CHALLENGES IN BUILDING THE INFORMATION INFRASTRUCTURE OF SMART CITIES

Mark S. Fox

Distinguished Professor of Urban Systems Engineering
Professor of Industrial Engineering and Computer Science

University of Toronto
msf@eil.utoronto.ca
Outline

• Problems with City Data
• Standards for City Data
• Behavioural view of city information systems
PROBLEMS WITH CITY DATA
A Data Analytics Perspective
IBM Smart City Components
Data Sparseness

- The data we collect is only a small subset of what we need
  - Metrolinx does not use Presto to collect exit data
- What we collect is often stored in “city silos” and not easily accessed by others
- Often what we collect is effectively lost
  - Where did all the street sensor data go that we have been “collecting” since the 1960s?
Data Reliability

• Example: London
  • Sensors reporting trains heading in the opposite direction
  • Multiple trains with the same number
  • Over 1000 bus stops not recorded in the transport database
Premature Data Aggregation

The tendency to aggregate raw data with the belief that you know what people want.

- Spatial
- Temporal
- Categorical
GTA Census Dissemination Areas
Toronto Police Station Boundaries
# Data Interoperability

## Toronto

<table>
<thead>
<tr>
<th>ID</th>
<th>Creation Date</th>
<th>Service Request Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>TO-Request1</td>
<td>01-02-2010</td>
<td>Missing/Damaged Signs</td>
</tr>
<tr>
<td></td>
<td>10:35:00</td>
<td></td>
</tr>
</tbody>
</table>

## New York

<table>
<thead>
<tr>
<th>ID</th>
<th>Creation Date</th>
<th>Agency</th>
<th>Complaint Type</th>
<th>Descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>NYC-Request1</td>
<td>4/30/2013 12:00:00AM</td>
<td>DOT</td>
<td>Street Sign – Damaged</td>
<td>Stop</td>
</tr>
</tbody>
</table>

## 311

<table>
<thead>
<tr>
<th>ID</th>
<th>Division</th>
<th>Section</th>
<th>Service Request Name</th>
<th>Problem Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAM-01</td>
<td>Transportation Services</td>
<td>TMC – Signs &amp; Markings</td>
<td>Missing/Damaged Signs</td>
<td>SAM-01</td>
</tr>
</tbody>
</table>
Data Completeness

ISO 37120: Sustainable Development and Resilience of Communities – Indicators for City Services and Quality of Life

100 indicators defined

**City Indicator Themes**
- Economy
- Education
- Energy
- Environment
- Finance
- Fire and Emergency Response
- Governance
- Health
- Recreation
- Safety
- Shelter
- Solid waste
- Telecommunication and innovation
- Transportation
- Urban Planning
- Wastewater
- Water and sanitation

**Profile Indicator Themes**
- People
- Housing
- Economy
- Government
- Geography and climate
Problem

Toronto

14.6

Melbourne

24.1
Problem

Toronto

14.6

School

Student

School

Teacher

Melbourne

24.1
What Data?

6.4 Primary 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 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."

- Grades, courses, Catholic school, …
Data Complexity

Toronto Well Being
Edmonton: Crime Analysis
233 distinct data sources

Contextual Analysis of Crime

Geospatial
Census
Corporate Databases + Crime Data
Open Data
Google Places API
STANDARDS FOR CITY DATA
Semantic Interoperability

- Many smart city applications require the integration of many data sources, both internal and external to the city.
- Early applications, such as visualization of "hot spots" (e.g., crime), assumed that the attributes contained in the data sources were independent, and only a single shared attribute (e.g., geo-location) was needed to integrate them.
- With more sophisticated applications, the independence assumption is no longer valid.
  - A review of 311 datasets across North American cities identified the presence of semantic ambiguity and overlap in important attributes and values (Nalchigar & Fox, 2017).
- Integration of the datasets requires **Semantic Interoperability**
  - Semantic interoperability is the ability of computer systems to exchange data with unambiguous, shared meaning.
ISO JTC1 WG 11

• What the Acronyms mean
  • ISO: International Standards Organization
  • JTC: Joint Technical committee of ISO and IEC
  • IEC: International Electrotechnical Commission
  • WG 11: Working group that focuses on Smart City Frameworks and Data standards

• Three standards projects of relevance to City Data Semantic Interoperability
  • 30182: Generic concept model for city data
  • 21972: Standard for city indicator data, e.g., ISO 37120
  • 30145-2: Standard for city meta data
• Achieving interoperability across cities through the adoption of common concept (ontology)
• Based on BSI PAS 182, which is based on ontology research over the last 2-3 decades
Event Concept and Relationships

- Observation
- Case
- Account
- Decision
- Item
- Plan

- Event
  - at
  - has
  - sub-concept of
  - contains
  - contains
  - outcome of
  - has role in
  - plans event

Place
State
Upper Level Ontology for Smart City Indicators

A standard for representing indicator definitions (machine readable) and the data used to derive them


**Initial Draft:** to be distributed at WG 11’s June 19th meeting in Southampton, UK.
Indicator Pattern

Indicator unit of measure (measurement ontology)

Year of measurement (temporal ontology)

Place of measurement (placename and geospatial ontologies)

Statistic being measured (Statistics ontology)

Population being measured (Population ontology)

Person being counted (Theme ontology)
30145-2

- Ontology for representing meta-data for Smart City data
  - A standard for representing meta data for individual pieces of data (not data sets)
  - **Initial Draft:** to be distributed at WG 11’s June 19th meeting in Southampton, UK.
Define a standard for representing Meta-data.

- Place names
- Provenance
- Measurement
- Validity
- Trust
Future Standards
Service Specific Concepts

- ISO 37120 Education Theme Indicators

6.1: % of female school-aged population enrolled in schools
6.2: % of Students Completing Primary Education: Survival Rate
6.3: % of Students Completing Secondary Education: Survival Rate
6.4: Primary Education Student/Teacher Ratio
6.5: % of male school-aged population enrolled in schools
6.6: % of school-aged population enrolled in schools
6.7: # of higher education degrees per 100 000 population
Student Pattern

Pattern Competency

• What is the age range for school age women?
• Did the student attend a public school or private school?
• What schools did student attend in year Y?
• Was the student part time or full time?
• What grade level did they attend?
Service Specific Concepts

- Example: 311 (complaint management)
UofT Service Ontologies

- **Completed**
  - Foundation
  - Education
  - Energy
  - Environment
  - Finance
  - Innovation & Telecom
  - Shelter
  - Transportation

- **Under Development**
  - Fire & Emergency
  - Governance
  - Health
  - Recreation
  - Safety
  - Water & Sanitation
BEHAVIOURAL VIEW OF CITY INFORMATION SYSTEMS

A New Perspective on City Information Systems
Complexity

IT’S NOT ROCKET SCIENCE!
“composed of a system of systems (i.e., services) acting as an organic whole with ICTs providing an artificial nervous system.” Kanter & Litow (2009)
Smart City Fallacy

• Viewing a city as a **system** we can control.
  • Complexity and uncertainty works against us!

• Instead it is composed of millions of agents and billions of activities operating in loose coordination according to a set of rules and laws imposed by society, and subject to a set of other “laws” not under our control:

  • **Murphy’s Law**: What can go wrong, will go wrong
  • **Molte’s Law**: No battle plan survives contact with the enemy
  • **Billings’ Law**: “The squeaky wheel gets the grease”
  • **Fox’s Law**: There will always be people who work the “system” to their own benefit, or for every rule, there is someone who is ready to break it
Managing Complexity and Uncertainty

- If we are to manage the complexity and uncertainty inherent in a city, we need a different way of thinking about the digital systems that increasingly operate and manage our cities.

What are the BEHAVIOURS that we wish our digital systems to exhibit?
Aware

More than sensing the environment

• Know what is to be expected and the limits or constraints on what is being observed.
• Know whether deviations are significant.
• Know who/what has to be informed when deviations occur.

1. **Sense** - Digitization of the city
   • IOT, Human Sensor Networks, eGovernment
2. **Remember** – Removing walls between data
3. **Standardize** – Achieving interoperability
4. **Detect** – Operational deviations
   • Sensors, Services, HR, Finances, Governance
Responsive

Ability to do the right thing by flexibly responding to events with a focus on outcomes and not means.

1. **Situation Understanding** – all relevant information available.
2. **Shared Knowledge** – access to prior experiences.
3. **Flexible** – non-rigid responses to varying situations.
4. **Empowering People and Systems**
5. **Teleological** – replacing process with goals.
   - Operate within constraints
Introspective

- Ability to examine the systems performance and find new ways of achieving goals.

1. **Recognize Failure:** Recognizing that goals are not being achieved.
2. **Diagnose root cause of performance.**
3. **Construct globally better solutions.**
4. **Predict events, behaviours.**
5. **Self-Aware:** Security, privacy, governance.
Accountable

Cities must be held accountable for their actions.

1. **Determining Responsibility**
   - Do we have a digital trail of information, activities and agents that contribute to a particular decision or action?
     - Provenance
   - Can we determine what factors contribute to the situation and who was responsible for those factors?

2. **Determining underlying basis for decisions**
   - Looking for recurring patterns of behaviour
   - Looking for connections between interested parties

3. **Digital Ombudsman**
   - Who is the system accountable to?
   - Who decides what it must be accountable for?
   - Who decides how to respond to accountability issues?
Conclusion

• A lot of issues with the data has city has or does not have
• The need for standards to achieve interoperability
• A behavioural perspective on a Smart City’s digital infrastructure
  • Aware
  • Responsive
  • Introspective
  • Accountable