Recommendations from the U.S. for TC 268 Data Exchange Standard for Smart and Sustainable Cities

Introduction

The U.S. provides the following recommendations as focus topic areas for a Data Exchange Standard (or sub-parts for a multi-part standard). These recommendations reflect the fact that:

- A Smart City may not necessarily require digital computing (ICT) technologies, but the application of ICT to city management and operations is becoming increasingly common. This is not only reflective of the fact that such technology is becoming more affordable, but also the growth of opportunities for commodification. Data exchange practices must take into account potential for public-private partnerships or other means of monetization.
- Not only do city entities themselves generate vast amounts of data that are integral to the city, so do city inhabitants, businesses, and organizations. Data exchange practices must account for differences among the parties who generate or own data that may be integral or useful for smart city applications.
- While cities may be located in a defined physical space, communities may be less restricted. Moreover, the Internet, which moves data within and between communities, is not contained within physical boundaries. Data exchange practices must reflect the fact that city data is not limited by the physical bounds of the city.
- While a city can be viewed, and understood as an ecosystem with a physical structure, a society, and a flow of interactions, data and information[1][2] have become essential city infrastructure. The Interactions between the physical structure and society effectively reflect the activities in the city. These activities can be analyzed and measured as flows of information including city functions, city economy, city culture, and even the flow of data and information itself. Data exchange practices must facilitate this analysis.

Recommendation 1: Data Groups/Master Data Management

This standard should create some basic data groupings that would help parties engaged in information sharing understand the meaning of the information to each party. While there has been work done within other ISO committees on data groups related to specific applications (e.g. ISO/IEC 19661:2011, which defines a data group a "distinct, non-empty, non-ordered and non-redundant set of data attributes where each included data attribute describes a complementary aspect of the same object of interest; and ISO/TS 16785:2014, which defines a data group as a "class of closely related attributes"), there is no international standard on the different groupings of information one might find in a large organization, like a city. Table 1 below shows example data categories, how they might be defined, and some policy considerations that could be applicable to each.

[1] Data, Information and Knowledge https://www.youtube.com/watch?v=QsP5WGv0aQc
[2] Data-Information-Knowledge in 3 minutes or less https://www.youtube.com/watch?v=slJSY05JE9Q

Table 1. Data groupings applicable to a large organization, such as a city.

Data Category		Possible policy, sharing, other notes
Master Data	"Single source of truth". Relatively static about people, places and things involved in an organization or community.	Usually kept in the financial systems (SAP Financials, Oracle Financials) or depending on the organization an 'Operations System' (Inventory Mgmt, DRP, ERP, CRM, ..)
Reference Data	Sets of values / classification schemes referred to by systems, people, processes and transactional and master records	
Transactional Data	Describes the internal or external events or business transactions that take place as an organization conducts its business.	Customers want a capability for real time order status. Business transactions need to be discoverable; e.g. legal.
Historical Data	Significant facts/events as of a certain point in time that should not be altered; e.g. security reporting, compliance reporting, invoices.	Historical data needs to be discoverable.
Metadata	1. Audit Trail Metadata: for security, compliance and forensics e.g. time stamps, creator, create data, update date, .. 2. Business Metadata: non-technical aspects of data and their use e.g. field definitions, report names, organizational vocabulary .. 3. Technical Metadata: used to describe technology and data structures. **Clarification NOTE**: there may be more common types of Metadata	
Temporary Data	Traditionally data kept inside of an organization to speed up business processing, typically not by people but by systems.	Data telemetry, local 'cache stores', …

As an example of how these groupings might be used, consider telemetry data collected directly from environmental sensors. In order to attach the proper terms of use in a data sharing context, the owner or custodian of that data would want to consider whether it should be considered "temporary data," "transactional data," or "historical data"

Recommendation 2: Data Sharing Spectrum

A data exchange standard should also contain a smart city data sharing spectrum that can guide users on when and how certain types of data may be shared. A data sharing spectrum helps answer the question "What can we share with OTHERS?" and "What can others share with US?" These questions become increasingly important as city entities generate more data, and as they make use of data generated by third parties, including citizens. Table 2 provides an example of what a data sharing spectrum could look like for a smart city or community.

Table 2. A Smart Community Data Sharing Spectrum allows users to classify data in terms of the proper audience (who can see and use the data) and decide the right enforcement needs for that classification.

Illustrative example of a Smart Community Data Sharing Spectrum					
	CLOSED (within an Entity)		**SHARED (between Entities)**		**OPEN (with any Entity)**
Possible Classification	INTERNAL ONLY	NAMED ACCESS (Individual Access)	GROUP ACCESS	PUBLIC ACCESS	PUBLIC
audience	**Internal Audience Only** (within a legal entity)				**General Public** (anyone with access)
enforcement	"Formal Agreement"; e.g. legal contract / Law / Regulation	"Explicit Assignment"; e.g. Agreement	Via "Authorization"; e.g. Administrator	"Limited License"; e.g. Terms of Use License	Open License; Privacy statement,
example	Company sales report document	Vendor reports	Company emails	LinkedIn account	Company website

Recommendation 3: Guidance to Better Understand and Implement Information Security Classifications

Current data classification schemes may create a false sense of security, as described in a recent article by the consulting firm Gartner:

> From Gartner's article **"How to Overcome Pitfalls in Data Classification Initiatives"**[3], "*Organizations often perceive a false sense of security when they implement a data classification scheme or process. There is a mistaken belief that because a data item has a classification label assigned to it, it is secured.* ***The reality is that a classification is just a label, or an additional metadata tie to a data item.*** *What users or systems do with this information is what matters — if they disregard it, it will not provide any additional protection. This is* ***the fundamental linchpin for data classification programs — the data classification program must be tied to concrete guidelines and actions to be taken for each classification, otherwise it is of little to no use at all.***"

Thus, standalone classifications based on risk, like "High Business Impact," "Medium Business Impact," or "Low Business Impact," without more, will not support secure data exchange. Moreover, as the hardware and software of the Internet of Things (IoT) develops (and in fact, is already deployed in some cities), there will be a growing need for convergence of standards and practices from the domain of virtual information technology (IT) data management with physical operational technology (OT) management principles (see Figure 1, which illustrates the safety, privacy, security, reliability, and resilience needs that converge as cyber-physical systems become common in city infrastructure).

[3] https://www.gartner.com/doc/3247517/overcome-pitfalls-data-classification-initiatives

In the safety arena, standards like in IEC 61508 "Functional Safety of Electrical/Electronic/Programmable Electronic Safety-Related Systems" may provide helpful guidance. In addition, data classification schemes must also address the data security concerns (including data in motion, data at rest, data in use, and data in storage), the classification scheme and related guidelines/policies must account for data reliability (i.e. is the data sufficiently complete), personally identifiable information (PII) protection (i.e. whether the collection and dissemination of data matches public expectations around privacy), and data resiliency (i.e. whether the data is in a state that enables it to remain available for applications).

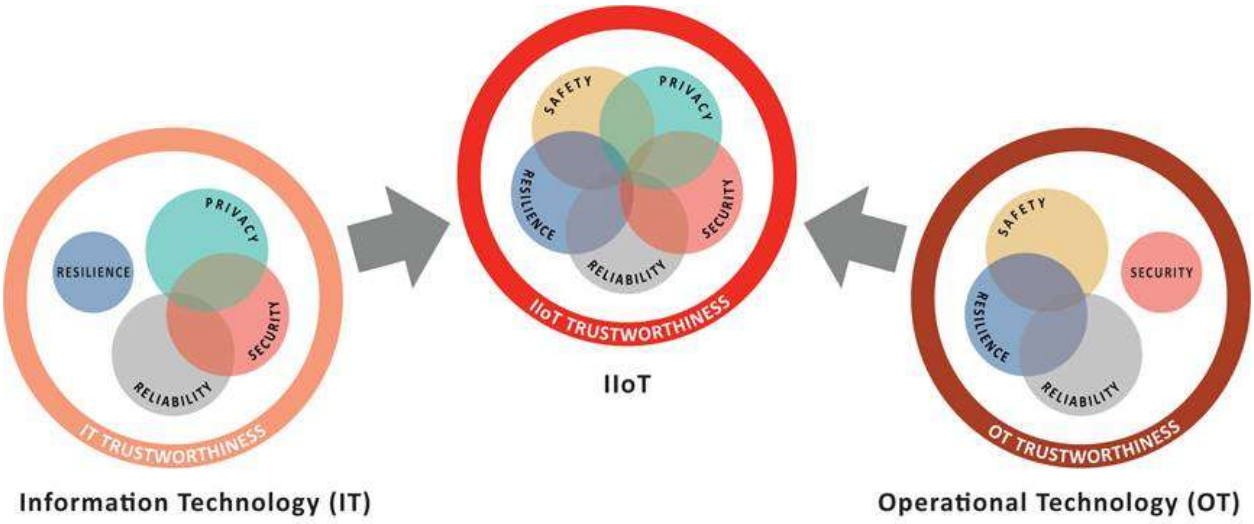

Figure 1. The Industrial Internet of Things (IIoT) represents a convergence of digital IT with Operational Technology (OT). *Image from the Industrial Internet Consortium.*

Recommendation 4: Data Maturity Guidance

Communities, like other organizations, are likely not immune to "redundant data" or other issues related to data management. Some type of guide, like a "data maturity index," would be useful to help communities navigate these issues. Following a Plan, Do, Check, Act scheme, such a guide could help communities understand their data management capacity and ability at present, identify areas in need of improvement, and provide guidance in determine how much the community would need to invest to improve its data management practices.

Recommendation 5: Support for Multiple ICT Data Management and Data Collection Technologies

Regarding underlying ICT data management technologies, in any given city or community a variety of communications and security protocols may be available, depending on various considerations including; available spectrum, communications reliability, data latency, and low power devices in very large numbers which may or not have a known physical location. Given that, it will be necessary for cities to support a variety of technical protocols to realize the general Internet of Things (IoT) both in the areas of Cloud Computing[4] and for Edge Computing[5], for a sustainable data sharing economy in cities.

[4] https://en.wikipedia.org/wiki/Cloud_computing
[5] https://en.wikipedia.org/wiki/Edge_computing

Recommendation 6: Guidance for Applying the Above Recommendations to Specific Smart City Opportunities

The chart below (Table 3) was developed by the ITU-T to illustrate potential IoT opportunities for cities within different domains. A data exchange standard could develop guidance around data management considerations that should apply within the domains. This would necessarily include developing an understanding of the type of data to be collected and the general goal for its use (such as better real-time planning; emergency response needs; long-term planning, etc).

Table 6. IoT Opportunities for Smart Cities.