# Completeness of Open City Data for Measuring City Indicators

UNSW AUSTRALIA

Never Stand Still | Built Environment | City Futures Research Centre

Chris Pettit
Professor of Urban Science
c.pettit@unsw.edu.au

CITYFUTURES

# Open Data + Open Cities = Smart Cities

*Smart Cities are enabled through the use of data and technologies to realise  sustainable, productive, resilient and liveable cities.*

# Global City Indicators

- The Economist Intelligence Unit – Global Liveability Ranking
- The Mercer Quality of Life Survey
- 
- Monocle annual quality of life Survey
- OECD Better Life Survey
- ISO37120 Sustainable development of communities
  - *Indicators for city services and quality of life.*

# Attributes of Good City Indicators

- **Objective:** clear, well defined, precise and unambiguous, simple to understand.
- **Relevant**: directly related to the objectives.
- **Measurable and replicable**: easily quantifiable, systematically observable.
- **Auditable:** valid, subject to third-party verification, quality controlled data (legitimacy across users).
- **Statistically representative** at the city level.
- **Comparable/ Standardized longitudinally** (over time) **and transversally** (across cities).

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.

# Attributes of Good City Indicators

- **Flexible**: can accommodate continuous improvements to what is measured and how. Have a formal mechanism for all cities and interested parties to comment on.
- **Potentially Predictive**: extrapolation over time and to other cities that share common environments.
- **Effective**: tool in decision making as well as in the planning for and management of the local system.
- **Economical**: easy to obtain/inexpensive to collect. Use of existing data.
- **Interrelated**: indicators should be constructed in an interconnected fashion (social, environmental and economics).
- **Consistent and sustainable over time**: frequently presented and independent of external capacity and funding support.

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.

# ISO37120 – Sustainable Development of Communities

Indicators for cities and quality of life

First edition 5th for May 2014

"ISO 37120 **establishes definitions** and **methodologies** for a set of city indicators to **steer** and **measure delivery of city services** and **quality of life**." p3.

# ISO 37120 Sustainable Development of Communities

"The indicators can be used to **track** and **monitor a city's progress** on city **service performance and quality of life** and **assist cities in setting targets** and **monitoring achievements**." p3.

"In order to achieve sustainable development, **the whole city system needs to be taken into consideration**. Planning for future needs must take into consideration current use and efficiency of resources in order to better plan for tomorrow." p3.

# ISO 37120 Sustainable Development of Communities

City indicators have been developed to:

1. measure performance management of city services and quality of life over time;

2. learn from one another by allowing comparison across a wide range of performance measures; and,

3. share best practices.

# ISO 37120 Sustainable Development of Communities

**Definitions**

**City:**

“urban community falling under a specific administrative boundary, commonly referred to as a city, municipality or local government.”

**Indicator**

“a quantitative, qualitative or descriptive measure”

# ISO 37120 Sustainable Development of Communities

**Type of Indicators:**

a) **core indicators**: indicators that are required to demonstrate performance in the delivery of city services and quality of life.

b) **supporting indicators**: indicators that are recommended to demonstrate performance in the delivery of city services and quality of life.

c) **profile indicators:** indicators that provide basic statistics and background information to help cities determine which cities are of interest for peer comparisons. Profile indicators are used as an informative reference.

.

# ISO 37120 Sustainable Development of Communities

**17 Themes:**

- Economy (7)
- Education (7)
- Energy (7)
- Environment (8)
- Finance (4)
- Fire and Emergency Response (6)
- Governance (6)
- Health (7)
- Recreation (2)
- Safety (5)
- Shelter (3)
- Solid waste (10)
- Telecommunication and Innovation (3)
- Transport (9)
- Urban Planning (4)
- Waste Water (5)
- Water and sanitation (7)

Total number of indicators: 100 (core and supporting)

# Open Data Challenges

- **Cities and state governments are publishing in a range of forms**
  - Documents (.docx)
  - Spreadsheets (.csv, json)
  - Spatial Data (.kmz, shp, geojson)
  - XML (CityGML, SDMX)
  - Semantic Web Resource Description Framework (RDF)

- **Agencies are publishing “low hanging” fruit**
  - Easiest to release
  - Low utility value

- **‘babel’ of open data**
  - Department in and between cities do not use the same data models
  - Difficult to combine, analyse and compare

# Problem

How can we perform longitudinal and transversal analysis of city data?

- The convergence of global standards for city indicators, open city data and the semantic web provides a unique opportunity for cities and researchers to analyse and compare city performance

- If cities, and agencies with published data on cities, openly publish and link their indicators (using semantic web standards), can we truly understand and believe these indicators without the supporting data also being openly published?

# Developing an Openness Model

**Goal**

Develop a method that can measure the “degree of openness” of a city indicator.

*By “degree of openness” we mean the extent to which the supporting data upon which a particular city indicator is based, is publically available.*

# Introducing CIDOM

**C**ity

**I**ndicator

**D**ata

**O**penness

**M**easure

# CIDOM

- CIDOM-1: is the **percentage of nodes** in the **dependency graph** that are openly published by the city and related government agencies. It provides a general measure of the totallity of the data published for an indicator.

- CIDOM-2: is the **number of levels** of the **dependency graph** that are openly published by the city and related government agencies.

- CIDOM-3: is the **average** of the **format levels** for each node in the dependency graph (i.e., pdf, json, service, etc.) at which the data is openly published by the city and related government agencies. It determines the dominating format  used to publish.

# ISO 37120 6.4
# Primary Education 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.

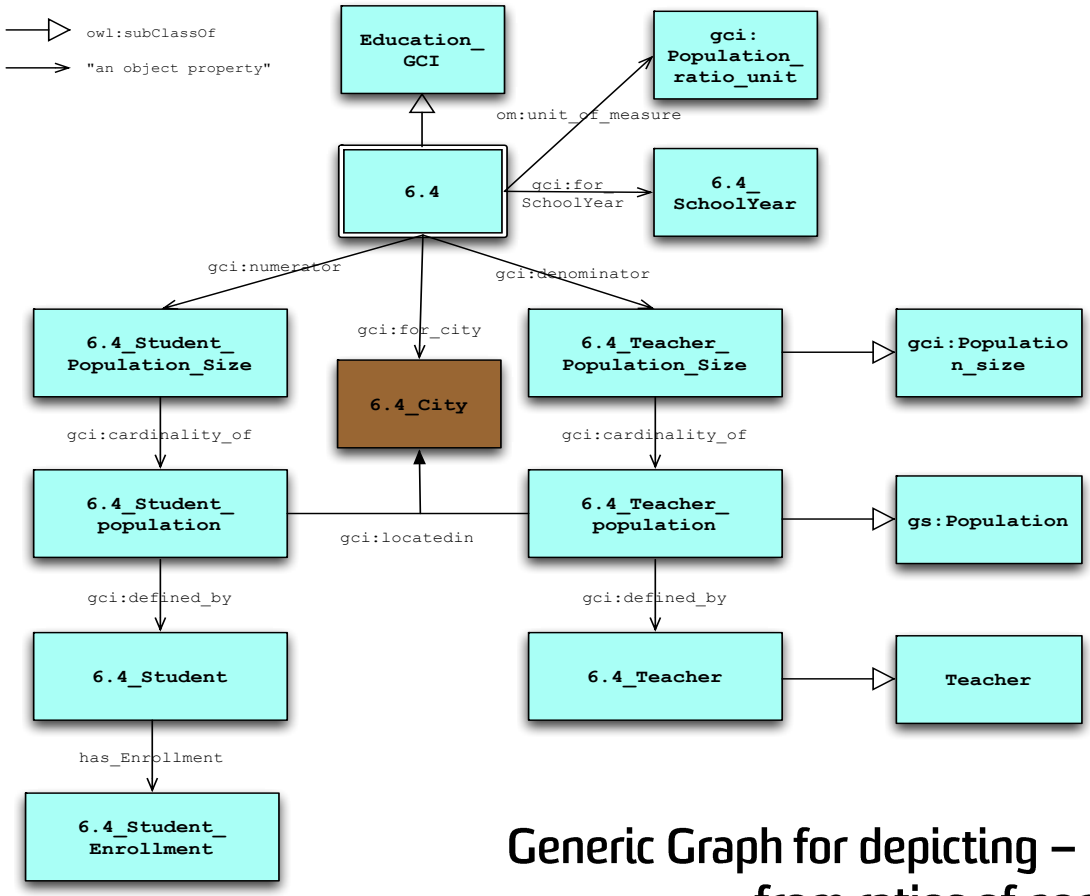
- **One part-time student enrolment shall be counted as one full-time enrolment;** in other words a student who attends school for half a day should be counted as a full-time enrolment. If a city reports full-time equivalent (FTE) enrolment (where two half day students equal one full student enrolment),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

# 6.4 Primary Education Student/Teacher Ratio

## Definition

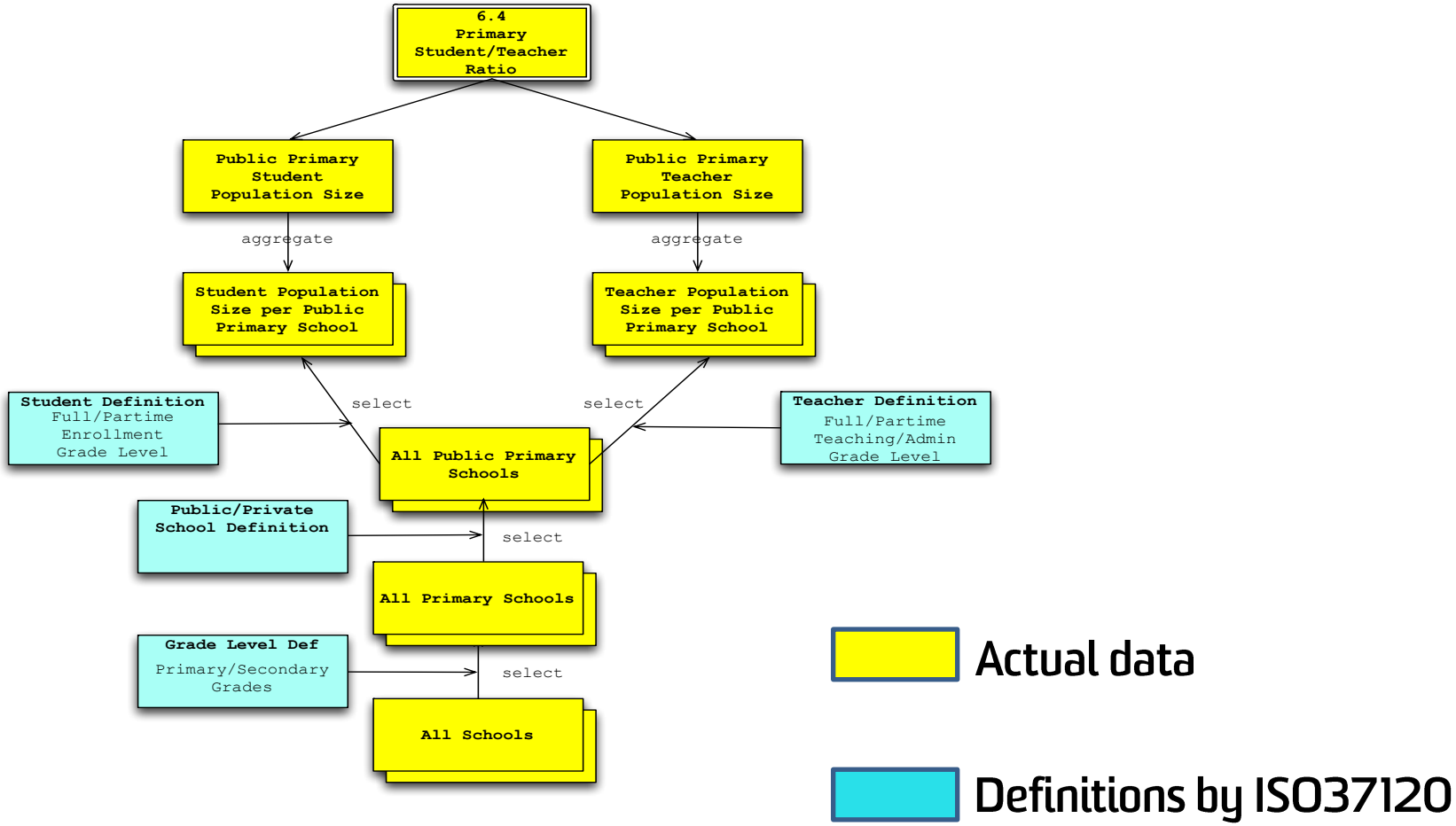



Generic Graph for depicting – indicators derived from ratios of populations

Fox, M.S., (2014), “An Education Ontology for Global City Indicators (ISO 37120)”, Working Paper, Enterprise Integration Laboratory, University of Toronto, 14 October 2014.

# 6.4 Primary Education Student/Teacher Ratio

## Dependency Graph



# The CIDOM Approach

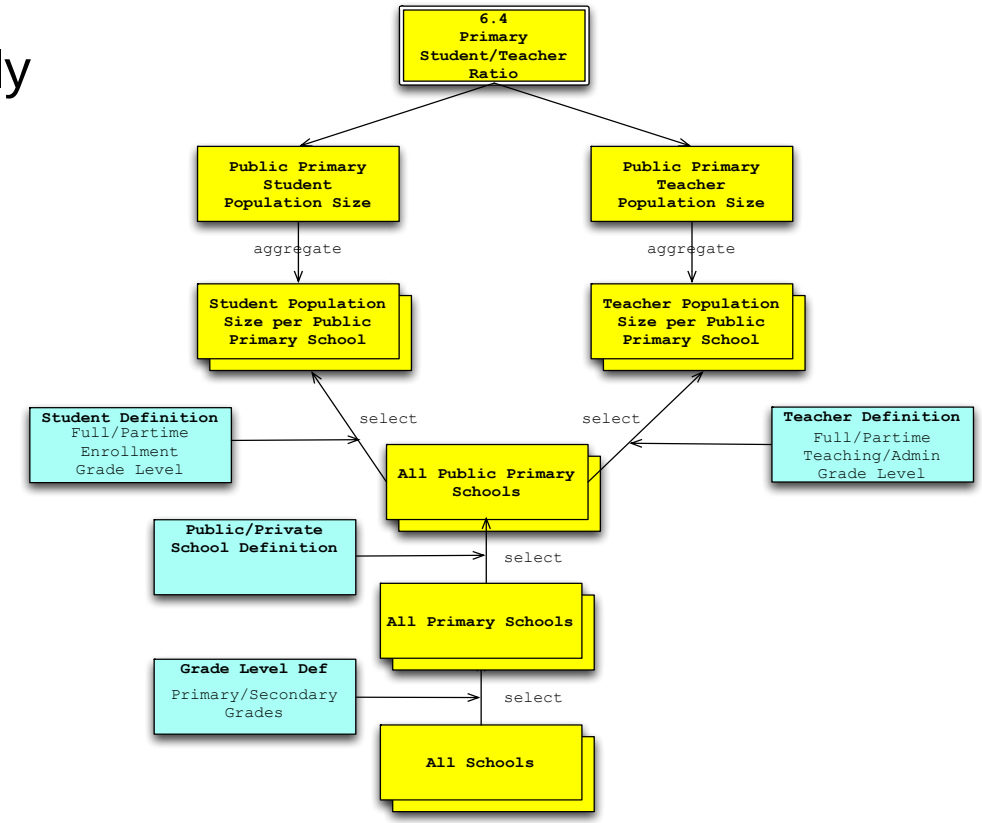
- The indicator **"Student/Teacher Ratio" is the root of a dependency** tree where the supporting definitions and data branch out below it.

- 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.

- **CIDOM-1 is based on the nodes of the dependency tree** that have been openly published. The more nodes published in a more open format, the higher the CIDOM-1 score.

# CIDOM -1

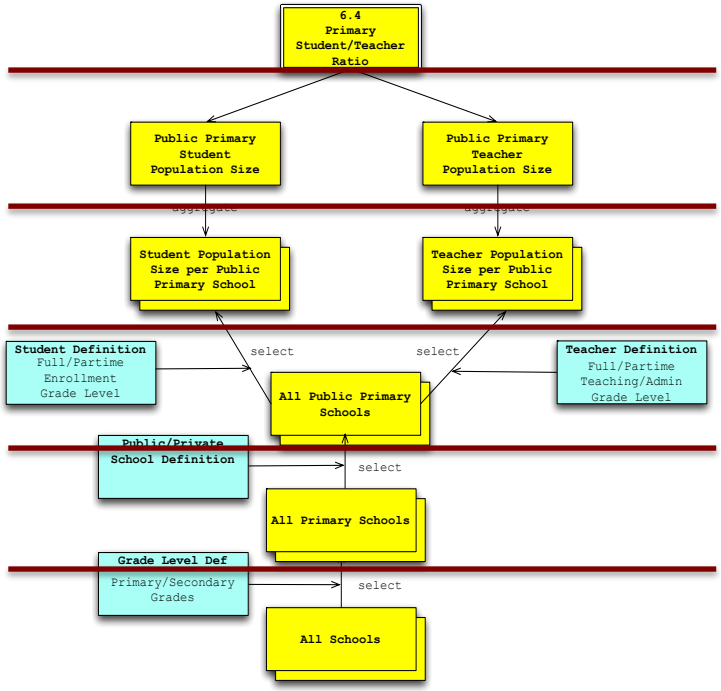
Percentage of nodes in the dependency graph that are openly published by the city and related government agencies.



# CIDOM -2

Number of levels of the dependency graph that are openly published by the city and related government agencies. More precisely:

- **Level 0:** Indicator value is not openly published.
- **Level 1:** Only the indicator value is openly published, e.g., on the student teacher ratio value is published without any supporting data.
- **Level 2:** All of the data that the indicator directly depends on is openly published, e.g., numerator and denominator of the student teacher ratio are published.
- **Level 3:** All of the data directly supporting level 2 is published.
- **Level n:** All of the data directly supporting level n-1 is published.



# CIDOM-3

Average of the levels at which the data is openly published by the city and related government agencies.

1. Published **reports containing data**, on a city website. For example, PDF report.
2. Publish **data openly on a city web site**, can be any format such as csv, json, shp, geo json.
3. Data is made **available as a service**, rather than as download.
4. Data **contains metadata** which conforms to a standard e.g Dublin Core, IS019115…
5. Publish open data that **conforms to linked data standards** for the semantic web, namely RDF triples, but uses internal vocabulary.
6. Publish open data using **commonly accepted vocabulary** for city indicator data.
7. Publish open data using **commonly accepted ontology** for city indicator data.

# Case Study Approach

- Our case study focuses on a single indicator, ISO 37120 6.4, Primary Student Teacher Ratio.
- **Two cities** were chosen: **Melbourne and Toronto**.
- We investigated both the availability of **supporting open data** and the format in which it is published.

# Case Study: Melbourne

- City of Melbourne publish ISO 37120 indicators. However this data is not available via there website or the City's Open Data web site yet.
- Teacher student ratios not maintained by the City, rather The Victorian Government Dept of Education and ACARA have this data.
- Victorian Dept of Education has some data on data.vic.gov but not data that can create 37120 indicator 6.4
- Department contacted but are weary of releasing data.
- ACARA have data on each school on *myschool website*
- Still in negotiations to get the data from ACARA….

# Case Study: Melbourne

- **CIDOM-1:** 0% as none of the 8 nodes in the dependency graph are openly available.
- **CIDOM-2:** is Level 0 as only the indicator value is currently not available.
- **CIDOM-3:** is Level 0 as the indicator value is currently not available in any format.

# Case Study: Toronto

- All ISO 37120 indicators for 2014 is available in PDF on the City of Toronto web site. No other data available on the City's Open Data web site.
- There are two school boards in Toronto
  - TDSB: Toronto District School Board
  - TCDSB: Toronto Catholic District School Board
- TDSB contacted and written request for data provided.
- TDSB suggested contacting City of Toronto.
- City of Toronto suggested contacting TDSB.
- 15 months have elapsed and still waiting.

# Cast Study: Toronto

- **CIDOM-1**: 12% as only one of 8 nodes in the dependency graph is openly available.
- **CIDOM-2:** is Level 1 as only the indicator value is available.
- **CIDOM-3:** is Level 1 as the indicator value is available as a PDF only.

# Conclusions

- The definition and adoption of city indicators on a global basis is a major step forward – ISO 37120.
- But this is only the first step in enabling the transparent, longitudinal and transversal analysis of city performance.
- What is missing is the publishing of the supporting data upon which indicators are based, using an appropriate format and ideally ontology.
- Without this step, meaningful analysis, whether manual or automating is a long way off.

# Conclusions

- **CIDOM** is meant to motivate cities and provide a method to assist in opening more data.

- Through opening up the underlying data for City Indicators this will result in the release of more useful data to support indicator creation and importantly other city wide analysis.

  *Smart Data Release = Smart City*

- **CIDOM** could be applied to other domains… not just city indicators….

# Recommendations

- IS0 37120 be endorsed and adopted by Australian Cities and supporting data agencies (State, Federal).
- A cookbook be created to support Australian Cities in implementing ISO 37120. CIDOM can play and important role to support this.
- ISO 37120 city indicators are published on Data.Gov.au and//or other open data portals.
- Australian support for urban city indicator ontology work and the creation of an urban data dictionary, under the auspices of RDA.

# Next Steps

- Furth develop, test, implement and refine CIDOM
- Apply across more Countries / cities.
- Apply across broader range of indicators
- Develop ontologies for all indicator Themes
- Develop an urban data dictionary

# Collaborators and partners

Professor Mark Fox,
University of Toronto

Fox, M. and Pettit, C. (2015) On the Completeness of Open City Data for Measuring City Indicators, IEEE First International Smart Cities Conference ISC2),Guadalajara, Mexico, 25-28 October 2015.

# Want to get involved?

Research Data Alliance – Quality of Urban Indicators Working Group

https://rd-alliance.org/groups/urban-quality-life-indicators.html

RDA
RESEARCH DATA ALLIANCE

c.pettit@unsw.edu.au