'Program'.
4. The 'Grade' attended in an 'Enrollment' is consistent with the 'Grade's taught at the corresponding School.
5. The Person counted in each Population resides in the Population's city.

The definition of the denominator can be found in the OWL file.

7.2. Percentage of Students Completing Primary Education (ISO37120:6.2)

This indicator relies upon the definition of 'Cohort'. The basic structure of the ratio is the same as in '6.1', but the definition of the 'Enrolled Primary Ending Grade Person' that defines the population we are taking the size of is where it differs. In particular, it is constrained by the 'Program' they attend being ending year of their grade level. Similarly for the starting grade.

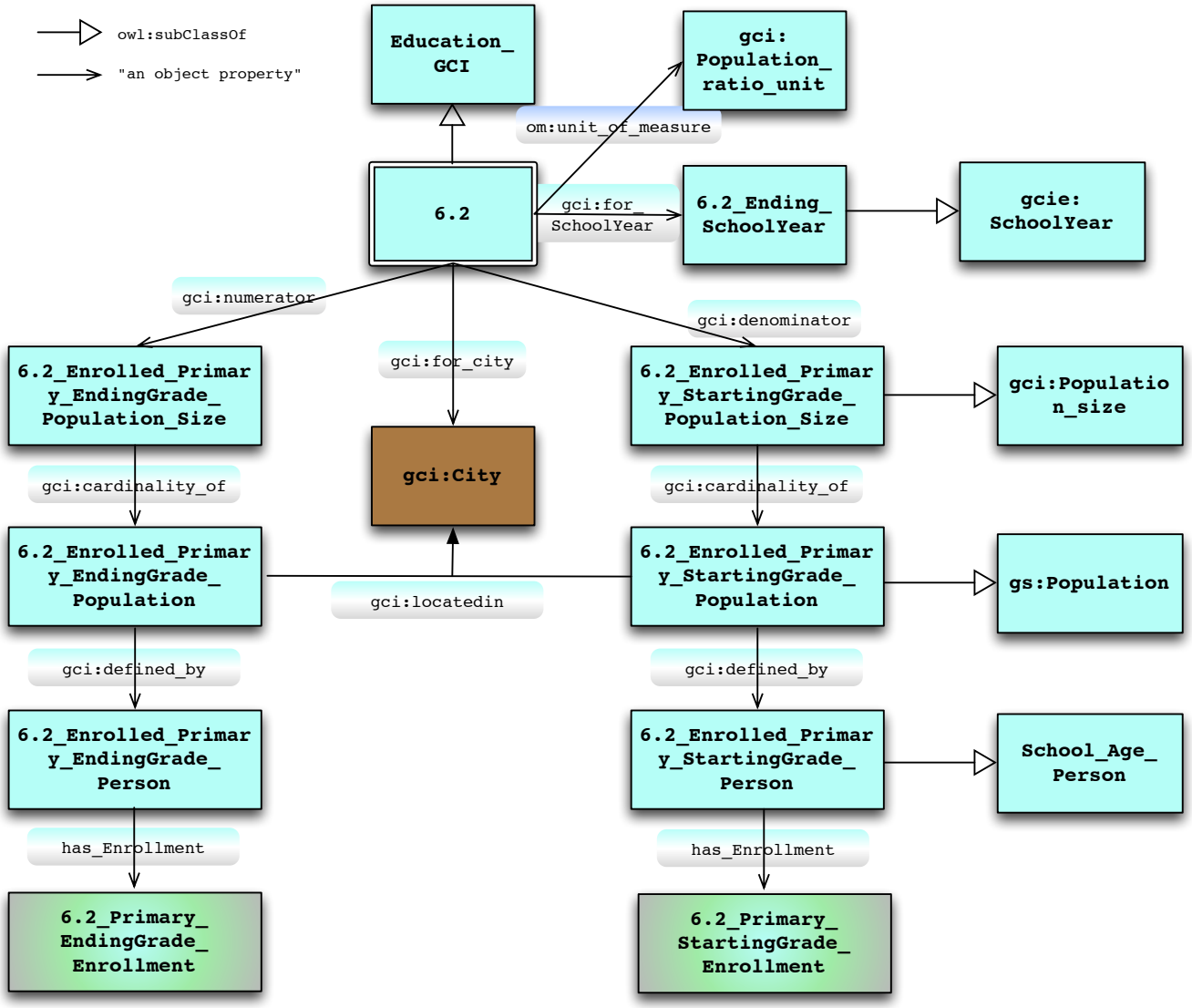


Figure 9: 6.2

The definitions of the Enrolled Primary Starting and Ending Grade Programs are found in Figure 10. The definitions of these 'Program's are where both the 'Grade Level' and 'Cohort' classes come into play.

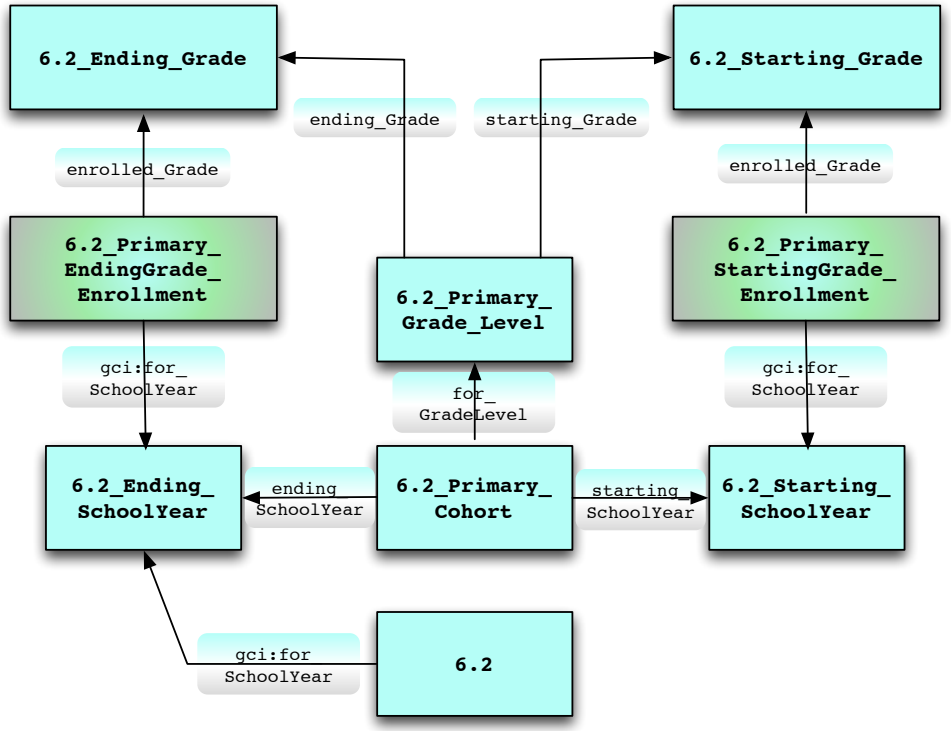


Figure 10: 6.2 Cohort Definition

An 'Enrollment' defines both the 'Grade Enrolled' and the 'Year' of enrollment. To satisfy the definition of '6.2', the Starting and Ending Grades, and the Starting and Ending years have to be consistent with the 'Cohort' specification that includes the 'Grade Level'. In order for this to work properly, we have to define the following axioms:

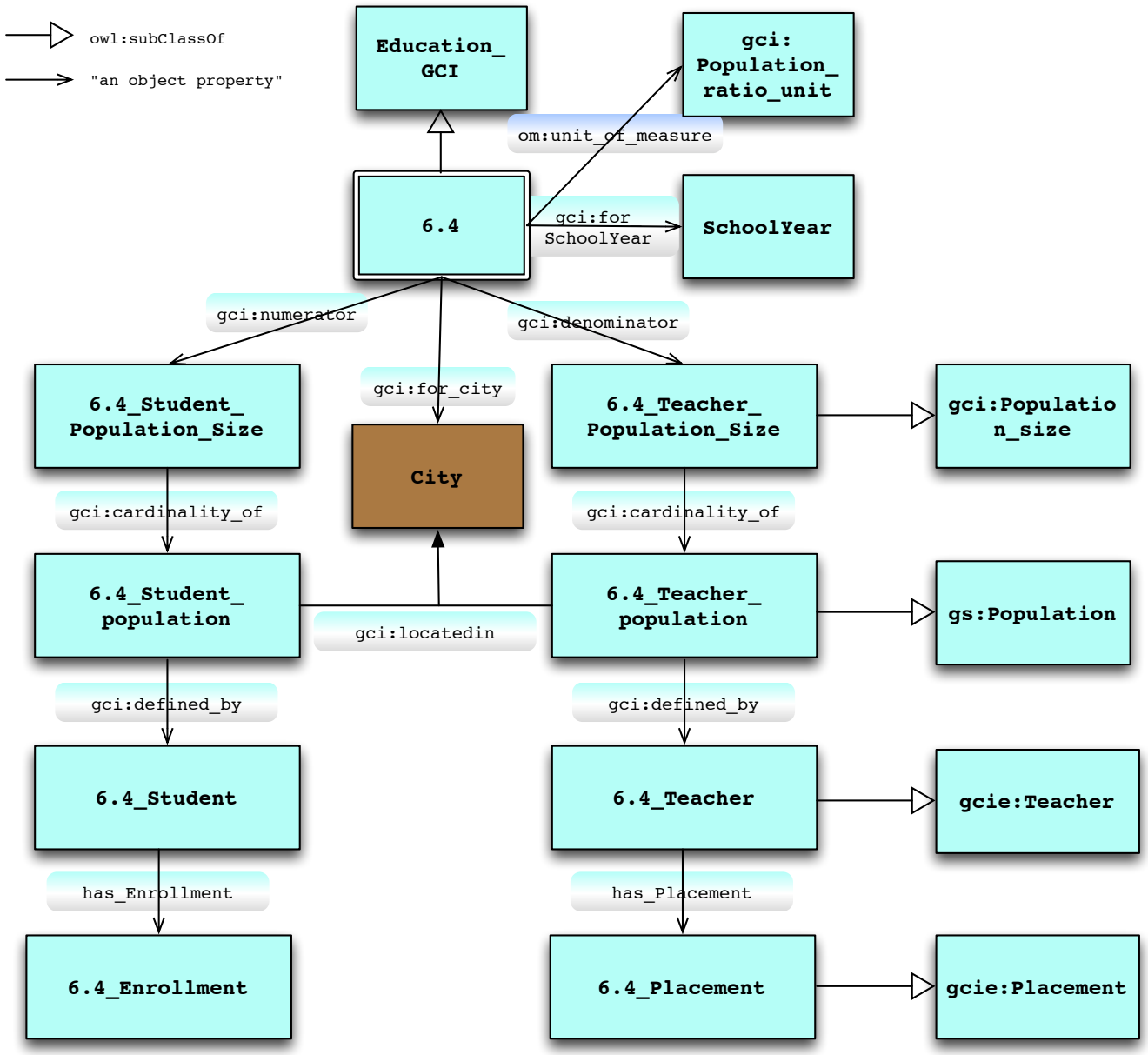
1. The ending school year of the 'Primary Ending Grade Enrollment' is the same as the 'School Year' of '6.2' and the ending school year of the '6.2 Primary Cohort'.
2. The starting school year of the 'Primary Starting Grade Enrollment' is the same as the starting school year of the '6.2 Primary Cohort'.
3. The ending grade of the '6.2 Primary Ending Grade Enrollment' has to be the same as the ending grade of the'6.2 Primary Grade Level'.
4. The starting grade of the '6.2 Primary Starting Grade Enrollment' has to be the same as the starting grade of the '6.2 Primary Grade Level'.

7.3. Percentage of Students Completing Secondary Education (ISO37120:6.3)

6.3's definition is similar to 6.2 except for the substitution of Secondary for Primary. The OWL 2 implementation can be found in:
http://ontology.eil.utoronto.ca/GCI/ISO37120/Education.owl.

7.4. Primary Education Student/Teacher Ratio (ISO37120:6.4)

6.4 has the same structure as 6.1 but varies in the definition of Student and Teacher.



The numerator is the cardinality of the 'Student Population'. The denominator is the cardinality of the 'Teacher Population'. '6.4 Student' is defined to be a subClassOf 'Student'. The restriction that they attend a 'Public Primary School' for the designated 'School Year' is defined by their enrollment:

Class	Property	Value Restriction
6.4_Student	subClassOf	Student
	has_Enrollment	6.4_Enrollment
6.4_Enrollment	owl:subClassOf	Enrollment
	for_SchoolYear	exactly 1 6.4_SchoolYear
	attends	exactly 1 PublicPrimarySchool
	enrolled_Grade	some PrimaryGrade
	enrolled_Status	exactly 1 (Full_Time or Part_Time)

	enrolled_Program	exactly 1 GradeLevelPrimary
	enrolled_Courses	some Course

6.4_Teacher is defined as follows as having at least one Placement in a Public Primary School.

Class	Property	Value Restriction
6.4_Teacher	owl:subClassOf	EducationalStaffInstructional
	has_Placement	6.4_Placement
6.4_Placement	owl:subClassOf	Placement
	days_Worked	exactly 1 positiveInteger
	min_Days_Worked	Value 1
	org:memberOf	PublicPrimarySchool

Axioms

1. Each teacher has to satisfy the minimum days worked requirement.
2. A teacher is counted as one fifth for each day worked.

7.5. Percentage of male school-aged population enrolled in schools (ISO37120:6.5)

This is defined in the same way as 6.1, except for substituting Male for Female. The OWL 2 implementation can be found in: http://ontology.eil.utoronto.ca/GCI/ISO37120/Education.owl.

7.6. Percentage of school-aged population enrolled in schools (ISO37120:6.6)

This is defined in the same way as 6.1, except for removing the Female restriction. The OWL 2 implementation can be found in:
http://ontology.eil.utoronto.ca/GCI/ISO37120/Education.owl.

7.7. Number of higher education degrees per 100 000 population (ISO37120:6.7)

The structure of this indicator is similar to '6.1'. There are two significant differences. First the '6.7' City Population Size has its unit of measure constrained to hectokilopc (100,000) in order to assure that when we take the ratio of number of people with tertiary degrees in the city to total population of the city, it is to 100,000 of population. Second, the definition of the '6.7 Tertiary Degree Resident' is constrained to having a tertiary degree as defined by the city.

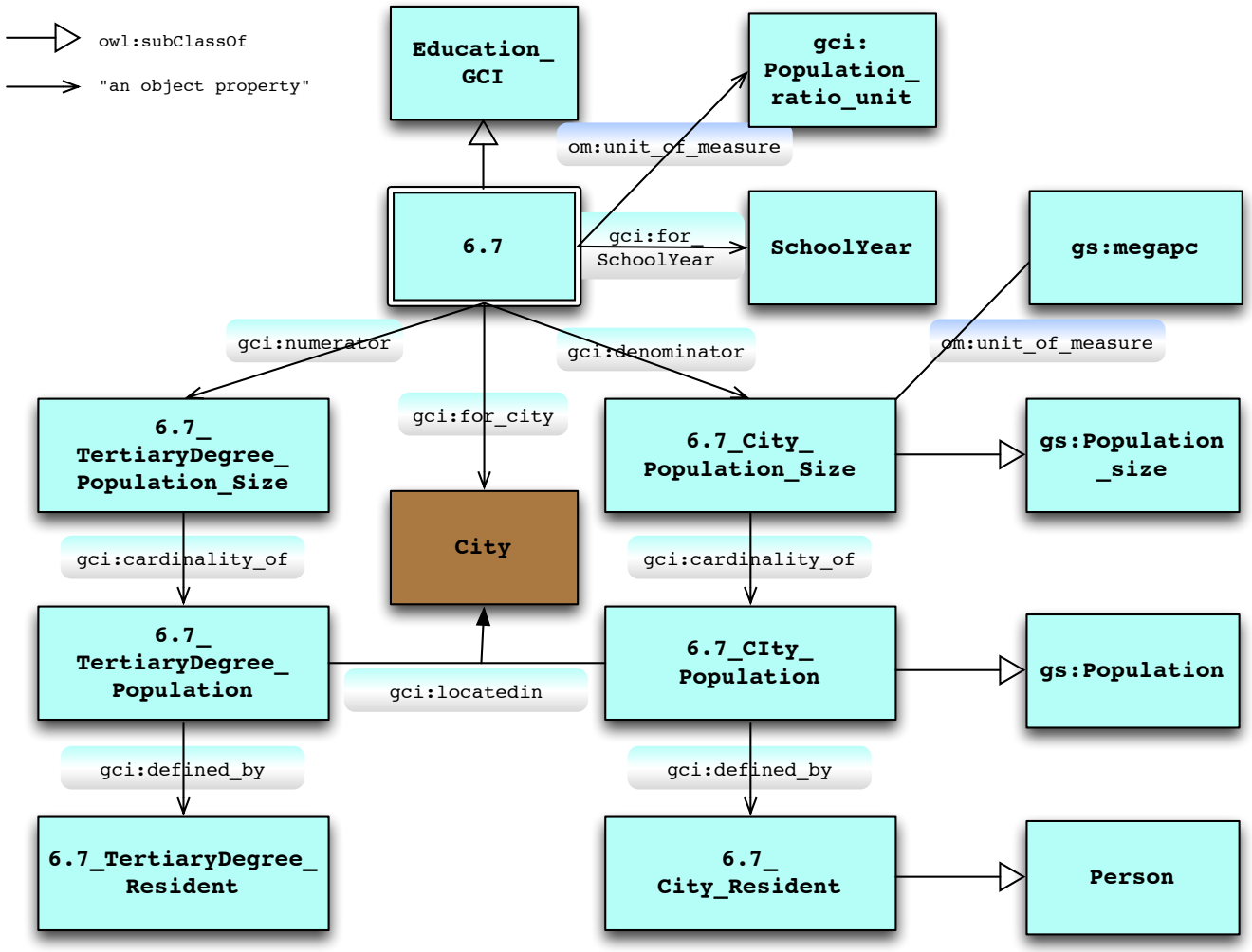


Figure 11: EDU7

The key difference with '6.7' is the definition of the numerator which depends upon the a resident of the city having a tertiary degree:

Class	Property	Value Restriction
6.7_TertiaryDegree _Resident	owl:subClassOf	Person
	has_EducationalDegree	some TertiaryDegree

The following axioms are defined to satisfy the definition:

1. Resident tertiary degrees are restricted to those defined by the city.
2. The date of the degree awarded has to be on or before the year for the indicator.
3. The city that the resident resides in is the same as the city for the indicator.

8. Evaluation

Verification and validation remain research issues in ontology engineering[12]. In this section we verify the Education Ontology by testing its consistency and demonstrating it can be used to answer the competency questions. We then evaluate the ontology by confirming that our definitions of the ISO 37120 Education indicators can be used to test the consistency of city data.

In the next two subsections, we use the City of Toronto in the Province of Ontario, Canada to illustrate the competency questions. For ease of understanding we will show the instances in table form. Prefixes are defined as follows:

- iso: http://ontology.eil.utoronto.ca/ISO37120.owl#
  - URIs for each ISO37120 indicator
- isoe: http://ontology.eil.utoronto.ca/GCI/ISO37120/Education.owl#
  - The ISO37120 education indicators definitions defined in section 7.
- gcie: http://ontology.eil.utoronto.ca/GCI/Education/Education.owl#
  - The Education ontology defined in section 5
- gn: http://sws.geonames.org/
- sc: http://schema.org/
- ic: http://ontology.eil.utoronto.ca/icontact.owl
  - An internationalized address ontology.

This first table defines the instances that provide background information on the city of Toronto, schools, grades, etc.

Instance	Property	Value
gn:6251999	rdfs:label	Canada
	rdfs:type	gn:Feature
	rdfs:type	sc:Country
gn:6093943	rdfs:label	"Ontario"
	rdfs:type	gn:Feature
	rdfs:type	sc:Province
gn:6167865	rdfs:label	"Toronto"
	rdfs:type	gn:Feature
	rdfs:type	sc:City
ontarioPrimaryProgram	rdfs:type	gcie:GradeLevelPrimaryCanada
	gcie:has_Certification	opp_certification
	gcie:has_Fulltime_Hours	35
	gcie:has_Fulltime_Period	om:week
	gn:parentCountry	gn:6251999
	gcie:starting_Grade	ontarioGradeOne
	gcie:ending_Grade	ontarioGradeSix
	gcie:starting_Age	6
	gcie:ending_Age	13
opp_certification	rdfs:type	ProgramCertification
	gcie:certified_By	omet
	gcie:certification_Date	1951-01-01
ontarioGradeOne	rdfs:type	gcie:GradeOne

[12] See Ontology Summit 2013 at http://ontolog.cim3.net/cgi-bin/wiki.pl?OntologySummit2013

	gn:locatedIn	gn:6093943 (Ontario)
ontarioGradeSix	rdfs:type	gcie:GradeSix
	gn:locatedIn	gn:6093943 (Ontario)
cedar_grove	rdfs:type	gcie:PublicPrimarySchool
	gcie:delivers_Program	ontarioPrimaryProgram
	gcie:has_Certification	cg_certification
omet	rdfs:type	GovernmentOrganization
	rdfs:label	“Ontario Ministry of Education and Training”
cg_certification	rdfs:type	SchoolCertification
	gcie:certified_By	omet
	gcie:certification_Date	1951-01-01

The following table defines the instances that instantiate the 6.1 indicator.

Instance	**Property**	**Value**
6.1_ex *(instance of 6.1)*	rdfs:type	iso:6.1
	gci:numerator	6.1_EF_size
	gci:denominator	6.1_F_size
	gci:for_City	gn:6167865
	om:phenomenon	gn:6167865
	om:value	6.1_ex_value
6.1_ex_value *(the value of 6.1)*	rdfs:type	om:Measure
	om:numerical_value	30
	om:unit	gci:Population_Ratio_Unit (change to instance)
6.1_EF_size (numerator of 6.1)	rdfs:type	isoe:6.1_EnrolledFemaleSchoolAge_Population_Size
	gci:cardinality_of	6.1_EF_pop
	om:phenomenon	6.1_EF_pop
	om:value	6.1_EF_size_value
6.1_EF_size_value *(value of the numerator of 6.1)*	rdfs:type	om:Measure
	om:numerical_value	1000
	om:unit	gci:Population_size (change to instance)
6.1_F_size (denominator of 6.1)	rdfs:type	isoe:6.1_FemaleSchoolAge_Population_Size
	gci:cardinality_of	6.1_F_Pop
	om:phenomenon	6.1_F_Pop
	om:value	6.1_F_size_value
6.1_F_size_value *(value of the denominator of 6.1)*	rdfs:type	om:Measure
	om:numerical_value	30000
	om:unit	om:Population_size (change to instance)
6.1_EF_pop *(Numerator population)*	rdfs:type	isoe:6.1_EnrolledFemaleSchoolAge_Population
	gci:locatedin	gn:6167865
	gci:defined_by	6.1_EF_person
6.1_F_pop *(Denominator population)*	rdfs:type	isoe:6.1_FemaleSchoolAge_Population
	gci:locatedin	gn:6167865
	gci:defined_by	6.1_F_person
6.1_EF_person	rdfs:type	isoe:6.1_EnrolledFemaleSchoolAge_Person
	gcie:has_Enrollment	6.1_EF_enrollment
6.1_F_person	rdfs:type	isoe:6.1_FemaleSchoolAge_Person
6.1_EF_enrollment	rdfs:type	isoe:6.1_Enrollment
	gcie:for_SchoolYear	6.1_SchoolYear
	gcie:attends	cedar_grove
	gcie:enrolled_Courses	oc1, oc2, oc3, oc4, oc5, oc6, oc7
	gcie:enrolled_Grade	og1, og2, og3, og4, og5, og6
	gcie:enrolled_Program	ontarioPrimaryProgram

	gcie:enrolled_Status	(fulltime or parttime)
6.1_SchoolYear	rdfs:type	SchoolYear
	starting_Year	2014
jane_smith	rdfs:type	FemaleStudent
	org:memberOf	6.1_F_pop
	org:memberOf	6.1_EF_pop
	has_Primary_Residence	js_home
js_home	rdfs:type	HomeAddress
	ic:has_City	gn:6167865

8.1. Verification

We take two approaches to verification, i.e., what we have implemented conforms to the ontology specifications. The first is to determine whether the ontology is consistent. The consistency of our Education ontology is dependent upon the ontologies it imports. The following diagram depicts the ontology import hierarchy.

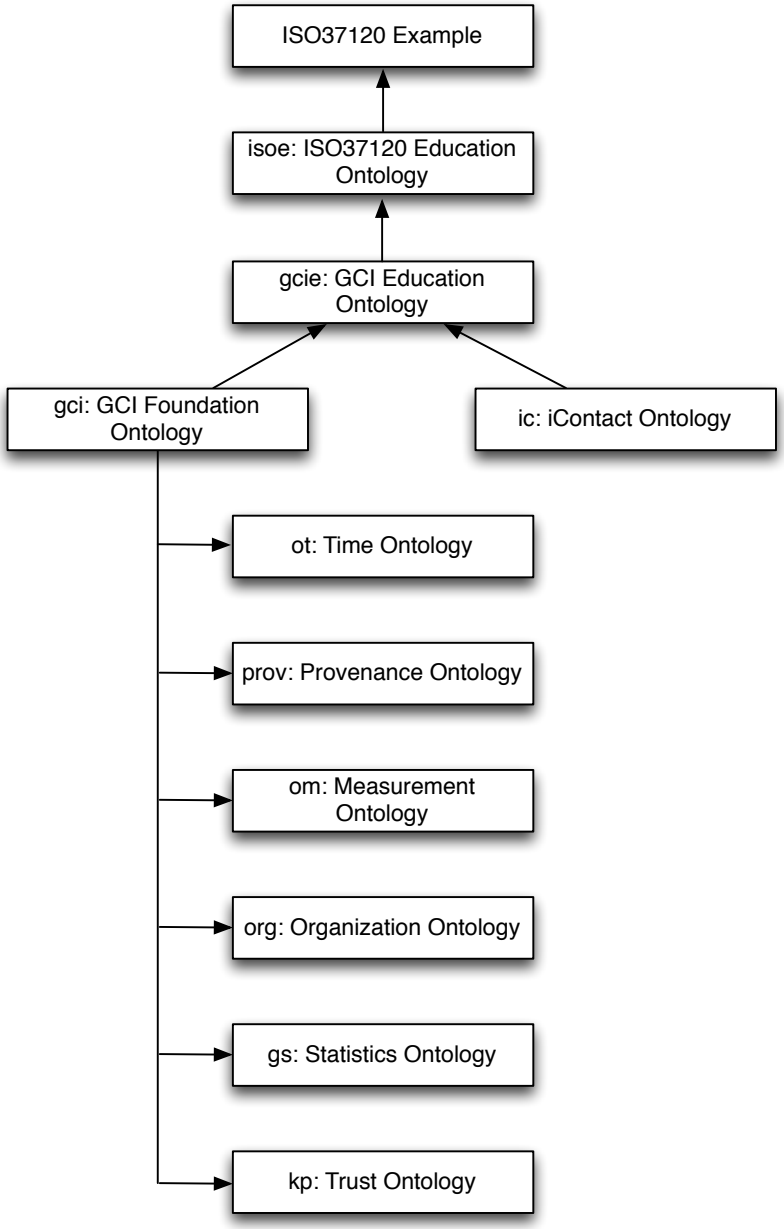


Figure 12: Education Ontology import hierarchy

The following identifies the URI for each of the imported ontologies:

- isoe: http://ontology.eil.utoronto.ca/GCI/ISO37120/Education.owl#
  - The ISO37120 education indicators definitions defined in section 7.
- gcie: http://ontology.eil.utoronto.ca/GCI/Education/Education.owl#
  - The Education ontology defined in section 5
- gci: http://ontology.eil.utoronto.ca/GCI/Foundation/Foundation.owl#
  - The Foundation ontology defined in (Fox, 2013)
- ic: http://ontology.eil.utoronto.ca/icontact.owl#
  - International contact ontology
- ot: http://www.w3.org/2006/time#

    - Time ontology
- prov: http://www.w3.org/ns/prov#
    - Provenance ontology
- om: http://www.wurvoc.org/vocabularies/om-1.8/
    - Measurement ontology
- org: http://ontology.eil.utoronto.ca/organization.owl#
    - Organization ontology
- gs: http://ontology.eil.utoronto.ca/govstat.owl#
    - Statistics ontology
- kp: http://ontology.eil.utoronto.ca/trust.owl#
    - Trust and validity ontology

Using Protégé's Hermit reasoner, we can test an ontologies consistency. In this case, the ontologies in Figure 12 were found to be consistent.

The second approach we have taken to verification is to use competency questions as specifications. The following shows how the consistency questions for '6.1' are implemented in SPARQL.

1. (F) What city is the indicator for?

   SELECT ?city WHERE {6.1-ex gci:for_City ?city}

2. (CD) Are the students residents of the city?
   Identifies each student that is a member of the Enrolled Femail Population and checks to see if their primary residence is the same city as the indicator instance.

   SELECT ?student WHERE
   { 6.1_ex gcie:for_City ?city .
   ?student org:memberOf 6.1_EF_pop .
   ?student gcie:has_Primary_Residence ?PR .
   ?PR ic:has_City ?city }

3. (D) What is the age range for school age women?

   SELECT ?start ?end WHERE
   { 6.1_ex gci:for_City ?city .
   ?city gcie:has_Primary_Grade_Level ?pgl .
   ?pgl gcie:starting_age ?start .
   ?city gcie:has_Secondary_Grade_Level, ?sgl .
   ?sgl gcie:ending_age ?end }

4. (F) Is a school a private or public institution?

   SELECT ?status WHERE { cedar_grove org:has_Ownership ?status }

5. (F) Does a school teach Primary or Secondary courses?

   SELECT ?ctype WHERE { cedar_grove gcie:has_SPType ?ctype }

6. (F) Is a school a home school? Religious school?

   SELECT ?type WHERE { cedar_grove gcie:has_SchoolType ?type }

7. (D) Is the private school certified by the government?

   SELECT ?govorg WHERE
   { cedar_grove has_Ownership privately_owned.
   cedar_grove has_Certification ?cert.
   ?cert certified_By ?govorg.
   ?govorg subclassOf GovernmentOrganization }

8. (F) What grades comprise primary (and secondary) school?
   In order to answer this question properly, we would have to loop through the grades from the starting to the ending grade. We do not show that looping here.

   SELECT ?sgrade ?egrade WHERE
   { ?gradelevel for_City toronto.
   ?gradelevel starting_Grade ?sgrade.
   ?gradelevel ending_Grade ?egrade. }

9. (F) How many hours of school do you have to attend to be full time?
   The following will print out the hours and period for every program associated with the school cedargrove.

   SELECT ?hours ?period WHERE
   { cedargrove delivers_Program ?program.
   ?program has_Fulltime_Hours.
   ?program has_Fulltime_Period ?period }

10. (D) What school did person X attend in year Y?
    We answer this for a specific person, johnsmith, for the primary grade level for school year 2010.

    SELECT ?school WHERE
    { johnsmith has_Enrollment ?enrol.
    ?enrol enrolled_Program ?gradelevel.
    ?gradelevel subclassOf GradeLevelPrimary.
    ?enrol for_SchoolYear 2010.
    ?enrol attends ?school }

11. (D) What proportion of the students are in private schools for school year x?
    The following will return a count of students who enrolled in a primary grade level and taught at a private school for year 2010.

    SELECT (COUNT(?studentpriv) AS ?Num) WHERE
        { ?studentpriv has_Enrollment ?enrol.
          ?enrol enrolled_Program ?gradelevel.
          ?gradelevel subclassOf GradeLevelPrimary.
          ?enrol for_SchoolYear 2010.
          ?enrol attends ?school.
          ?school has_Ownership privately_owned }

8.2. Validation

Validation refers to whether software (in this case ontology) fulfills its intended purpose. Our research has two goals:

1. To determine that the data provided by a city is consistent with the definitions provided in ISO37120, and
2. To determine the root causes for why a city's indicator changes over time (i.e., longitudinal analysis), or why it differs from another city (i.e., transversal analysis).

In this section we focus on using the educational ontology for consistency analysis. Root cause analysis will be addressed in future research.

There are two types of consistency we are concerned with. The first has to do with whether the data submitted by a city for a specific indicator, in the form of OWL instances/individuals, is consistent with the definition of the indicator. In other words, does each individual contain all of the necessary properties and satisfy the property restrictions defined in the class it is a member of?

We implemented a set of prolog rules that determine whether an individual is consistent with the class it is a member of. Given an individual and its corresponding class, the rules determine whether:

- the individual contains all of the necessary properties as defined by the class it is a member of, and
- the corresponding value for the individual's property is consistent with the restrictions defined by the class for that property.

Running the example through the rule set detected both types of errors. Though these errors are conceptually simple and easy to fix, we believe that they will represent a significant portion of the errors we will find in real data. Note that these rules are independent of whether the data is about indicators or not.

The second type of consistency is depicted by definitional constraints that cannot be represented in OWL. Examples of these constraints appear at the each of each subsection in Section 7. For example, at the end of Section 7.1, the following constraints are identified:

1. For the specified year, the age of the ‘Student’ is within the range defined by the grade levels.
2. The school year of the ‘6.1_SchoolProgram’ is the same as the year for the ‘6.1’ indicator.
3. The ‘School’ teaches the enrolled ‘Program’.
4. The ‘Grade’ attended in an ‘Enrollment’ is consistent with the ‘Grade’s taught at the corresponding School.

Each of these constraints are represented as prolog rules and can be applied to any individuals/instances of 6.1 indicator data.

9. Conclusions

Upon embarking on the development of an Education ontology for representing ISO37120 education indicators, it was not expected how broad and deep an ontology would be required. In order to represent what appear to be rather simple indicators, an ontological infrastructure spanning educational institutions, programs, certification, cohorts, etc. was required. It was also unexpected that the existing education ontologies, at least the ones we could find, would provide very little of what was needed.

In summary, this research makes four contributions:

1. Defines an education ontology that is broader and deeper than existing education ontologies, but still focused on supporting the definition of ISO37120 education indicators;
2. Defines each ISO37120 education indicator using the foundation and education ontologies, thereby providing a computationally precise definition;
3. Publishes the ISO37120 education indicator definitions using Semantic Web standards, thereby making it possible to reason about the definitions and instances using existing ontology tools; and
4. Demonstrates that the ontology-based definitions of indicators can be used to automatically validate that indicator data supplied by cities conforms (or not) to the indicator definitions.

10. Acknowledgements

This research is sponsored by the Natural Science and Engineering Research Council of Canada. We acknowledge the support of the Global Cities Institute, University of Toronto.

11. References

Fadel, F.G., (1994), "A Resource Ontology for Enterprise Modelling", M.A.Sc. Thesis, Enterprise Integration Laboratory, University of Toronto.

Fadel, Fadi G., Fox, M.S., and Gruninger, M. (1994) "A Resource Ontology for Enterprise Modelling", *Proceedings of the Third Industrial Engineering Research Conference*, Institute of Industrial Engineers, pp. 455-460.

Fazel-Zarandi, M., Fox, M.S., (2012), "An Ontology for Skills and Competency Management." *Proceedings of the 7th International Conference on Formal Ontologies in Information Systems (FOIS 2012)*, Graz, Austria

Fox, M.S., Barbuceanu, M., Gruninger, M., and Lin, J., (1998), "An Organisation Ontology for Enterprise Modeling", In *Simulating Organizations: Computational Models of Institutions and Groups*, M. Prietula, K. Carley & L. Gasser (Eds), Menlo Park CA: AAAI/MIT Press, pp. 131-152.

Fox, M.S., and Grüninger, M., (1998), "Enterprise Modelling", *AI Magazine*, AAAI Press, Fall 1998, pp. 109-121.

Fox, M.S., and Huang, J., (2005), "Knowledge Provenance in Enterprise Information", *International Journal of Production Research,* Vol. 43, No. 20., pp. 4471-4492. http://www.eil.utoronto.ca/km/papers/fox-ijpr05.pdf

Gruninger, M., and Fox, M.S. , (1994), "An Activity Ontology for Enterprise Modelling", *Workshop on Enabling Technologies - Infrastructures for Collaborative Enterprises* , West Virginia University.

Grüninger, M., and Fox, M. S., (1995), "Methodology for the Design and Evaluation of Ontologies." *Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing*, IJCAI-95, Montreal, Canada.

Hobbs, J.R., and Pan, F., (2006), "Time Ontology in OWL", http:www.w3.org/TR/owl-time/.

Hoornweg, D., Nunez, F., Freire, M., Palugyai, N., Herrera, E.W., and Villaveces, M., (2007), "City Indicators: Now to Nanjing", World Bank Policy Research Working Paper 4114.

Hitzler, P., et al., (2012), "OWL 2 Web Ontology Language Primer (2nd Edition)", http://www.w3.org/TR/owl-primer.

Horridge, M., and Patel-Schneider, P.F., (2013), "OWL 2 Web Ontology Language Manchester Syntax (2nd Edition), W3C Working Group Note 11 December 2012, http://www.w3.org/TR/owl2-manchester-syntax/.

Huang, J., and Fox, M.S, (2006), "An Ontology of Trust – Formal Semantics and Transitivity," *Proceedings of the International Conference on Electronic Commerce,* pp. 259-270. http://www.eil.utoronto.ca/km/papers/huang-ec06.pdf

ISO37120, (2014), "ISO 37120: Sustainable Development of Communities – Indicators for City Services and Quality of Life", *International Organization for Standardization*, First Edition, 2014-05-15, ISO37120:2014(E).

Kim, H., and Fox, M.S., (1994), "Formal Models of Quality and ISO 9000 Compliance: An Information Systems Approach", *American Quality Congress (AQC) Conference*, American Society for Quality Control, Las Vegas NV.

Lin, J., Fox, M.S., and Bilgic, T.,, (1997), "A Product Ontology", Enterprise Integration Laboratory Technical Report .

Lin, J., Fox, M.S., and Bilgic, T., (1996), "A Requirement Ontology for Engineering Design", *Concurrent Engineering: Research and Applications* , Vol. 4, No. 4, pp. 279-291, Sept. 1996.

Matuszek, C., et al., (2006), “An Introduction to the Syntax and Content of Cyc”, In *Proceedings of the 2006 AAAI Spring Symposium on Formalizing and Compiling Background Knowledge and Its Applications to Knowledge Representation and Question Answering*.

McCarney, P. L., (2012), “Global City Indicators (GCI)©: Standardization of City Services and Quality of Life Indicators, Methodologies and Definitions”, Draft submission to the International Organization for Standardization (ISO) for the standardization of the Global City Indicators and corresponding definitions and methodologies entitled: Global City Indicators (GCI): Standardization of City Services and Quality of Life Indicators, Methodologies and Definitions©.

McCarney, P. L., (2013), “Why Cities? Why Metrics? – Tracking Cities in the 21st Century” Global Cities Institute Working Papers, Vol. 1 No. 1, pp. 1-32, University of Toronto.

Mohamed, A., Liu, D., Summerfield, J., Rissen, P., Lee, R., and Rose, Z., (2013), “Curriculum Ontology”, http://www.bbc.co.uk/ontologies/curriculum/, 2013-04-03.

Niles, I. & Pease, A., (2001), “Towards a Standard Upper Ontology”, In *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems FOIS-2001*, Ogunquit, Maine. New York: ACM Press.

Pattuelli, M.C., (2003), “The GovStat Ontology: Technical Report”. The GovStat Project, Integration Design Laboratory, School of Information and Library Science, University of North Carolina at Chapel Hill, http://ils.unc.edu/govstat/papers/govstatontology.doc.

Rijgersberg, H., Wigham, M., and Top, J.L., (2011), “How Semantics can Improve Engineering Processes: A Case of Units of Measure and Quantities”, *Advanced Engineering Informatics*, Vol. 25, pp. 276-287.

Tham, D., Fox, M.S., and Gruninger, M., (1994), "A Cost Ontology for Enterprise Modelling", *Proceedings of the Third Workshop on Enabling Technologies - Infrastructures for Collaborative Enterprises*, West Virginia University.

Uceda-Sosa, R., Srivastava, B., and Schloss, B., (2012), “Building a Highly Consumable Semantic Model for Smarter Cities”, *In Proceedings of the workshop on AI for an Intelligent Planet,* ACM.

Appendix

The Global City Indicator Foundation ontology can be found in:
http://ontology.eil.utoronto.ca/GCI/Foundation/GCI-Founation.owl.

The Global City Indicator Education ontology can be found in:
http://ontology.eil.utoronto.ca/GCI/Education/GCI-Education.owl.

URIs for all of the ISO37120 indicators can be found in:
http://ontology.eil.utoronto.ca/ISO37120.owl.

Definitions of the ISO37120 education indicators, using the GCI Foundation and Education ontologies can be found in:
http://ontology.eil.utoronto.ca/GCI/ISO37120/Education.owl.