Appeared in: Proceedings of the Workshop on Semantic Cities: Beyond Open Data to Models, Standards and Reasoning, AAAI, Quebec City, 28 July 2014.

Foundation Ontologies Requirements for Global City Indicators

Mark S. Fox
Department of Mechanical & Industrial Engineering,
Global Cities Institute, University of Toronto
msf@eil.utoronto.ca

Abstract

City Indicators are metrics used to measure city performance. Global City Indicators, as developed by the Global Cities Institute at the University of Toronto, are metrics that have been agreed to by over 250 cities world wide and have been approved as ISO 37120. The definitions of the indicators exist only in written form. The purpose of this research is to provide an ontology for representing the definition of these indicators and their instantiation by cities worldwide so that they can shared across the Semantic Web. This paper describes the requirements for the ontology and provides an example of its use.

Introduction

Cities are moving towards policy-making based on data. But as Hoornweg et al. (2007) state: "Today there are thousands of different sets of city (or urban) indicators and hundreds of agencies compiling and reviewing them. Most cities already have some degree of performance measurement in place. However, these indicators are usually not standardized, consistent or comparable (over time or across cities), nor do they have sufficient endorsement to be used as ongoing benchmarks."

In response to this challenge, the Global City Indicator (GCI) Facility[1] was created by the World Bank to define a set of city indicators that can be consistently applied globally. This requires agreement on what performance indicators are of interest, and providing a clear, concise and unambiguous definition of each indicator. Over 250 cities are currently participating in both defining and implementing these indicators. Over 100 indicators spanning City Services and Quality of Life have been defined and approved as ISO 37120 (McCarney, 2013).

[1] http://www.cityindicators.org/

Our goal is to formalize the definition of city indicators using the technology of Ontologies as implemented in the Semantic Web. By doing so we will:

- enable the creation of more precise definitions thereby reducing the ambiguity of interpretation,
- take indicators out of the realm of humans and into the realm of computers where the world of Big Data, open source software, mobile apps, etc., can be applied to analyze and interpret the data, and
- achieve interoperability, namely the ability to access, understand, merge and use indicator data available from datasets spread across the web.

What makes this task interesting is the need to select, extend and integrate many existing ontologies to represent a single indicator. These ontologies form a hierarchy from very generic foundational ontologies to more applied ontologies specific to the indicator category.

In this paper we focus on identifying the foundation ontologies required to represent indicators, and use the Student-Teacher Ratio indicator (STR) as an example.

General Requirements and Competency

Consider the definition of the education indicator Student-Teacher ratio (ISO 37120):

> "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 preschool 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."

Our ontology has to satisfy two requirements:

1. We have to be able to represent the definition of the indicator as provided in ISO 37120, and
2. We have to be able to represent a specific city's indicator and the data that supports it.

From a competency perspective, we have three categories of questions we would like to answer:

1. **Consistency**: Is a city's reported indicator consistent with the ISO 37120 definition? If not, where does it deviate?
2. **Longitudinal Analysis**: What is the root cause of change in the value of a city's indicator over time?
3. **Transversal Analysis:** What is the root cause for the difference in the same indicator across two different cities?

In order to satisfy our requirements and be able to answer the competency questions, we need to identify, extend and integrate the ontologies that will form the foundation for representing the indicators.

Foundation Ontologies Requirements

In this section we analyse each component of the STR and based on this analysis identify requirements and a foundation ontology that is needed to represent it.

Placename Ontology

The STR is computed over a geographic area. In the case of GCIs, it would be a city. Hence, a requirement of the GCI ontology is the ability to identify the geographic area over which the indicator has been calculated. That is, to associate a "placename" with a geographic area. For example, a reference to Toronto should cover the city of Toronto but a reference to the Greater Toronto Area should cover the larger area encompassing neighbouring cities. But it must be clear which each refers to. A second requirement is that when two indicators are supposed to be computed over the same geographic area, they are in fact the same area. This means that an area has to have a unique identifier.

There are a number of ontologies that represent geographic and place information. Schema.org[2] provides classes of placenames such as sc:City, sc:Country, and sc:State. It also provides classes for sc:GeoCoordinates (i.e., elevation, latitude, and longitude) and sc:GeoShape denoted by a polygon or circle. The Linkedgeodata.org ontology[3] extends what can have a placename by providing classes for gd:neighborhood, gd:building, gd:bridge, gd:hospital, gd:airport, gd:prison, etc.

The GeoNames project provides over ten million placenames spanning the world. It provides an International Resource Identifier (IRI) for every placename so that they can be uniquely referred to. The GeoNames' placenames are instantiations of the Geonames Ontology[4] that integrates a number of ontologies, including Schema.org and Linkedgeodata.org, to provide a broad set of classes that span almost every conceivable type of place.

At the core of the Geonames ontology is the geo:Feature. A geo:Feature includes the following properties:

- name: text name of the feature, e.g., "Toronto".
- featureClass – Class of feature such as Administrative (e.g., state, parish), Hydrographic (e.g., stream, lake), and Area (e.g., Parks).
- population – Population of the feature.
- wgs84_pos:lat – Latitude of the feature.
- wgs84_pos:long – Longitude of the feature.

The unique IRI for the city of Toronto is: http://www.geonames.org/6167865.

Measurement Ontology

A city indicator is a measure of some property of a city. At the core of an indicator lies a number. The question is what does that number represent? Measurement ontologies provide the basic concepts that underlie numbers. They divide measurement into a Quantity such as length (*the what*) and a Unit of Measure such as meters (*the how*). A Unit of Measure has a scale classified as nominal, ordinal, interval or ratio, and whether the number is the composition of dimensions such as velocity being composed of speed and direction, and whether it has a starting point such as absolute zero on the Kelvin scale.

In the case of the STR, the purpose of a measurement ontology is to provide the underlying semantics of the number, such as what is being measured and the unit of measurement. The importance of grounding an indicator in measurement ontology is to assure that numbers are comparable, not that they are measuring the same thing (which is dealt with later), but the actual measures are of the same type, e.g., ratio of student and teacher population counts, or that the counts of the student and teacher populations are of the same magnitude (i.e., thousands vs millions).

Upper level ontologies such as SUMO (Niles and Pease, 2001) and CYC (Matuszek et al., 2006) provide classes for

[2] The Schema.org ontology is available at: http://schema.org/. We will use the prefix "sc:" to identify classes and properties from the ontology.

[3] The Linkedgeodata.org Ontology is available at: http://www.linkedgeodata.org/ontology/. We will use the prefix "gd:" to identify classes and properties from the ontology.

[4] The Geonames Ontology is available at: http://www.geonames.org/ontology/ontology_v3.1.rdf. We will use the prefix "geo:" to identify classes and properties from the ontology.

representing quantities, but the OM ontology[5] (Rijgersberg et al., 2011) provides a more rigorous ontology based on measurement theory. In the following, we review some of the basics:

- **Quantity**: Refers to what is being measured. It links the phenomenon (e.g., an object) being measured to the value of the measurement (Measure). E.g., Length, Diameter.
- **Unit of Measure**: "A unit of measure is a definite magnitude of a quantity, defined and adopted by convention and/or by law. It is used as a standard for measurement of the same quantity, where any other value of the quantity can be expressed as a simple multiple of the unit of measure. For example, length is a quantity; the meter is a unit of length that represents a definite predetermined length."
- **Measure**: "Combines a number to a unit of measure on an interval or ratio scale." For example, 3 meters, 10 kilograms.

Before we can represent the concept of a STR, there are several building blocks that need to be put in place. First, we need to represent the cardinality of a set. The STR is the ratio of Student to Teacher, which is the ratio of the number of students to the number of teachers. Both students and teachers represent sets, i.e., the set of all students within a placename and the set of teachers within the same placename. The size of each set is its cardinality.

Figure 1 depicts the GCI classes required to represent the number of students and teachers. We start by defining a unit of measure: gci:Cardinality_unit. As the meter is the unit of measure for length, a gci:Cardinality_unit is the unit of measure for the size of a set. The gci:Cardinality_unit has a ratio scale: gci:Cardinality_scale, which is a subclass of om:Ratio_scale and is has a zero element (namely zero). We specialize the gci:Cardinality_unit to the class gci:Population_cardinality_unit which is the unit of measure for the cardinality of set defined by a Population (defined in the next section), and associate the symbol "pc" with it. For example, 1100pc represents a population cardinality (or size) of 1100. We can take full advantage of prefix notations available in OM to scale the numbers by defining units of measures: gci:kilopc, gci:megapc and gci:gigapc which are multiples of gci:Population_cardinality_unit. 1.1 kilopc represents 1100 pc.

With the above defined, we can now introduce the unit of measure for measuring a population ratio such as STR. gci:Population_ratio_unit is defined to be a subclass of om:Unit_division. It has two properties:

- om:numerator whose range is restricted to being a gci:Population_cardinality_unit.
- om:denominator whose range is restricted to being a gci:Population_cardinality_unit.

In other words, a population ratio is derived from two population cardinalities.

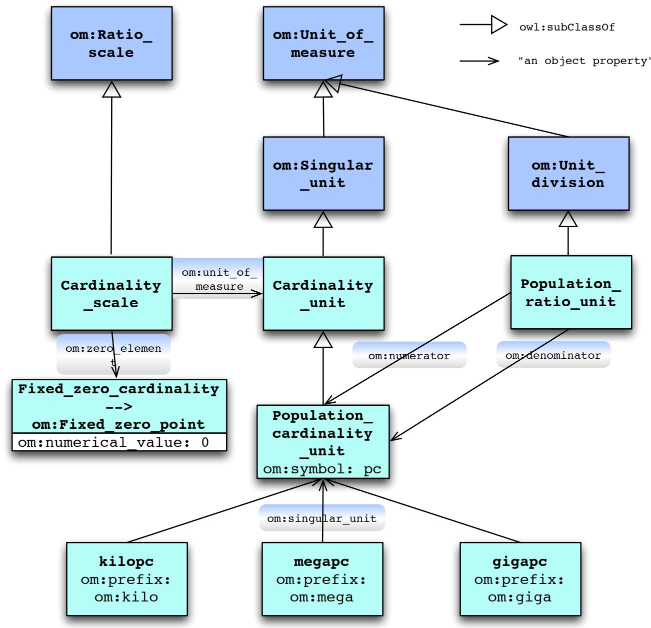

Figure 1

The above, provides the unit of measures for populations (pc) and population ratios (pc/pc) (the *how*). We now have to define *what* we are measuring which is referred to as a Quantity in the OM ontology. First, we need to define the om:Quantity for the size of the teacher and student populations from which the STR is derived. We introduce gci:Population_size as a subclass of om:Quantity. Its om:unit_of_measure is the gci:Population_cardinality_unit. We now have the requisite infrastructure to define GCIs (Figure 2). First we define the class of GCIs, gci:Global_city_indicator, as a subclass of om:Quantity. All GCIs will be a subclass of gci:Global_city_indicator. gci:Education_GCI is introduced as a subclass of gci:Global_city_indicator with a property that it is a gci:for_city_service gci:Education_city_service. Simply, this denotes that this indicator is for the education city service.

The actual value for a city's STR will be an instance of the quantity gci:Student_teacher_ratio_GCI class, which is a subclass of gci:Education_GCI. It has the following properties:

- **om:unit_of_measure,** whose range is the gci:Population_ratio_unit. This signifies that the quantity is a ratio with a numerator and denominator that are restricted to being gci:Population_cardinality_unit's.

[5] The OM ontology can be found at: http://www.wurvoc.org/vocabularies/om-1.8/. We will use the prefix "om:" to identify classes and properties from the ontology. Definitions and examples are taken directly from the ontology where quoted.

- **gci:numerator & gci:denominator,** whose ranges are gci:Student_population_size and gci:Teacher_population_size classes respectively, which satisfy the gci:Population_ratio_unit numerator and denominator constraints.
- **gci:city,** whose range is a geo:Feature that uniquely identifies the city for which this is an indicator.
- **gci:teacher_def & gci:student_def,** whose range are a subclass of Teacher and Student respectively. These define the properties of the teachers and students that we are measuring. For example, all full time students in grades 1 through 12.

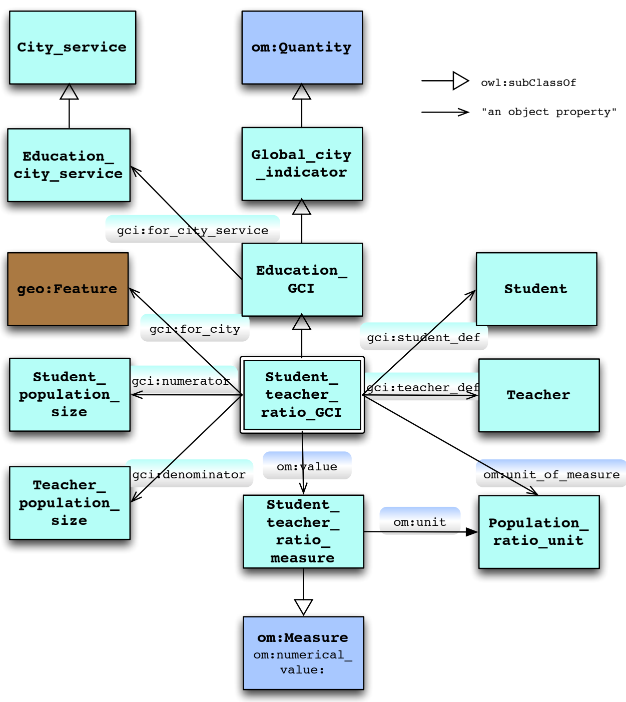

Figure 2

The Quantity instance would link the object being measured (i.e., City) with the actual measurement being an instance of a Measure. The instance of Measure then contains the measurement's numeric value and a link the Unit of measure.

At this point you may have noticed that neither the gci:Student_population_size nor gci:Teacher_population_size have been linked to the students nor teachers within a city. We do so in the next section where we introduce the statistics ontology.

Statistics Ontology

The STR indicator is based on a measure of the number of students and teachers within a population designated by a Placename, namely a city. One can view both sizes as a statistical measurement in the sense that there is a population that we want to perform a measurement of, namely a school population, and we are counting the number of members that satisfy a description of a Student and a Teacher, respectively. While the STR requires a count of the population, other measures would require statistical measures of mean, deviation, etc. of other characteristics of the population.

Anticipating the larger requirements of the Global City Ontology, we have adopted the GovStat[6] general statistics ontology (Pattuelli, 2009). The core class is the gs:Population to be measured. (A definition of the population is not provided and is part of our extension to GovStat.) A gs:Population is linked to a parameter (e.g., mean, standard deviation) by the gs:is_described_by property, and the parameter is a sub class of gs:Parameter. In statistics it is almost always the case that only a portion of the population is measured. This portion is represented by the class gs:Sample, and the parameter being measured is represented as a subclass of gs:Statistic. Finally, the variable for which the parameter is being measured is defined by the class gs:Observation which gs:Statistic links to via the property gs:is_composed_of, and the actual variable which is a subclass of gs:Variable is linked to gs:Observation via the property gs:is_a_characteristic_of.

What we are missing at this point is a definition of the population that we are measuring or from which a sample is to be taken. For the STR indicator the gs:Population must identify the area in which the population resides, i.e., the city, and what characterizes a member of the population, namely the characteristics of a Student or Teacher. For example, the characteristics of a Teacher could be: fulltime, defined as teaching 30 or more hours per week, and teaches at the primary or secondary level, where primary spans grades 1 thru 8 and secondary spans 9 thru 12.

As depicted in Figure 3, we have extended the GovStat ontology as follows:

- Added a property to gs:Population, gs:located_in, that identifies the area that the Population is drawn from.
- Added a property to gs:Population, gs:defined_by, that identifies the class that all members of the Population are subsumed by.

In order to complete definition of gci:Population_size, we need a further constraint. The property of (gs:is_property_of) the gci:Population must be a gs:Count parameter.

Provenance Ontology

Up to this point we have focused on the representation of indicator itself. But another important aspect of an indicator is its providence, namely:

[6] The GovStat Ontology is not available online, but a version with the GCI extensions can be found at: http://ontology.eil.utoronto.ca/govstat.owl. We will use the prefix "gs:" to identify classes and properties from the ontology.

- Who created it,
- What activities were performed to generate it,
- What datasets were used in its generation, and
- When was it generated?

Much of the research into provenance has grown out of workflow management where the focus has been the evolution of a document as it proceeds through a sequence of edits, perhaps by different people and/or systems. Tracking the various versions created, who did what and when has been the primary concern. This research has culminated in the proposed Semantic Web standard called the PROV Ontology[7] (Belhajjame et al., 2012), which has based on the work of Hartig & Zhao (2010) and Moreau et al. (2010). In the following we outline the basic concepts of the PROV ontology and indicate how it is incorporated into the GCI ontology.

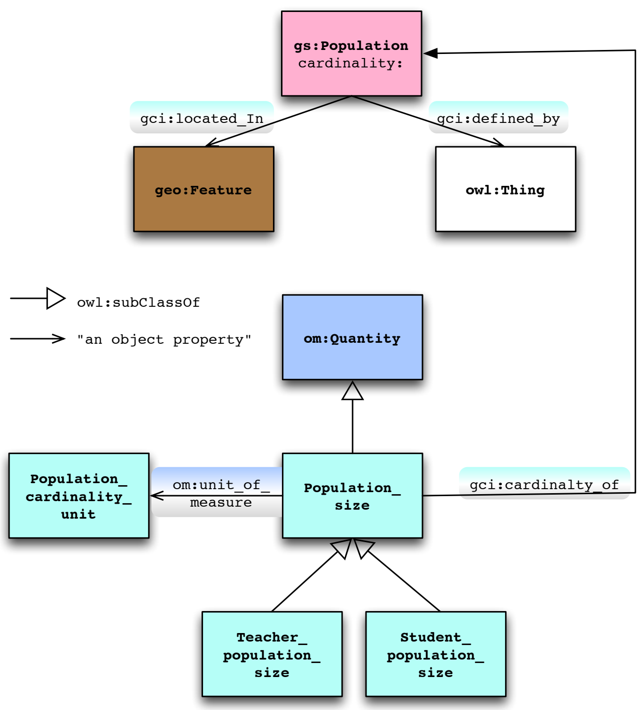

Figure 3

At the heart of the PROV ontology are three classes:

- **pr:Entity**: represents any artifact for which we want to specify its provenance. In our case it would be an indicator or the data from which the indicator was directly or indirectly derived.
- **pr:Activity**: the action (or sequence of actions) that creates or transforms an entity. In our case it may be a computation performed over some data set such as census data.
- **pr:Agent**: the person, organization, or system that performs or plays some role in the activity that transforms an entity. In our case it may be a software application that mines a data set or a person who reviews a data set.

Along with these classes are defined a set of properties that define the causal relationship among entities and activities:

- **pr:wasGeneratedBy**: It links an pr:Entity (domain) to a pr:Activity (range), identifying the activity that generated the entity.
- **pr:used**: It links an pr:Activity (domain) to an pr:Entity (range), identifying the entities used by an activity.
- **pr:wasAssociatedWith**: It links an pr:Activity (domain) to a pr:Agent (range), identifying the agents that play a role in the activity.
- **pr:wasAttributedTo**: It links an pr:Entity (domain) to an pr:Agent (range), identifying the agents that had a role in creating the entity.
- **pr:wasRevisionOf**: Links two pr:Entity's where domain entity is a revision of the range entity.

Finally, the PROV ontology provides a time property that specifies the time an entity was created.

- **pr:generatedAtTime**: It links a pr:Entity (domain) to a pr:time (range), identifying the time the entity was generated.

We integrate the PROV ontology into the GCI ontology as follows (Figure 4). First, we make gci:Global_City_Indicator a owl:subClassOf pr:Entity. Consequently, every indicator we create will be treated as a pr:Entity and inherit its properties, including pr:generatedAtTime which provides us with the time that the indicator was created, and pr:wasRevisionOf which allows us to track revisions to the value of the indicator. It also allows us to link the GCIs to a pr:Activity via a pr:wasGeneratedBy to show what activity generated the GCI, and to a pr:Agent via a pr:wasAttributedTo to show who the source of the GCI was. Finally, the gci:numerator and gci:denominator are made to be owl:subPropertyOf pr:wasDerivedFrom to show what entities were used to derive the GCI.

Time Ontology

Fundamental to the concept of provenance is the time at which measurements are taken, computed or derived. Questions may arise regarding the temporal relationship among indicators and among measurements. Not just at what time something occurred, but whether something occurred before, after or during some external event. For example, was "Total Employment" of New Orleans determined before or after Hurricane Katrina? Or did Katrina

---

[7] The PROV Ontology can be found at: http://www.w3.org/ns/prov. We will use the prefix "pr:" to identify classes and properties from the ontology.

take place during the interval that the indicator was determined? To answer these questions, we need a much richer notion of time that supports reasoning about time points, time intervals and the relationships amongst them. Many time ontologies have been developed. We have chosen OWL-Time[8] for its simplicity and ability to represent time as a point or interval (Hobbs & Pan, 2006). See Figure 4.

Validity Ontology

An ongoing issue with the web is whether information/data found on a page is correct (true) or incorrect (false). Whether the creator of the information deliberately makes false statements, or unknowingly copies false information from another site, there is no way to discern what is correct from incorrect. The same holds with city indicators. Data and analyses that are believed to be true at the time they are gathered or computed, may be found over time to be incorrect. Or it may not be clear whether the information is true or not, especially if the indicator is based on a sampling of a population, but one can assign a degree of validity to the information. In addition, in the case where data is derived from other data, and the latter is no longer valid at some point of time, then the former becomes invalid for that same point of time. For example, gci:Student_teacher_ratio is derived from gci:Student_population_size and gci:Teacher_population_size, if gci:Student_population_size is valid only within an interval of time such as the year for which it is gathered, then outside of that interval, both gci:Student_population_size's and its dependent gci:Student_teacher_ratio's validity are unknown.

Fox & Huang (2005) provide an ontology, called the Knowledge Provenance Ontology[9] (KP), for representing the validity (certainty) of a proposition. It assigns to a "proposition" a validity between [0,1] or "unknown." Validity may be dynamic in that it changes over time. An example of the latter is any population count is valid only at a point of time or for an interval of time. The time interval during which the proposition's validity is known is called the "effective" time interval.

At the core of KP is the kp:KP_prop class which identifies a proposition to which a validity, effective time interval and dependencies can be assigned. We add to the definition of gci:Global_city_indicator that it is a owl:SubClassOf kp:KP_prop. Hence any gci:Global_city_indicator is also a proposition to which we can assign a validity, effective time interval and dependencies.

The following properties are associated with a kp:KP_prop and are inherited by all gci:Global_city_indicator's (Figure 4):

- **kp:assigned_certainty_degree**: This is a data property that maps a kp:KP_prop (domain) onto a number [0,1] (range) or unknown. It is the degree of certainty that the proposition is valid (true) from the perspective of the creator of the gci:Global_city_indicator instance.
- **kp:effective**: This is a data property that maps a kp:KP_prop (domain) onto a time interval (range). It is the time during which the kp:assigned_certainy_degree is valid for the gci:Global_city_indicator instance.

Dynamic Placenames

Consider the unique placename for the City of Toronto. If we wish to do a longitudinal analysis of an indicator for Toronto, we run into a problem. The geographic definition of Toronto changed in 1998 after its amalgamation with five adjacent municipalities.

Yet in the Geonames ontology there is a single Toronto; there is no representation for how placenames evolve over time. Kauppinen and Hyvönen (2007) have addressed this problem. They propose an ontology based on Spatial Temporal Regions. A placename has associated with it a spatial region, defined by a polygon, and a time interval over which the placename and the region do not change.

In the Global City Ontology we will refer to placenames whose spatial regions can change over time as Dynamic Placenames. Rather than adopt Kauppinen and Hyvönen's terminology directly, we adapt their ideas by reusing the provenance, time and validity ontologies to represent how place names change over time and the cause of their change.

Trust Ontology

The final piece of the GCI ontology "puzzle" is the representation of trust. The problem we wish to address is how to represent the degree of trust we have in the creator of indicator values and the data from which they are derived.
Huang & Fox (2006) and Huang (2008) provide an ontology of trust[10]. The ontology views trust as occurring between two agents, where $agent_1$ has or has not trust in $agent_2$. Trust arises out of direct experience or the experience of others whom you may trust. Trust is also context dependent. For example, $agent_1$ may trust $agent_2$ in providing information on topics relevant to their expertise, such as a meteorologist characterizing the climate of a city, but lacks trust in $agent_2$ outside of their field of expertise. Fi-

[8] The OWL-Time Ontology can be found at: http://www.w3.org/2006/time. We will use the prefix "ot:" to identify classes and properties from the ontology.
[9] The Knowledge Provenance Ontology can be found at: http://ontology.eil.utoronto.ca/kp.owl. We will use the prefix "kp:" to identify classes and properties from the ontology.

[10] The Trust Ontology can be found at: http://ontology.eil.utoronto.ca/trust.owl. We will use the prefix "tr:" to identify classes and properties from the ontology.

nally, they identify two types of trust: 1) *trust in belief*, where agent$_1$ believes what agent$_2$ believes, and 2) *trust in performance*, where agent$_1$ believes that agent$_2$ will perform an activity properly.

This representation of trust differs from validity as it refers to our trust in the agent that produced the data. The obvious example is how to represent the trust we have in an organization that has a history of "cooking the numbers." The consequence of not having trust in the producer of data is that the validity one assigns to data or indicator will be reduced by this lack of trust.

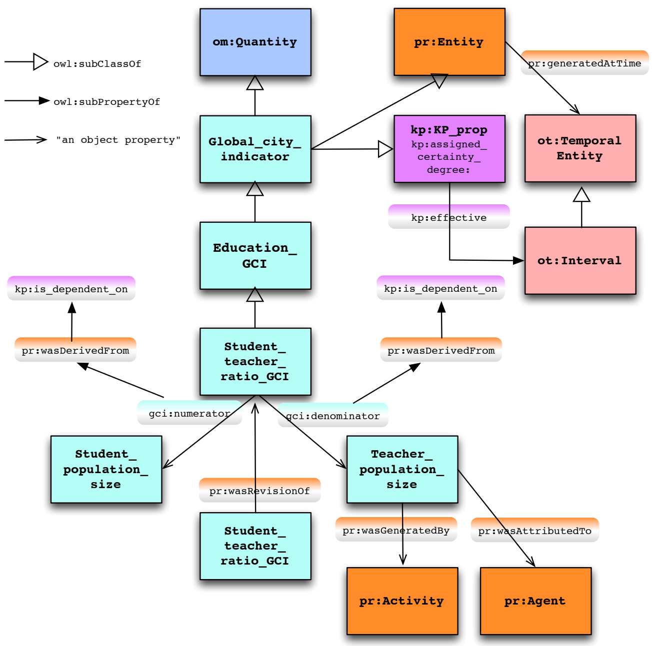

Figure 4

Conclusion

By analyzing the educational indicator Student-Teacher Ratio, we have identified requirements for foundational ontologies spanning placenames, measurement, statistics, provenance, validity and trust. We have also identified existing foundation ontologies that satisfy these requirements directly or by extension. One issue we do not have space to address is to what extent are these foundation ontologies relevant to the rest of the Global Cities Indicators. In Fox (2013) we show that these ontologies underlie the full set of indicators in ISO 37120.

In the section of Requirements and Competency, we identify three categories of competency: consistency, longitudinal analysis and transversal analysis. Work has been done on addressing consistency. Due to lack of space we cannot address it here, but in Fox (2013), a number of consistency questions are identified along with their implementation using the foundation ontologies.

Finally, competency related to longitudinal and transversal analysis is the current focus of our research.

Acknowledgements

This research is supported by the Natural Sciences and Engineering Research Council of Canada.

References

Belhajjame, K., Deus, H., Garijo, D., Klyne, G., Missier, P., Soiland-Reyes, S., and Zednik,S., (2012), "PROV Model Primer", http://www.w3.org/TR/prov-primer.

Fox, M.S., (2013), "An Ontology for Global City Indicators", Working Paper #3, Global City Institute, University of Toronto.

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

Hartig, O., and Zhao, J., (2010), "Publishing and Consuming Provenance Metadata on the Web of Linked Data", *Proceedings of the Third International Provenance and Annotation Workshop.*

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.

Huang, J., (2008). "Knowledge Provenance: An Approach to Modeling and Maintaining The Evolution and Validity of Knowledge", PhD Thesis, Dept. of Mechanical and Industrial Engineering, University of Toronto.

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

Kauppinen, T., and Hyvönen, E., (2007), "Modeling and Reasoning About Changes in Ontology Time Series", In *Ontologies: A Handbook of Principles, Concepts, and Applications to Information Systems*, Integrated Series in Information Systems, Vol. 14, pp. 319-338, Springer.

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., (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.

Moreau et al., (2010), "The Open Provenance Model Core Specification (v1.1)", *Future* Generation Computer Systems. Also see openprovenance.org.

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.