Open Data Standards Pilot Project

Final Report

June 2018

Acknowledgements

The pilot project has been a success based on the participation and contribution of the following organizations and individuals:

MISA Ontario Pilot Project Team

Morgan Calvert, Project Liaison
Connie McCutcheon, Project Sponsor
Catherine Baldelli, Director-at-Large
Ron St. Onge, Director Northern Region
Anne Babej, Executive Director
Jacqueline O'Hara, Community Manager

Pilot Project Municipal Participants

City of Brampton - Katherine Kulson, Gaea Oake
Greater Sudbury - Ron St-Onge
City of Guelph - Sasha Einwechter
City of Kitchener - Dianne Adams
Town of Milton - Catherine Baldelli
Niagara Region - Connie McCutcheon
City of Ottawa - Darrell Bridge
City of Toronto - Denis Carr
City of Welland - Jamie Leitch
City of Windsor - Robert Price

Appreciation to all from the OpenNorth Pilot Project Team

Jean-Noé Landry, Executive Director
Peck Sangiambut, Research and Policy Officer
Jury Konga, Associate

Overview and Background

Canadian municipal open data is expanding quickly. Since the formation of the G4 group of Canadian cities (Toronto, Ottawa, Vancouver and Edmonton) in 2009, open data has become a priority at the Federal and provincial levels, while municipal open data initiatives have grown to include smaller cities and towns now numbering over 100 in Canada. The value of standards are recognized in many disciplines and is a business imperative for the Information Technology sector. Standardization in how datasets are presented in open data catalogues remains an issue, which affects efforts to create federated search and efficient use by end users (internal and external). Federated search, a technique for searching multiple databases simultaneously to return a single result,[1] is one goal of data standardization and interoperability, and is recognised by the Federal Government as an important objective.[2]

This pilot project[3] is aimed at assessing the current municipal open data situation and potentially setting the stage for a collaborative cross-jurisdictional standardization effort on open data. It does this by diagnosing discoverability in ten datasets, highlighting initial findings of gaps in standardization, and proposing a draft pilot dataset definition. As this initiative potentially scales in the future, the inclusion of data modelling and ontological (e.g. properties and relationships) interoperability will also need to be addressed.

The project looks at two main aspects of data in catalogues: data discoverability, and standards compliance and interoperability.

- For data discoverability, there is an exploration of some of the public-facing aspects of datasets, which include: naming conventions (for file/dataset names), categorisation of datasets, keyword tagging, metadata standards.
- For standards compliance and interoperability, there is an exploration of the varying use of standards as they relate to the content of a given dataset and its metadata, which include: domain standards, metadata standards, and atomic standards (standards that define basic attributes such as a date format). This exploration is necessarily contextualised within larger data ecosystems of regional and national data infrastructures.

The report includes a differentiation between types of standards where necessary, such as vocabulary (e.g., Data Catalog Vocabulary), schema. By looking at the standards coverage of these datasets (i.e. which datasets use which standards) across a number of municipalities, and identification of gaps in standardization for each category.

---

[1] Shokouhi, M., & Si, L. (2011). Federated search. *Foundations and Trends® in Information Retrieval*, *5*(1), 1-102.

[2] Dusseault, Pierre-Luc. (2014) Open Data: the Way of the Future. Report of the Standing Committee on Government Operations and Estimates. http://publications.gc.ca/collections/collection_2014/parl/xc70-1/XC70-1-1-412-5-eng.pdf

[3] https://www.misa-asim.ca/news/news.asp?id=388366

Open Data Standards

Benefits of Open Data Standards

The following provide a summary of some key benefits derived from data standards including open data:

- “Standards facilitate development, sharing, and use of geospatial data and services”, U.S. Federal Geographic Data Committee (FGDC);
- Being machine readable in common data format standards (e.g.CSV, JSON, XML) allows effective use of data;
- Standard data definitions facilitate the interoperability among business systems (databases) and potential use of APIs;
- Standards for describing data (e.g. names, tags and metadata) are most effective discovery mechanism for internal and external users;
- Reduction in time-consuming data transformation efforts; and
- Economic development benefits from entrepreneurs accessing and utilizing standardized municipal data.

Prior standardization such as the Canada wide BizPal project among all levels of government should be viewed as an example of success in inter-jurisdictional standards co-development.

Standards Review and Resources

Standardization can occur at multiple levels, ranging from the data about a dataset (metadata) to the formatting and range of attributes for a given data field.

The following resources were researched to consider in the exploration of datasets and develop a recommended dataset definition.

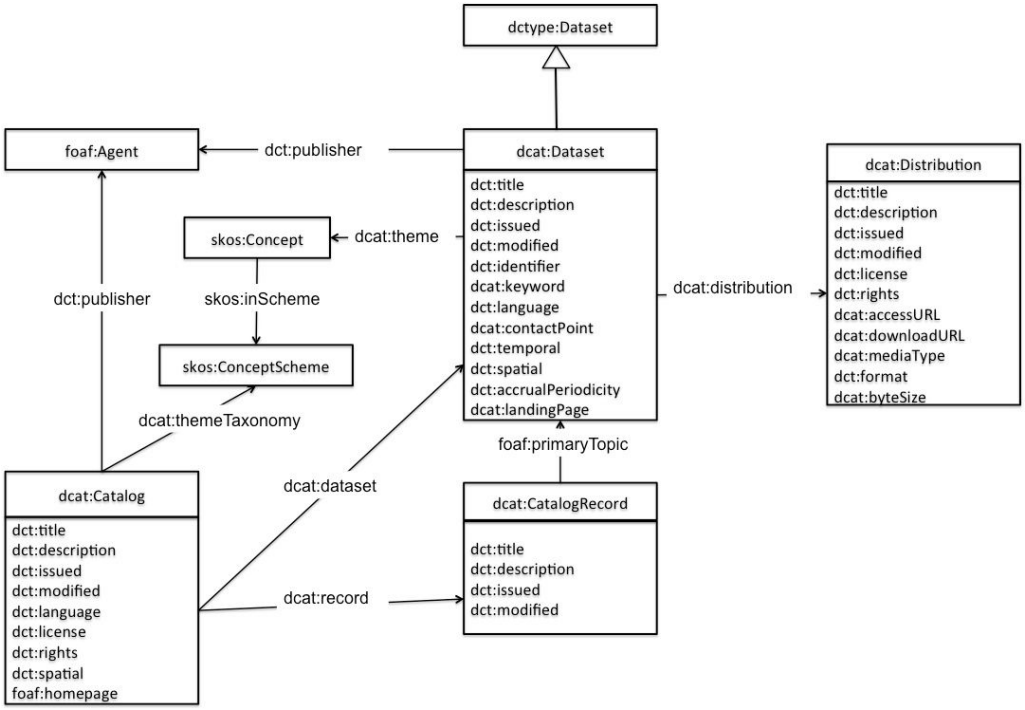
Name	URL
W3C - Data classification	www.w3.org/TR/dwbp/
W3C - Data best practices	www.w3.org/TR/dwbp/
ISO/IEC Odata standard	https://www.oasis-open.org/news/pr/iso-iec-jtc-1-approves-oasis-odata-standard-for-open-data-exchange
Open Science taxonomy	www.fosteropenscience.eu/taxonomy/term/112
Open Data Standards Directory	https://datastandards.directory/
Open Referral - Health Services example	http://openreferral.readthedocs.io/en/latest/hsds/reference/#hsds-spec
European PSI - GIS data	https://www.w3.org/2013/share-psi/bp/sgd/

Open Geospatial Consortium	http://www.opengeospatial.org/docs/is
US National Information Exchange Model (NIEM)	https://www.niem.gov/
ESRI Canadian Municipal Data Model	https://www.arcgis.com/home/item.html?id=e0a1b07b9e594dffba8ce9ad8f79c3e2
Municipal DIY Open Data Toolkit	https://open.canada.ca/en/do-it-yourself-open-data-toolkit
Open Knowledge International	https://okfn.org/projects/
Open Data Institute (UK)	https://theodi.org/publications

Catalogue standards

Part of this project addresses the visible standards compliance for a given dataset when viewed in an open data catalogue (e.g. naming convention, keyword tags). Therefore, the exercise considers working on operationalising data catalogue standards.

Standards exist for data catalogues. DCAT (Data Catalog Vocabulary), and the Dublin Core Metadata Initiative (DCMI) describe how catalogues themselves should be structured and attributes to datasets assigned. DCAT is particularly important because it is used at the international level (such as the UK and EU), as well as locally the including Government of Canada and Government of Quebec. It is also used at municipal level.



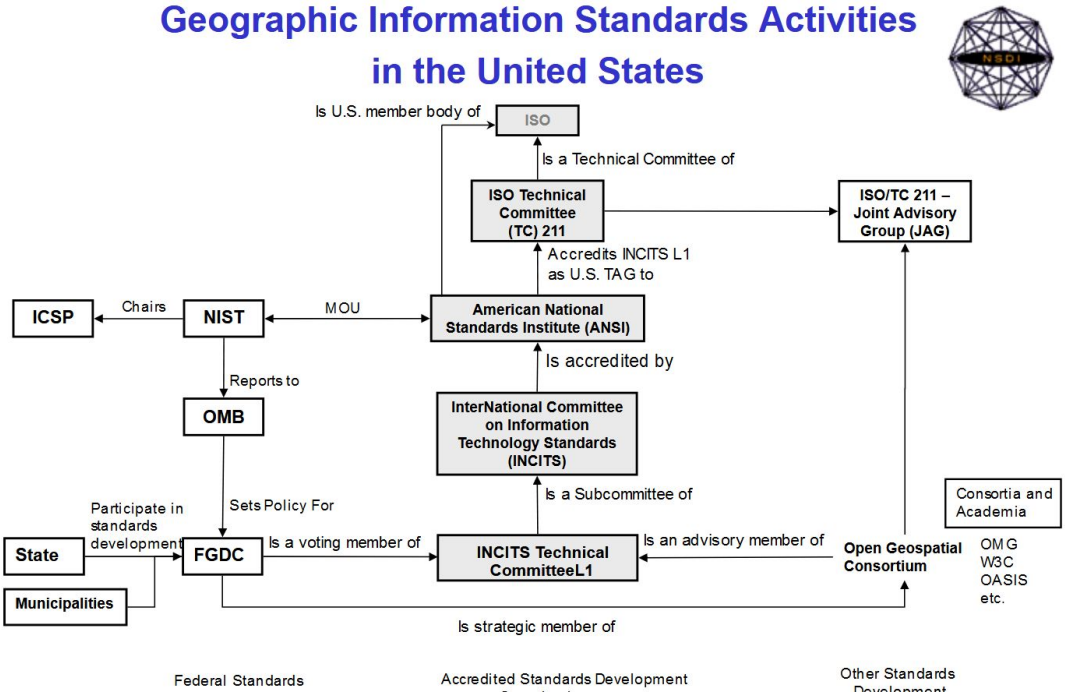
Source: https://www.w3.org/TR/2014/REC-vocab-dcat-20140116/

The figure above describes the relationship between catalogues and datasets. In this pilot project, the focus is on the dataset itself (dcat:Dataset) and its components including title and keyword.

Standards dependencies

Notably, standards are also built on top of other standards. Supplementary standards may be needed when a single standard does not provide enough specificity. For example, DCAT uses Dublin Core metadata terms to increase standards interoperability, which defines terms it does not.

Geospatial data standards are important because of the dominance of spatial data in the data that municipalities collect. Spatial data can be found ranging from coordinates and addresses to polygons and topology. Geospatial or geographic data standardization is a massive endeavour, which has resulted in a myriad of institutions such as the case of the USA (image below). The National Spatial Data Infrastructure (NSDI) of the United States has been a long time leader in standards development as has the European Union's INSPIRE (Infrastructure for Spatial Information in Europe). The schematic below illustrates the many organizations involved in the standards setting process. Having multiple committees (some of which may be specific to domains or jurisdictions) and institutions complicates the formal standards development and adoption process.

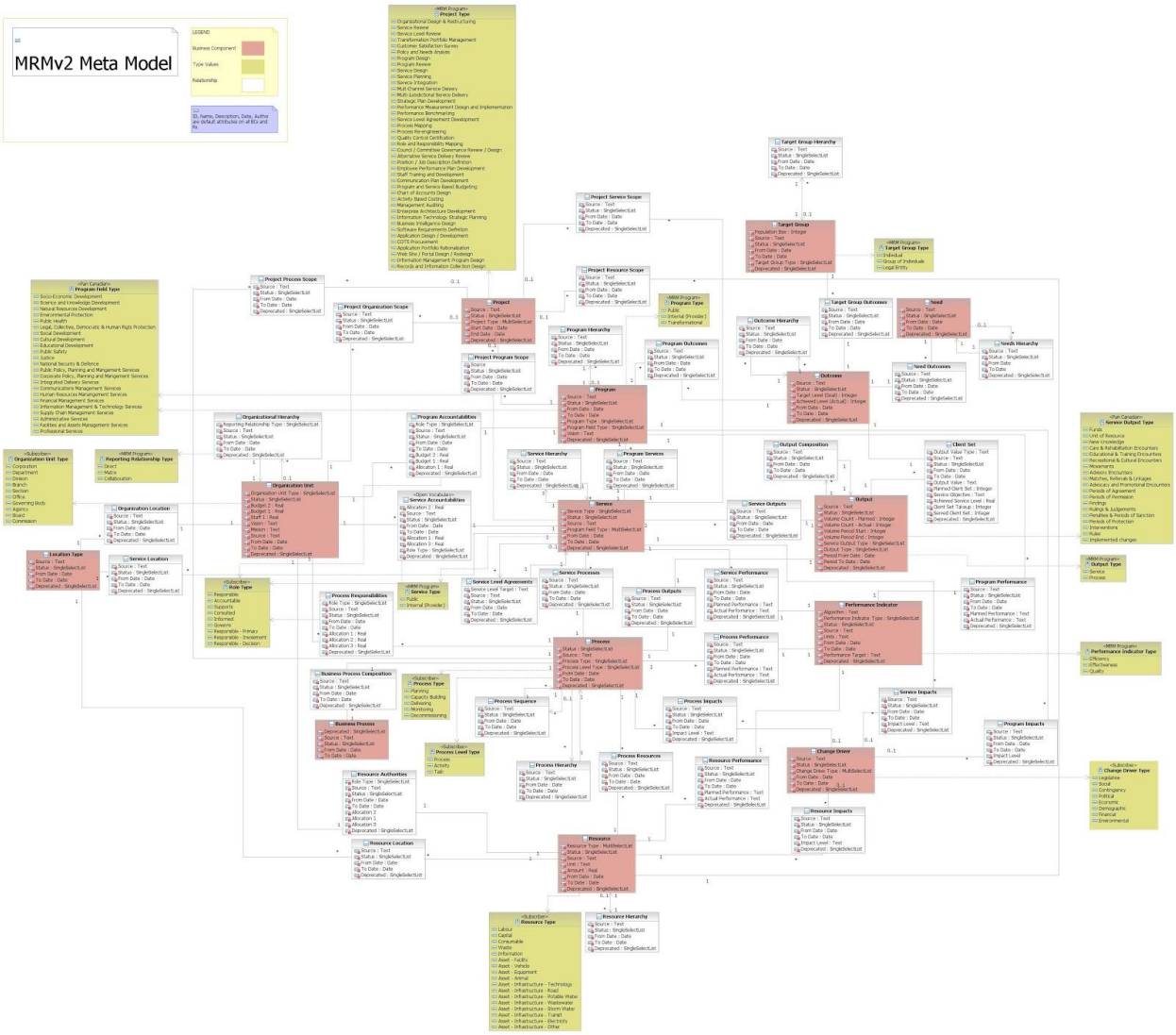


Source: "Development and Implementation of FGDC Standards" http://slideplayer.com/slide/7922445/ Slide 12

Municipal Reference Model

MISA Canada and its partners developed a model entitled the Municipal Reference Model (MRM). This model describes a city's services, processes, programs, customers, and the relationships between all of these entities.[4]

The following visual illustrates the complexity of the MRMv2 Meta model at its highest level. This is included to provide context on the high level of complexity and therefore resourcing required to undertake comprehensive standards development.

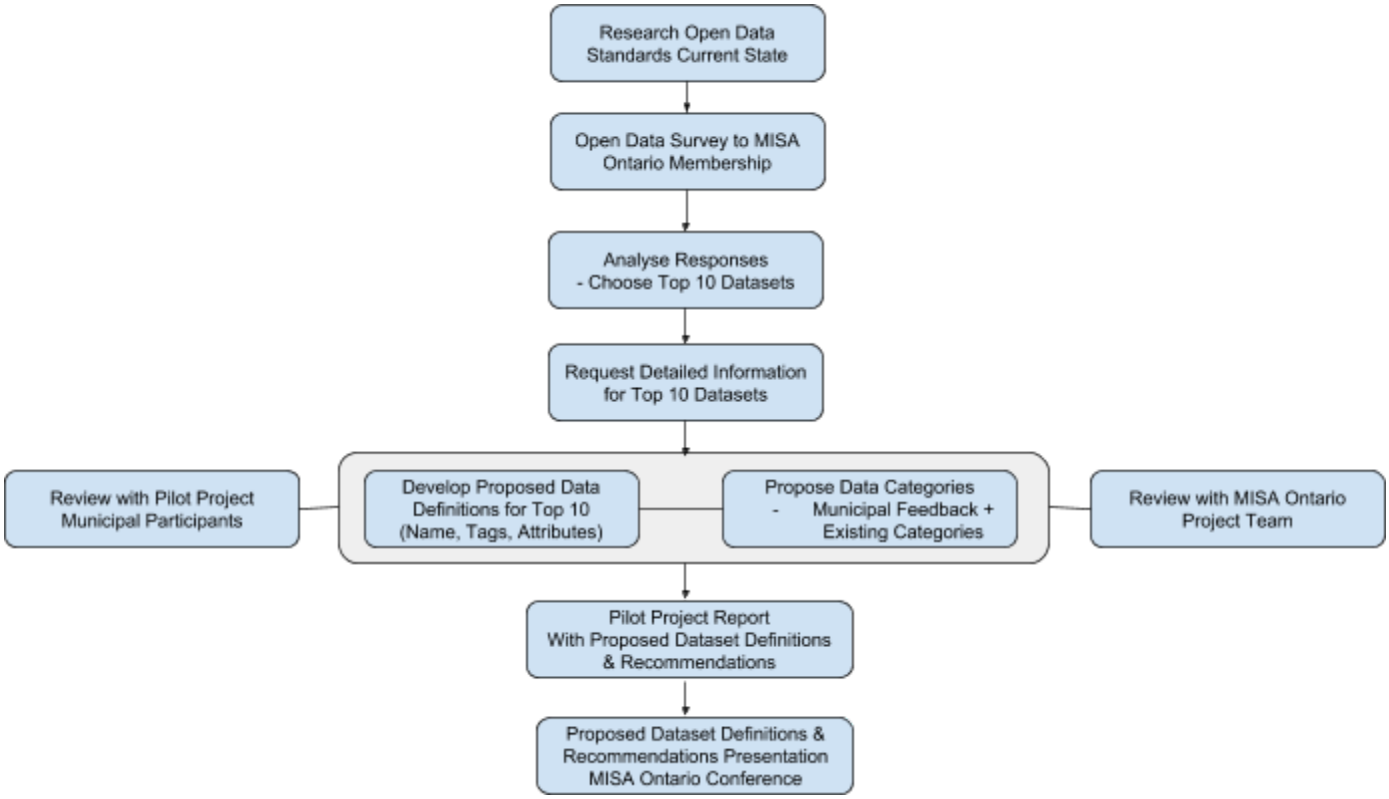


Source: Roy Wiseman, Executive Director , MISA Canada

[4] KPMG International (2014). Services to Local Government: Bringing clarity to city services with the Municipal Reference Model.
https://assets.kpmg.com/content/dam/kpmg/nz/pdf/March/mrm-for-local-government-kpmg-nz.pdf

Approach and Insights

The following illustrates the process flow for the pilot project. This project began with an initial environmental scan of the current state of data standards in Canada. It then proceeded to survey MISA members on their standards practices and top ranked datasets. Ten datasets and a sample of municipalities were chosen to analyse. From this, a set of categories and a dataset definition was developed. Recommendations for future work in standardization of these datasets and expansion of the pilot were also developed.



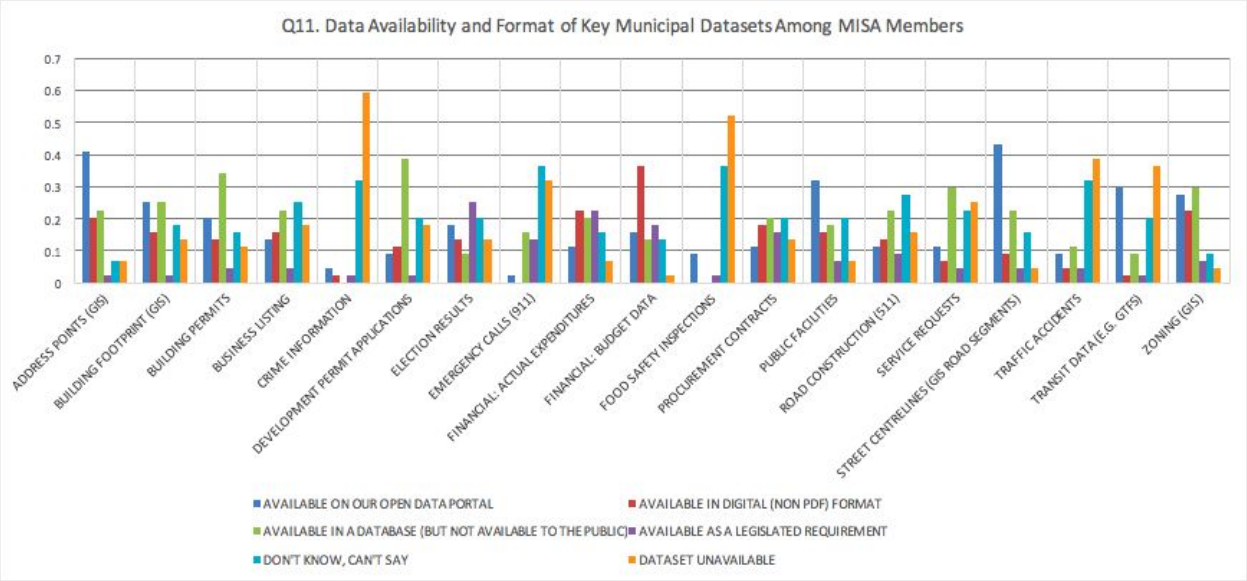
Open Data Survey Insights

The open data survey was sent to the MISA Ontario membership and had a return of 75 respondents. A summary of the survey questions may be found in Appendix 1 of this report. The following are some general observations derived from the survey:

- Respondents varied from municipalities of population of 20 - 100,000 (40%) to those less than 20,000 (22%) and more than 100,000 (38%);
- Their knowledge of open data, from some knowledge to advanced knowledge, was 84% of the respondents;
- Municipalities with open data portals (existing and planned) accounted for 43% of respondents while only 20% of the same group had an open data policy; and
- A corporate information management strategy existed for 22% of respondents while an additional 39% were “working on it at the moment”.

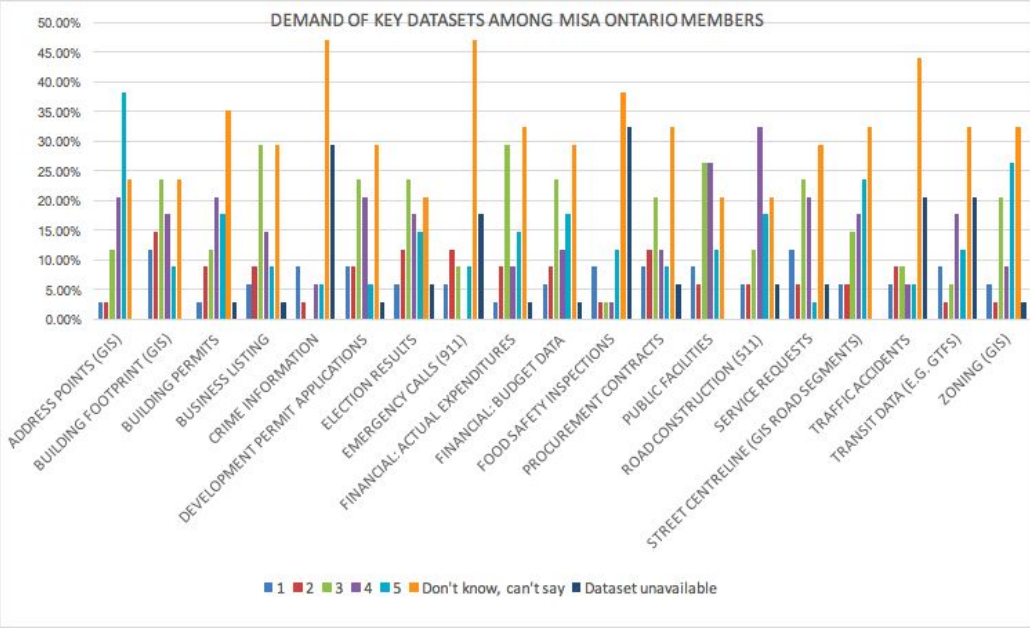
These findings illustrate the diversity of respondents which provides some confidence in the responses to data specific questions. Following are insights impacted open data pilot choices.

Availability of Key Municipal Datasets



As noted in the results, a couple of key datasets such as crime information and food safety inspections were not available as open data. The data that was available as open data was led by **“Street Centrelines (GIS & Road Segments) at 43% (32/75)**.

Demand for Key Municipal Datasets



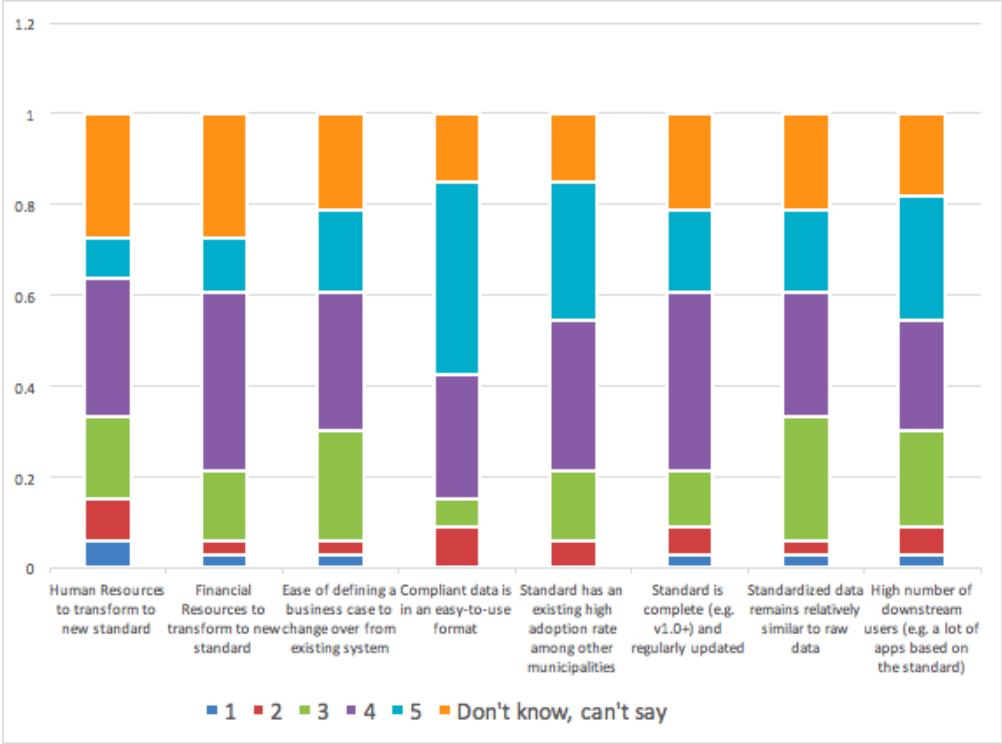
The survey revealed that of the datasets, the demand was identified as **highest for Street Centrelines and Address Points**. The results were fairly similar in the question on perceived benefits for these datasets. These two results led to defining the top 10 datasets used in the pilot project.

The top 10 datasets chosen for a detailed review were (alphabetical order):

- Address Points; Budget; Building Permits; Business Listing; Election Results; Public Facilities; Road Construction; Street Centreline; Transit ; and Zoning (GIS)

Please note that some of the above dataset names have proposed dataset name changes from the pilot project.

Factors in Consideration of Adoption of Standards



For data standards adoption, the above chart illustrates a variety of factors considered by IT personnel. The **most prominent factors related to adoption were adoption rate by other municipalities and compliant data in an easy-to-use format**. Most of the other factors were similar in ranking of importance with the human and financial resource factors being the least significant. These factors were considered in defining the adoption strategy for the proposed new dataset definitions.

Use of the Open Data Survey

The survey provided the input required to define the **top 10 datasets** which had more detailed information provided by participating municipalities. From a demand and availability perspective, the **Street Centreline (now Road Network) dataset was chosen to have a detailed assessment** as found later in the report. Adoption factors were considered in the development of the adoption strategy.

General Dataset Observations

There were significant differences among the municipal participants in the use of names, categories, tags, and attributes. Some municipalities did not have categories or tags for their open data.

Regarding metadata, some did not provide any and several used one or more elements of the recognized geospatial metadata standard. For this reason, the metadata link is provided in each data definition but not the individual elements.

Dataset Name	Issues/Opportunities	Comments
Road Network	Inconsistency with National & Ontario road network standards (see later detailed assessment)	There seems to be a separation of what municipalities need from a road network versus what the senior levels of government are doing. It suggests a need for further work in harmonizing the dataset specification.
Public Facilities	Most municipalities have multiple datasets under this overarching dataset and will be treated differently than other datasets.	Public facilities is being dealt with as a “superset” or an aggregated dataset comprised of multiple sub-aggregates and individual datasets.
Address Points	Recent new address standard NG 9-1-1 (US based, Canada adopted) needs to be considered.	Some municipalities included attributes related to the use of the property or building at the address. These are not included as they may be better attributed to either a building or a land parcel.
Transit Data	Public transit data should be the same for every jurisdiction including regional ones like Metrolinx	Transit data is being dealt with as an aggregated dataset comprised of multiple individual datasets. Some municipalities provide GTFS live data while others do not.
Building Permits	Explore opportunity with Statistics Canada	There is an opportunity to standardize building permit data exchange with Statistics Canada. There has not been a response from the Chief Building Officer association yet.
Road Construction	The differences among municipalities suggests a need to confirm core attribute requirements by domain experts.	Significant differences in the attributes associated with this dataset. This may be linked to road maintenance software used by the organizations.

Zoning (GIS)	Inability to supply zoning data as open data due to licence agreement with Teranet.	Many municipalities license the use of Teranet parcel/property data. This is an open data challenge that should be discussed with the Provincial government. There are more map products than data products for zoning.
Election Results	Different voting counting systems and new rank ballot approaches need to be considered in this dataset definition.	This dataset has not considered differing values based on a ranked ballot system. There may also be variations of data availability depending on the ballot counting system used by the municipality.
Budget	This is another aggregated dataset comprised of operating and capital budgets which require more rigor. Only Toronto had operating budget data for use in the dataset definition. Capital budgets were of two types: project oriented or service oriented.	There are not many municipalities that have released their annual operating and 10 year capital budgets. Additional rigor should be applied based on feedback from municipal financial officers association and financial reporting requirements for Municipal Affairs.
Business Directory	Opportunity to work with EDCO and Province to define core attributes that allow comparison and aggregation.	This dataset has many common core attributes and a similar number of additional attributes among cities. The focus was on the core attributes. Consideration might be given to work with EDCO to standardize this data for their purposes and that of the province from an economic development perspective.

Exploring Gaps in Standardization

Road Networks

In this section, there is an exploration of interrelated issues of discoverability and standards compliance (and standards interoperability) for road network data. Highlighting some of the differences in how road network datasets are catalogued, suggests a deeper question of the hierarchy of terms used to describe the dataset. It is suggested that differences in how a dataset is presented is also linked to fundamental differences in how municipalities work with certain data. Roads are given classifications and are therefore ranked or ordered as part of a municipality's operations. The use of words such as 'lane' or 'roadway' to describe a dataset may also be embedded with the same hierarchy (explicitly or implicitly). Furthermore, roads are separated and therefore classified based on jurisdictional oversight. Discussion of roads, in whatever form, therefore necessarily imply a hierarchy. Assuming that open data stewards have a working knowledge of the data they output, dataset names and descriptions are likely influenced by the working context of a data steward, which in turn is influenced by data models and classification schemes. It is acknowledged that some of the differences are defined by software data models which can vary among municipalities.

Discoverability

Road networks are often described as street or road centreline files. This is because data collected on roads will be measured at the centre of the municipal road allowance and/or paved surface. The term 'centreline' (or centerline) is common in industry and the GIS field, with esri defining a centreline as: "A line digitized along the center of a linear geographic feature, such as a street or a river, that at a large enough scale would be represented by a polygon."

Dataset	Different Names Used	Different Categories	Number of Tags (range)	Number of Attributes (range)
Road Network	5	8	0 - 9	8 - 18

As seen in the table above, a variety of names, categories, tags, and attributes are used across our sampled municipalities.

Naming convention

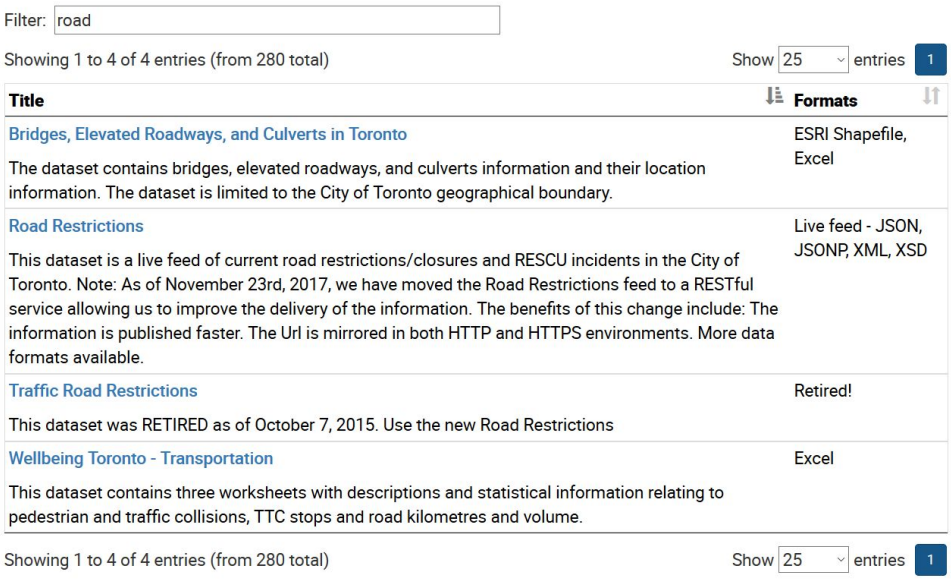
Three municipalities use the term centreline in their dataset names. While this is an accurate name for the dataset, its name does not induce discoverability for non-specialists wishing to find roads or streets. Guelph, on the other hand, has a Guelph Streets dataset, which also happens to be a centreline file. However, this description of streets does not include mention of the word 'road'. Meanwhile, four municipalities, name their datasets with variations on 'road'. Niagara Region, and Ottawa (in the datafiles with its Road dataset), use the term road segment. This is similarly accurate, as roads are represented as line segments, not singular lines, in topology.

While some, such as the City of Welland, explicitly state that their road dataset represents road centrelines, others, such as City of Guelph, do not.

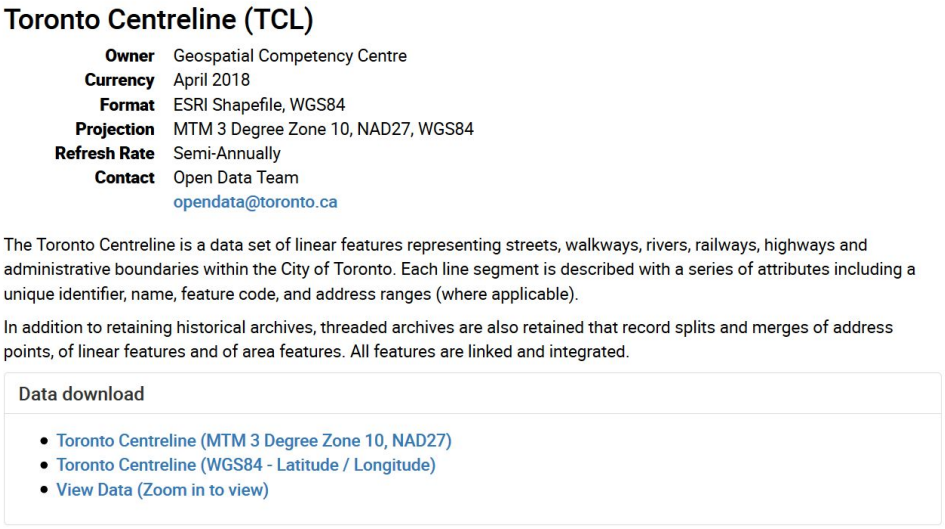
Naming variations in road network datasets

Jurisdiction	Name
City of Brampton	Streets (Centreline)
City of Greater Sudbury	N/A
City of Guelph	Guelph Streets
City of Kitchener	Roads
Niagara Region	Road Segments
City of Ottawa	Roads
City of Toronto	Toronto Centreline (TCL)
City of Welland	Single Line Road Network
City of Windsor	Street Centreline

This may also affect searchability of datasets. A quick search in Toronto's open data catalogue for 'road' reveals four datasets, none of which are the Toronto Centreline (TCL) dataset.



Source: https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-catalogue/



Toronto Centreline (TCL)

Owner Geospatial Competency Centre
Currency April 2018
Format ESRI Shapefile, WGS84
Projection MTM 3 Degree Zone 10, NAD27, WGS84
Refresh Rate Semi-Annually
Contact Open Data Team
opendata@toronto.ca

The Toronto Centreline is a data set of linear features representing streets, walkways, rivers, railways, highways and administrative boundaries within the City of Toronto. Each line segment is described with a series of attributes including a unique identifier, name, feature code, and address ranges (where applicable).

In addition to retaining historical archives, threaded archives are also retained that record splits and merges of address points, of linear features and of area features. All features are linked and integrated.

Data download

- Toronto Centreline (MTM 3 Degree Zone 10, NAD27)
- Toronto Centreline (WGS84 - Latitude / Longitude)
- View Data (Zoom in to view)

Source: https://www.toronto.ca/city-government/data-research-maps/open-data/open-data-catalogue/#e4ec3384-056f-aa59-70f7-9ad7706f31a3

Instead, searches for ‘street’ or ‘centreline’ will return the Toronto Centreline (TCL) dataset. A closer examination of Toronto Centreline (TCL) feature code descriptions reveals that streets exist, but are not defined within the list of objects under Linear Feature Definition. This example of a difference in use of similar terms (road vs street), at different levels of conceptual hierarchy is a potential explanation for inconsistency in how municipalities name their datasets.

Keyword tagging

Discoverability is significantly impacted by the choice of tags used for datasets.

Tagging variations in road network datasets

Jurisdiction	Tag
City of Brampton	street, streets, centreline, road, network, right-of-way, Topographic, Transportation, Infrastructure, Topography and Boundaries
City of Greater Sudbury	N/A
City of Guelph	N/A
City of Kitchener	Road, Roads, Transportation, Highways, City Streets, Municipal Roads, ROW
Niagara Region	Roads, Network, Road Network, Street, Highway, Regional Road, Road, Alley, Artery, Asphalt, Avenue, Boulevard, Byway, Drive, Driving, Expressway, Freeway, Lane, Main Drag, Pavement, Roadway, Route, Thoroughfare, Throughway, Thruway, Turnpike, Viaduct, Access Point, Corridor, Promenade

City of Ottawa	Road, Roadway, Route, Street, Streets, chemin, chemins, routes, rue, rues
City of Toronto	N/A
City of Welland	Road, roads, roadway, roadways
City of Windsor	N/A

A wide range of tagging practices can be found in the sampled municipalities. For example, Ottawa's Road dataset is covered by the following tags: road, roadway, street, streets, chemin, chemins, routes, rue, rues. A multiplicity of tags are used in this case as the city recognises that interpretations and searches for their Roads dataset may be made with any one of those road-related words, and in both English and French. Niagara Region's tag use is even more detailed, with 30 tags corresponding to a range of features not including roads. Niagara's tagging corresponds to multiple classes of road object, including road network. While this may be useful, as the tags correspond to both the purpose of the dataset and its content, an increase in the granularity of tags verges on descriptions of a dataset - a function that could instead be fulfilled by documentation or long descriptions. However, it is valuable to still include as many tags as possible from the end user perspective as portal software can vary in their searching approaches and may not include full text search of all data descriptions and metadata.

This variation is important because not all road datasets contain the same content. Municipalities may choose to disaggregate some road data, such as publishing a separate dataset for one-way streets. Certain types of roads may also be excluded. For example, Niagara Falls (not one of our sampled municipalities), notes that its Road - Centreline dataset "displays the centre line of the road allowance. Road centre line is only shown for roads that are considered 'driveable'. It does include Private Roads but not driveways." When such a distinction is not suggested in its name (Road - Centreline) or tags, data users must rely on municipalities to describe their datasets in metadata or documentation.

Dataset names may make sense in one context, particularly for GIS practitioners, but open data practitioners and the general public who do not have the requisite domain knowledge may not know to search for alternative dataset names. Data discoverability for these datasets would therefore benefit from the use of a range of descriptors in metadata and keyword tags.

Standards compliance and interoperability

The Ontario road network: Road Net Element is Ontario's database of roads for the entire province, created by the Ministry of Natural Resources and Forestry. Interoperability is important to ensure that roads from municipal datasets join up with roads in their surrounding areas. Common standards are therefore important to assess. The Ontario road network (ORN) has its own data model and schema called the Ontario Road Network (ORN) Data Standard for Geometry and Attributes. This standard describes roads, road events, and other characteristics including the proper formatting of addresses and street names. The ORN is also in compliance with the North American Profile of ISO 19115 Geographic Information Metadata, a geographic data schema. Importantly, the Government of Canada has also adopted this standard. While compliance with metadata is desirable, it is essential that the data standard/model for the actual datasets be harmonized to ensure the optimum degree of interoperability.

Metadata

Our sample of municipalities revealed variations in metadata provided for road datasets. While four municipalities provide details of on the datum and projection (important if grid coordinates in use) used in their road dataset (e.g., Niagara Region, Welland, Windsor, Toronto), four (e.g., Brampton, Guelph, Kitchener, Ottawa) did not. This suggests inconsistency in the application catalogue metadata standards. Ideally from a global perspective dataset definitions should include links to their respective data models and is another consideration for participant municipalities provide to consider.

Standards such as DCAT (Data Catalog Vocabulary) and the Dublin Core Metadata Initiative (DCMI) help define description of open data in a data catalogue. While they are content agnostic, they contain specific fields for describing spatial characteristics of datasets. A stronger linking to spatial standards may provide stronger evidence (to the end user) of a municipal dataset's interoperability with databases and data models such as the ORN. To achieve this, datasets need to provide links to the relevant spatial data standard they conform to (in this case NAP ISO 19115), in DCAT's dct:spatial property. It's interesting to note that there still exists the common interchangeable use of spatial and geospatial by practitioners with spatial including the CAD environment geospatial referencing the common geographic reference system (latitude and longitude).

Hierarchies of Terms

Different applications of similar road-related terms suggests inconsistency in how terms are defined. One area to turn to for explanation are the data models upon which road datasets are based. Because the software used to create data is based upon data models, data models likely influence the internal naming practices of institutions.

Table Name: ORN_ROAD_CLASS_LIST

ROAD_CLASS	ROAD_CLASS_DESCR	NRN_ROADCLASS
Alleyway / Laneway	A low speed thoroughfare dedicated to provide access to the rear of properties.	8
Arterial	A major thoroughfare with medium to large traffic capacity	3
Collector	A minor thoroughfare mainly used to access properties and to feed traffic with right of way.	4
Expressway / Highway	A high-speed thoroughfare with a combination of controlled access and intersections at grade level.	2
Freeway	An unimpeded, high speed controlled access thoroughfare for through traffic with typically no at grade intersections, usually with no property access or direct access and which is accessed by a ramp. Pedestrians prohibited.	1
Local / Strata	A low speed thoroughfare dedicated to provide access to properties with potential public restriction, trailer parks, First Nations, strata or private estates.	6
Local / Street	A low speed thoroughfare dedicated to provide full access to the front of properties.	7
Local / Unknown	A low speed thoroughfare dedicated to provide access to the front of properties but for which the access regulations are unknown.	5
Ramp	A system of interconnecting roadways providing for the controlled movement between two or more roadways.	9
Rapid Transit	A thoroughfare restricted 24 hours a day, for the sole use of public transportation buses.	11
Resource / Recreation	A narrow passage which has as a primary function access for resources extraction and also may have a role in providing an access for the public to back country.	10
Service	A stretch of road permitting vehicles to come to a stop along a Freeway or Highway. These include weigh scales, emergency lanes, lookouts and rest areas.	12
Winter	A road that is only useable during the winter months when conditions allow for passage over lakes, rivers and wetlands.	13

ORN Data Standard for Geometry and Attributes, p. 57 Source: https://dr6j45jk9xcmk.cloudfront.net/documents/1866/go-its-29-ontario-road-network-orn.pdf

Related to the issue of discoverability, the ORN standard defines the types of roads within a road network. In the figure above, it can be seen that roads are classified into into laneways, freeways, streets and more. This is a representation of the classification system of roads for Ontario. The Federal Government's National Road Network also uses the same classification system (and codes) to define roads.

Yet, our examples of road datasets named or described with the term 'street' suggests that this hierarchical differentiation is not applied consistently, hence a lack of standardization in naming convention. These issues are seen in differences in classification systems for roads themselves. Toronto's road classification system has only five categories, all of which use the term 'road': local road, collector road, minor arterial road, major arterial road, expressway. Notably, Toronto's road classification system does not define streets. Instead, street is used interchangeably with road. Other municipalities, such as the City of Ottawa's classification system, also differ in their definitions and terminology for streets or roads. Notably, road classifications appear to have no bearing on the naming or tagging of datasets - technical definitions of 'road' are not reflected in dataset names and keywords (from our sample of municipalities), which may not fulfill expert data user needs. Information on standards compliance and interoperability is important to aid a data user's data exploration.

As seen in the figure above, a variety of sources can be used to define a dataset's name, categories and keywords. For the follow-on to this pilot project, defining the proper use of names and tags will require an understanding of keyword definitions for each domain, some of which can be found in existing data models.

The project has explored some of the variations in which a single dataset can be described and catalogued in an open data catalogue. Fundamental differences in how cities answer the question, "what is a road?" can be seen when exploring data models and classifications. Different cities have different descriptions for 'road' and 'street', and the two terms can be found to be used interchangeably. This is important, as municipal software systems and data models represent a municipal government's view on how to describe a road dataset - these views may differ from open data users in the public, who may have limited understanding of these definitions.

Any effort to standardise an open dataset, such as road data, should take into account publicly perceived definitions of terms, which will influence the choice of keyword tags and descriptions, as well as the internal data practices of municipal governments. Choice of words for dataset names, categories, keyword tags, and even long descriptions, will influence user expectations of what is contained within a given dataset.

Aligning Open Data Content with Descriptors

From our exploration of data from the municipalities sampled, it was identified that a number of key elements required to effectively define a dataset, which is called a **dataset definition**. This includes: dataset name, revision history, dataset description, related standards, dataset category (see below), dataset tags/keywords, dataset formats, metadata and dataset attributes (data dictionary).

The dataset definition describes the content that should be included in a dataset, the structure they should adhere to, the standards the data should comply with, and surrounding information that describes the characteristics of this dataset (metadata).

It is hoped that inclusion of these fields raises the minimum level of quality for defining the content and structure of municipal datasets, and their metadata. These recommendations cover both discoverability and content of datasets, and attempts to align front-facing (catalogue/portal view) information on a dataset with the content of a dataset. If metadata surrounding a dataset accurately describes what is contained within, and if the content of a dataset can be standardised, datasets will match user expectations. In the larger view, for datasets to be internally consistent, there is a need to consider the larger enterprise data model and interoperability among various business systems. By user expectations, it is meant that a dataset's name (and other descriptors such as categorisation and keyword tags) is reflective of the content contained within it.

Data Categories

The definition of data categories should take into account a couple of lenses: i) the external end user, and ii) the public sector community. The former of these are the people that are trying to discover the open data and potentially use it and as such,it is important to endeavour to think in their terms how they might classify a specific dataset. The second perspective needs to realize that categories should facilitate ease of finding data in terms of public services. This latter perspective could bthereenefit from a review the Municipal Reference Model (MRMv2) for the meta model of the high level services.

Based on the foregoing, the following are the proposed data categories for this open data standards pilot project:

1. *Administrative Boundaries*. Location based boundaries are used to define legal (e.g. municipal boundary) or operational boundaries (e.g. election wards, garbage collection zones).
2. *Business and Economy*. Services and related data that supports the business community and the local and regional economic development.
3. *Community Services*. Social services that support the wellbeing of individuals, families and the community at large.
4. *Environment*. The entire natural ecosystem of air, water, land and biodiversity and all related inter-relationship data.

5. *Health and Safety*. Services data related to Fire, Police, Emergency Medical Services and hospital and health care data.
6. *Infrastructure and Facilities*. Data related to hard infrastructure assets (e.g. sewer and water networks) and municipal facilities such as community centres and City Hall.
7. *Land Development*. Data related to the land development process (e.g. zoning change application, building permits), zoning and official land development plans.
8. *Local Government*. Data related to the operation of local government including elections and contact information for public services, complaints and events.
9. *Location* (Geospatial). Address points, (centreline) road network, location markers (horizontal and vertical control), E9-1-1 locations.
10. *Maps and Imagery.* Municipal map products and interactive mapping together with terrestrial and remote sensed imagery.
11. *Parks and Recreation*. This category includes data on recreation facility assets (e.g. parks, tennis courts, arenas), cultural assets (e.g. museums), libraries and recreation program data.
12. *Transportation*. The entire transportation ecosystem including roads, bike paths, public transit, transportation planning and related statistical data.

These categories can contain datasets that are either assets and/or services. The value of utilizing categories for datasets is a tool to assist end users in finding that data they are looking for. It should be noted that in some cases, a specific dataset may be related to more than one category and this is not an issue but rather an assistance for the discoverability of that dataset.

It should be noted that these categories are defined primarily from the perspective of local government although a few categories are relevant from local to global (e.g. environment). Many open data portals employ category icons but this was not considered as part of this pilot project.

Dataset Definitions

In Appendix 2, there is a draft dataset definition for each of this pilot projects top ten datasets. It is intended that these definitions be extensible and developed through collaboration and data user engagement. Recommendations are made on the minimum details a municipality should include in their physical dataset and the metadata describing said dataset. These include keyword tags and names, but also generic information such as data formats and links to related standards. By defining related standards the dataset should conform to, a data user will have more confidence they can import and compare multiple datasets without requiring transformation. This may include important regulations and directives that govern a piece of data. For example, linking building permit data to the local building code act contextualises the dataset within the local regulatory environment. Geospatial data that provides information on the datum and coordinate system allows users to know whether spatial transformations are needed to compare two sets of data.

Standards Adoption Strategy

The proposed strategy for adoption of data set definitions focuses on the following key elements:

- Simple Process
- Education
- Ongoing community of practice support
- Vendor support

Simple Process
Understanding that adopting any new standard can be resource intensive. A set of initial steps could be taken to ease the process of standards adoption. For adoption of the dataset definitions proposed in this pilot project, the following actions are proposed:

- Leverage ETL (Extraction, Transformation, Loading) software to support conversion;
    - These software support the conversion of datasets from one database format to another. ETL will facilitate changing dataset naming for public open data catalogue/portal.
- Prototype the adoption process in two prominent portal environments;
- Prepare a step by step DIY guide for dataset definition adoption.

Education
Standardization also requires awareness of resources available and socialization of concepts. In this regard, it is recommended:

- Presentation of report and proposed dataset definitions at the MISA Ontario annual conference;
- Webinar open to all MISA Ontario municipal members to introduce the pilot, the dataset definitions and discussion of adoption approach;
- Webinar with applicable vendor community to review their role and discuss support for their clients in adopting the dataset definitions; and
- Webinar and local meetups with the open data user community to provide insight into the data field mapping that is proposed for the open data portal.

Community of Practice and User Support
Communities of practice already exist around data standards in different domains, but have yet to be coordinated around open data practices. Peer and user networks are needed to support standardization of municipal open data services. It is recommended:

- Create a broad user group of researchers, local community groups, application developers and internal government users of open data to provide feedback on proposed changes to the open data services. This may become an ongoing user support mechanism for municipal open data programs.
- Examine the role of the MISA Canada Open Data SIG and recent GCcollab open data standards group discussions in defining peer-to-peer support

Vendor Support

Success in this pilot project would benefit from support from software vendors that supply data management systems. A prior OpenNorth report on standards adoption has identified the need for software vendor support in promoting data standardization. Since software tools (that government relies on for data collection, processing, analysis, and publication) govern data structuration, vendor support for consolidated data categories and other aspects of our dataset definition is needed. Without vendor support, collaborative standards adoption is less likely to be successful.

Recommendations for Moving Beyond a Pilot

Based on our exploration of datasets and development of a 'dataset definition', we present a number of short and long term recommendations for working with municipalities to standardize their open data.

Expand Collaboration

1. *Expand collaboration and partnering to MISA Canada Chapters*. Subsequent to this pilot project and similar ventures, it is recommended that other Chapters of MISA Canada and the Open Data Special Interest Group be invited to provide feedback on the pilot project results and become involved in future open data standards initiatives.

2. *Create a Open Data Community of Practice*. Establishing a community around open data standardization will bring together stakeholders from outside the municipal sector, including domain experts, researchers, other levels of government, and international expertise. This community will support the shift towards greater discoverability and interoperability of open data.

3. Develop Agreements/MOUs with External Organizations. There are opportunities to develop "partnership" agreements or MOUs with external organizations including the provincial and federal governments, LOLA (Linked Organization of Local Authorities) and related subject matter expert organizations such as the Ontario Good Roads Association (OGRA) and the AMCTO.

Complete Products and Tools

1. *Finalize Version 1.0 of Dataset Definition*. Complete community engagement to finalize a version 1.0 of the pilot project open dataset definition.

2. *Undertake Prototyping to Validate Change Processes.* As part of the development of Version 1.0 (above), undertake several prototyping exercises with different platforms to develop change processes and assess existing tools or develop new tools if required to support adoption

3. *Develop a DIY guide for dataset definition adoption*. To simplify the adoption of open data standards, develop a Do-It-Yourself Open Data Standards Adoption toolkit that could include: step-by-step processes, database structures, templates and resources.

4. *Engage municipalities to increase their dataset metadata using the dataset definition*. This will allow municipalities to expand their metadata offerings in a documented process that can be refined and repeated.

Plan for the Future

1. *Establish a plan to develop a municipal data reference model*. Using this open data pilot project as a start point, develop a plan to create a comprehensive Municipal Data Reference Model that can be utilized for open data, shared data and closed data management.

2. *Develop an Integrated Multi-Jurisdictional Canadian Data Infrastructure*. There is a need to bring together disparate data related projects and individual data governances to establish a digital data infrastructure as the backbone for a “National Data Strategy” that includes both urban and rural municipalities.

3. *Design a multi-jurisdictional Canadian Data Infrastructure.* Expand scope of work to include all levels of government. With a goal of interoperability and discoverability of open datasets, inclusion of all levels of government, with a framing within the Open Data Charter principles (which have been adopted by Federal, provincial, and municipal governments in Canada) will enable work towards federated search.

Appendices

Appendix 1: MISA Open Data Member Survey Summary

The survey was conducted by MISA Ontario to its members in March 2018. The following provides the content of that survey.

Introduction
This survey is undertaken as part of a pilot project to define standards around key municipal open datasets according to their current availability status, perceived demand and value, together with drivers of standardization to promote greater interoperability. The datasets identified in the following survey are currently deemed to be important municipal open datasets. The survey was designed to take no more than 7-10 minutes to complete. Assisted by the results of the survey, MISA Ontario will select 10 municipalities to work on this pilot project to standardize "top10" datasets across these municipalities. If you have any questions or require clarification on the survey, please contact info@misa.on.ca . Thank you for taking the time to complete this survey.

Section 1: Participant Information

- Municipality Name, Population, MISA Ontario Region, IT employees, title/role of respondent, level of open data knowledge, open data group participation

Section 2: Open Data and Information Management Status

- Open data management
- Information Management strategy
- Availability of open datasets (list of 19 used)
    - Address Points, Building Footprints, Building Permits, Business LIsting, Crime Information, Development Permit Applications, Election Results, Emergency Calls (9-1-1), Financial: Actual Expenditures, Financial: Budget Data, Food Safety Inspections, Procurement Contracts, Public Facilities, Road Construction (511), Service Requests, Street Centreline (GIS road segments), Traffic Accidents, Transit data (e.g. GTFS), Zoning (GIS)

Section 3: Perceptions of Demand and Benefits
Using the list of 19 datasets from section 2, the following questions asked for rating:

- How would you assess the demand, whether internal of external, for the datasets?
- What are the perceived external benefits for the datasets?
- What are the perceived internal benefits for the datasets?

Section 4: Need for Data Standards

- Where do you currently get information about data standards?
- When making decisions on adoption of data standards, rate the following adoption drivers
    - Human resources, financial resources, ease of defining business case, compliant data is easy-to-use format, high adoption rate with municipalities, standard is complete, standardized data similar to raw data, high number of downstream users

Appendix 2: Dataset Definitions

The following are the proposed definitions for the datasets reviewed in this pilot project. They are defined as being Version 0.1 which invites further feedback towards fine tuning to a version 1.0. Here is the order of the dataset definitions provided (in alphabetical order of new dataset names):

1. Address Points
2. Budget (Operating and 10 Year Capital)
3. Building Permits
4. Business Directory
5. Election Results
6. Public Facilities (aggregated dataset)
7. Road Construction
8. Road Network
9. Transit (aggregated dataset)
10. Zoning (GIS)

Address Points

Open Dataset Definition - Address Points					
Dataset Name: Address Points					
Revision History					
Date (YYYY/MM/DD per ISO 8601)	Version	Updates	Contact		
2018-05-07	0.1	Initial Draft	info@opennorth.ca		
Dataset Description					
Short Version:	Dataset of all the address points existing within the municipality.				
Complete Description:	Municipal address points are important for many applications related to location searches. This dataset provides information on the municipal address value and the actual location of the point may reflect several methods of references the address. The address location may be referenced relative to the centroid of the property, the centroid of a building footprint or rooftop, the location of the physical entrance or a midpoint in the property frontage.				

Dataset Catgeory:	Location				
Related Standards					
Dataset Tags/Keywords (for publishing)					
Dataset Tags	address, points, addresses, addressing, property, property address				
Dataset Formats					
Original:	ESRI shp file, Geodatabase				
Published formats:	Shp, Json, GeoJson, GML, KML				
Dataset Metadata					
Geospatial Data - Reference ISO 19115 North American Profile http://www.nrcan.gc.ca/earth-sciences/geomatics/canadas-spatial-data-infrastructure/standards-policies/8912					
Data Dictionary of Attributes					
Attribute Name	Field Name		Description		
Municipality Name	MUN_NAME		Name of the municipality associated with the address point.		
Address Identification	ADD_ID		Civic Address identification number		
Street Name	STREET_NAME		Street Name		
Civic Address	ADDRESS		Civic Address on the street		
Civic Address Unit	ADD_UNIT		Unit number associated with civic address if applicable		
Building identification	BLDG_ID		Building identification associated with address point		
Parcel identification	PARCEL_ID		Parcel identification associated with address point		
Address Point Reference	ADD_REF		Location Reference: property centroid, building centroid, building entrance, frontage point		
Latitude	ADD_LAT		Latitude of civic address		
Longitude	ADD_LON		Longitude of civic address		

Budget: Annual Operating

Open Dataset Definition - Budget: Annual Operating

Dataset Name: Budget : Annual Operating

Revision History					
Date (YYYY/MM/DD per ISO 8601)	Version	Updates	Contact		
	0.1	Initial Draft	info@opennorth.ca		

NOTE: PRELIMINARY RESULT BASED SOLELY ON CITY OF TORONTO DATA

Dataset Description					
Short Version:	Annual operating budget of the municipality.				
Complete Description:					
Dataset Category	Local Government				

Related Standards

Dataset Tags/Keywords (for publishing)	
Dataset Tags	Local Government, operating budget, finances

Dataset Formats					
Original:	CSV				
Published formats:	CSV, JSON, XML				

Dataset Metadata	
Non-spatial: Dublin Core Elements	http://www.dublincore.org/documents/dces/

Data Dictionary of Attributes			
Attribute Name	Field Name		Description
Toronto has a service based budgeting approach and the following are the attributes.			

Program	BUDGET_PROGRAM		Name of program or division
Service	PROGRAM_SERVICE		A detailed account of key offerings that are associated with each program currently delivered by a division.
Activity	SERVICE_ACTIVITY		A sub-set of the defined service, which includes unique processes and a discrete output delivered to the client(s)
Category name	ACTIVITY_CATEGORY		Category Name of Expense
Expense/Revenue	BUDGET_TYPE		Budget amounts defined as an expense or a revenue
Year budget	YEAR_BUDGET		Recommended or Approved Budget of the budget year

Budget: Ten Year Capital

Open Dataset Definition - Budget:Ten Year Capital					
Dataset Name: Budget: 10 Year Capital					
Revision History					
Date (YYYY/MM/DD per ISO 8601)	Version	Updates	Contact		
2018-05-07	0.1	Initial Draft	info@opennorth.ca		
Dataset Description					
Short Version:	The municipal 10 capital costs budget.				
Complete Description:					
Dataset Category	Local Government				
Related Standards					
Dataset Tags/Keywords (for publishing)					
Dataset Tags	Local Government, operating budget, finances,				
Dataset Formats					
Original:	CSV				
Published formats:	CSV, JSON, XML				
Dataset Metadata					
Non-spatial: Dublin Core Elements	http://www.dublincore.org/documents/dces/				
Data Dictionary of Attributes					
Attribute Name	Field Name		Description		
NOTE: The attributes are based on City of Kitchener Capital Budget dataset					
Project idenitifcation	PROJECT_NUM		Project number		

Capital budget identification	CAPITAL_NUM		Capital budget assigned number
Project description	PROJECT_DESC		Capital project description
Funding Source	FUND_SOURCE		Capital budget funding source
Approved budget Year 1	BUDGET_YEAR1		Approved capital budget value year 1
Projected budget year 2	BUDGET_YEAR2		Projected capital budget year 2
Projected budget year 3	BUDGET_YEAR3		Projected capital budget year 3
Projected budget year 4	BUDGET_YEAR4		Projected capital budget year 4
Projected budget year 5	BUDGET_YEAR5		Projected capital budget year 5
Projected budget year 6	BUDGET_YEAR6		Projected capital budget year 6
Projected budget year 7	BUDGET_YEAR7		Projected capital budget year 7
Projected budget year 8	BUDGET_YEAR8		Projected capital budget year 8
Projected budget year 9	BUDGET_YEAR9		Projected capital budget year 9
Projected budget year 10	BUDGET_YEAR10		Projected capital budget year 10
Department name	BUDGET_DEPT		Department name associated with capital project
Division name	BUDGET_DIV		Division name associated with capital project
Section name	BUDGET_SECTION		Section name associated with capital project

Building Permits

Open Dataset Definition - Building Permits

Dataset Name: Building Permits

Revision History					
Date (YYYY/MM/DD per ISO 8601)	Version	Updates	Contact		
2018-05-07	0.1	Initial Draft	info@opennorth.ca		

Dataset Description	
Short Version:	This dataset contains information on the location of permits for either the construction, renovation and demolition of structures and/or facilities.
Complete Description:	"The body responsible for enforcing Ontario's Building Code in your area issues permits for the construction, renovation, demolition and certain changes of use of buildings, and for the installation, alteration, extension or repair of on-site sewage systems." This dataset contains information related to the type of permits, dates associated with the permitting process, location associated with the permit and information related to the work undertaken and value. (Adapted from Ontario Ministry of Municipal Affairs website)
Dataset Catgeory:	Land Development

Related Standards	
Building Code Act	https://www.ontario.ca/laws/statute/92b23

Dataset Tags/Keywords (for publishing)	
Dataset Tags	Land Development, Building, Housing, Building Permits, permits, property, parcel, building, permit activity, structures, pools, renovation, demolition, septic systems

Dataset Formats	
Original:	CSV
Published formats:	CSV, JSON, XML

Dataset Metadata					
Non-spatial: Dublin Core Elements	http://www.dublincore.org/documents/dces/				
Data Dictionary of Attributes					
Attribute Name	Field Name		Description		
Permit Number	PERMIT_NUM		System or manually assigned building permit number		
Permit Revision Number	PERMIT_REV_NUM		System or manually assigned permit revision number		
Permit Type	PERMIT_TYPE		Permit types for construction, renovation and demolition.		
Structure Type	STRUCTURE_TYPE		Type of building or facility that is being built or changed		
Work Type	WORK_TYPE		Type of work including new construction, renovation or demolition		
Work Sub-Type	WORK_SUBTYPE		Work sub-type such as plumbing, electrical, etc.		
Work Description	WORK_DESC		Description of work being undertaken for this permit.		
Work Contractor	WORK_CONTRACTOR		Name of work contractor		
Work Contractor Contact Information	CONTRACTOR_INFO		Contact information for work contractor		
Permit Application Date	APPLY_DATE		Date the application was provided to municipality		
Permit Issue Date	ISSUE_DATE		Date the municipality issued the building permit		
Permit Expiry Date	EXPIRE_DATE		Date the building permit expires		
Permit Completion Date	COMPLETE_DATE		Date the work has been completed and approved		
Permit Status	PERMIT_STATUS		Permit status : pending, active, closed		
Permit Address	PERMIT_ADDRESS		Civic address for the permit request		
Permit Postal Code	PERMIT_PC		Postal code for the permit request		
Permit Ward Number	PERMIT_WARD		Municipal ward that the permit related to		
Property Legal Description	LEGAL_DESC		Legal description of the property (lot, concession, register and reference plan numbers		
Current Land Use	LANDUSE_C		Existing land use zoning		

	URRENT		
Proposed Land Use	LANDUSE_PROPOSE		New land use zoning once approved
Dwelling Units Gained	UNITS_GAINED		Number of new units related to this permit
Dwelling Units Lost	UNITS_LOST		Number of existing units lost related to this permit
Existing Gross Floor Area (GFA)	GFA_EXIST		Existing gross floor area in square metres
Proposed Gross Floor Area (GFA)	GFA_PROPOSE		Proposed gross floor area in square metres from this work
Construction Value	CONSTRUCT_VALUE		Value of the construction project
Occupancy Permit Issued	OCCUPY_DATE		Work has been inspected and approved for occupancy
Location - Centroid Latitude	PERMIT_LAT		Latitude of the centroid of the project area
Location - Centroid Longitude	PERMIT_LONG		Longitude of the centroid of the project area

Open Dataset Definition - Business Directory

Dataset Name: Business Directory

Revision History					
Date (YYYY/MM/DD per ISO 8601)	Version	Updates	Contact		
2018-05-07	0.1	Initial Draft	info@opennorth.ca		

Dataset Description	
Short Version:	This dataset contains information on the businesses operating within the muncipal boundaries.
Complete Description:	The municipal business directory dataset provides an invenotry of all known businesses operaqting within the municipality. It includes data related to type of business, number of employees and contact information.
Dataset Catgeory:	Business and Economy

Related Standards	
Industry NAICS Codes	https://www.canada.ca/en/revenue-agency/services/tax/businesses/topics/sole-proprietorships-partnerships/report-business-income-expenses/industry-codes.html

Dataset Tags/Keywords (for publishing)	
Dataset Tags	Business and Economy, directory, economic, improvement, listings, survey, business directory, business, small business, entrepreneur, business license, business incubation

Dataset Formats	
Original:	CSV
Published formats:	CSV, JSON, XML

Dataset Metadata	
Non-spatial: Dublin Core Elements	http://www.dublincore.org/documents/dces/

Data Dictionary of Attributes			
Attribute Name	Field Name		Description
Company Name	COMPANY_NAME		Operating name of company
Road Name	ROAD_NAME		Road name of business location
Road Address Number	ROAD_ADDNUM		Road address number of business location
Address Unit Number	ADD_UNIT		Unit number associated with road address if applicable
Postal Code	POSTAL_CODE		Canadian Postal Code
NAICS code primary	NAICS_PRIMARY		Primary type of business - NAICS code.
NAICS code secondary	NAICS_SECONDARY		Secondary type of business - NAICS code.(if applicable)
Year started	YEAR_START		Year the business began in this municipality
Business park	BUSINESS_PARK		Business park associated with business (if applicable)
Total employees full time	EMPLOYEES_FULL		Total number of fulltime employees
Total employees part time	EMPLOYEES_PART		Total number of part time employees
Export percentage	EXPORT_PERCENT		Percentage of sales that are exported
Contact telephone	CONTACT_PHONE		Telephone number to contact company
Contact email	CONTACT_EMAIL		Email address for company
First Name 1	FIRST_NAME1		First name of primary contact
Last Name 1	LAST_NAME1		Last name of primary contact
Title 1	TITLE1		Title of the primary contact
First Name 2	FIRST_NAME2		First name of secondary contact
Last Name 2	LAST_NAME2		Last name of secondary contact
Title 2	TITLE2		Title of secondary contact
Business license number	BUSINESS_LIC_NUM		Business licence number

Election Results

Open Dataset Definition - Election Results					
Dataset Name: Election Results					
Revision History					
Date (YYYY/MM/DD per ISO 8601)	Version	Updates	Contact		
2018-05-07	0.1	Initial Draft	info@open north.ca		
Dataset Description					
Short Version:	The municipal election results from the most recent election.				
Complete Description:					
Dataset Category	Local Government				
Related Standards					
Municipal Elections Act	https://www.ontario.ca/laws/statute/96m32				
Dataset Tags/Keywords (for publishing)					
Dataset Tags	Local Government, Elections, polling, voter turnout, stations, polling stations, voting locations, election results				
Dataset Formats					
Original:	CSV				
Published formats:	CSV, JSON, XML				
Dataset Metadata					
Non-spatial: Dublin Core Elements	http://www.dublincore.org/documents/dces/				
Data Dictionary of Attributes					
Attribute Name	Field Name		Description		

Election/By-election year	ELECT_YR		Election or by-election year
Candidate List	ELECT-CANDIDATES		List of all municipal election candidates (includes regional Council, Chair {if applicable} and School Boards where applicable)
Ward	WARD		Municipal ward number or description
Poll number	POLL_NUM		Poll identification number
Poll location	POLL_LOCATION		Civic address for the poll
Eligible voters total by ward	WARD_VOTERNUM		Total eligible voters by ward
Votes from Advance voting	VOTE_ADVANCE		Number of votes cast in advance polls
Votes by proxy	VOTE_PROXY		Number of votes cast by proxy
Votes on election day	VOTE_ELECDAY		Number of votes cast on election day
Voter Turnout	VOTE_TURNOUT		Voter turnout overall and by ward
Voting method	VOTE_METHOD		In person or online (if applicable).
By-Election	BY_ELECTION		Yes or No response to whether this is a by-election
Election results	VOTE_RESULT		Election counts for all candidates

Open Aggregate Dataset Definition - Public Facilities: Community Services

Dataset Name: Public Facilities: Community Services

Revision History					
Date (YYYY/MM/DD per ISO 8601)	Version	Updates	Contact		
2018-05-07	0.1	Initial Draft	info@opennorth.ca		
Dataset Description					
Short Version:	This definition represents an aggregate dataset of multiple datasets. The individual dataset definitions will be found in future definitions.				
Complete Description:					
Dataset Category	Community Services				
Related Standards					
Not applicable as this is an aggregate dataset					
Dataset Tags/Keywords (for publishing)					
Dataset Tags	Community Services				
Dataset Formats					
Original:	ESRI shp file, Geodatabase				
Published formats:	Shp, Json, GeoJson, GML, KML				
Dataset Metadata					
Geospatial Data - Reference ISO 19115 North American Profile					

http://www.nrcan.gc.ca/earth-sciences/geomatics/canadas-spatial-data-infrastructure/standards-policies/8912					
Individual Dataset Names					
Dataset Name			Description		
Police Stations			Location and name/id of police stations		
Hospitals			Location and name of local hospitals		
Fire Stations			Location and name/id of fire stations		

Public Facilities: Infrastructure and Facilities

Open Aggregate Dataset Definition - Public Facilities: Infrastructure and Facilities					
Dataset Name: Public Facilities: Infrastructure and Facilities					
Revision History					
Date (YYYY/MM/DD per ISO 8601)	Version	Updates	Contact		
2018-05-07	0.1	Initial Draft	info@opennorth.ca		
Dataset Description					
Short Version:	This definition represents an aggregate dataset of multiple datasets. The individual dataset definitions will be found in future definitions.				
Complete Description:					
Dataset Category:	Infrastructure and Facilities				
Related Standards					
Not applicable as this is an aggregate dataset					
Dataset Tags/Keywords (for publishing)					

Dataset Tags	Infrastructure and Facilities				
Dataset Formats					
Original:	ESRI shp file, Geodatabase				
Published formats:	Shp, Json, GeoJson, GML, KML				
Dataset Metadata					
Geospatial Data - Reference ISO 19115 North American Profile http://www.nrcan.gc.ca/earth-sciences/geomatics/canadas-spatial-data-infrastructure/standards-policies/8912					
Individual Dataset Names					
Dataset Name			Description		
Cemeteries			Cemetery name and location		
Recreation Facilities			Recreation facilities		
Public Washrooms			Public washrooms		
Public Drinking Fountains			Public drinking fountains		
Parking Lots and Garages			Parking locations		

Open Aggregate Dataset Definition: Public Facilities: Parks and Recreation

Dataset Name: Public Facilities: Parks and Recreation

Revision History					
Date (YYYY/MM/DD per ISO 8601)	Version	Updates	Contact		
2018-05-07	0.1	Initial Draft	info@opennorth.ca		
Dataset Description					
Short Version:	This definition represents an aggregate dataset of multiple datasets. The individual dataset definitions will be found in future individual dataset definitions.				
Complete Description:					
Dataset Catgeory:	Parks and Recreation				
Related Standards					
Not applicable as this is an aggregate dataset					
Dataset Tags/Keywords (for publishing)					
Dataset Tags	Parks and Recreation				
Dataset Formats					
Original:	ESRI shp file, Geodatabase				
Published formats:	Shp, Json, GeoJson, GML, KML				
Dataset Metadata					
Geospatial Data - Reference ISO 19115 North American Profile					

http://www.nrcan.gc.ca/earth-sciences/geomatics/canadas-spatial-data-infrastructure/standards-policies/8912					
Individual Dataset Names					
Dataset Name			Description		
Parks			Parks		
Landmarks and Cultural Spaces			Landmarks and cultural spaces		
Public Beaches			Public Beaches		
Trails			Trails - single purpose, multi-purpose.		
Bike Routes			Bike routes		
Bike Parking			Bike Parking		
Tennis Courts			Tennis courts		

Road Construction

Open Dataset Definition - Road Construction					
Dataset Name: Road Construction					
Revision History					
Date (YYYY/MM/DD per ISO 8601)	Version	Updates	Contact		
2018-05-07	0.1	Initial Draft	info@opennorth.ca		
Dataset Description					
Short Version:	This dataset provides information related to road closures or traffic interruptions.				
Complete Description:	The road construction dataset provides detailed information on construction projects, lane closures, event traffic impacts. The information includes location of road closure, dates for road closures, type of construction or event causing the closure, the projected date of completion and municipal person to contact for further information.				
Dataset Category:	Transportation				
Related Standards					

Open 511	http://www.open511.org/documentation/1.0/				
Road Network Dataset Definition	MISA Ontario, OpenNorth				
Dataset Tags/Keywords (for publishing)					
Dataset Tags	Transportation, Roadworks, road construction, streets, Infrastructure, road closures, traffic closures, bridge construction, road restrictions, construction				
Dataset Formats					
Original:	ESRI Shp file, Geodatabase				
Published formats:	Shp, Json, GeoJson, GML, KML				
Dataset Metadata					
Geospatial Data - Reference ISO 19115 North American Profile http://www.nrcan.gc.ca/earth-sciences/geomatics/canadas-spatial-data-infrastructure/standards-policies/8912					
Data Dictionary of Attributes					
Attribute Name	Field Name		Description		
Project Name	PROJECT_NAME		Capital project or event name		
Project Identification Number	PROJECT_ID		Capital project or event identification number		
Project Type	PROJECT_TYPE		Roadway reconstruction or pavement resurfacing, bridges,culverts and other restrictions (including events)		
Project Description	PROJECT_DESC		Project description		
Project Status	PROJECT_STATUS		Project status (pending, active, completed)		
Project Start Date	PROJECT_STARTDATE		Project start date		
Estimated Completion percentage	COMPLETE_PERCENT		Project percentage complete		
Estimated Completion percentage last update	COMPLETE_UPDATE		Update date for percentage complete		
Revised Completion Estimate	COMPLETE_REVISE		Project completion revision update		
Revision date explanation	REVISE_DETAILS		Explanation for project delay or accelerated completion		

Project Completion Date	PROJECT_COMPLETEDATE		Project completion date
Road Closure From Date	ROADCLOSE_FROMDATE		Road closure commencement date
Road Closure To Date	ROADCLOSE_TODATE		Road closure completion date
Road Name	ROAD_NAME		Road name designated by the jurisdiction
Road Segment Identification Number	ROADSEG_ID		Unique identification for each road segment within project area
Construction Area Polygon	CONSTRUCT_POLY		For visual purposes, a polygon defining extent of project area
Detour Routes Name	DETOUR_ROUTE		Detour route(s) name
Detour Routes Road Segments	DETOUR_SEGMENTS		Detour route road segments
Municipal Project Contact	MUNI_CONTACT		Municipal contact who is accountable or supervising the project
Municipal contact phone	MUNI_CONTACT_CELL		Cell phone number of the municipal contact who is accountable or supervising the project
Municipal contact email	MUNI_CONTACT_EMAIL		Email of the municipal contact who is accountable or supervising the project
Work Contractor	WORK_CONTRACTOR		Name of work contractor
Work Contractor Contact Information	CONTRACTOR_INFO		Contact information for work contractor

Open Dataset Definition - Road Network

Dataset Name: Road Network

Revision History					
Date (YYYY/MM/DD per ISO 8601)	Version	Updates	Contact		
2018-05-07	0.1	Initial Draft	info@opennorth.ca		

Dataset Description	
Short Version:	Dataset of all the road centreline segments in the municipality.
Complete Description:	Road centreline segments combine to form a geospatial network of roads (streets) in the municipality. The road segments are defined as being from one intersection to the next or to a point when the road name changes.
Dataset Category:	Transportation, Location

Related Standards	
Ontario Road Network:	https://www.ontario.ca/data/ontario-road-network-road-net-element
National Road Network	http://ftp.maps.canada.ca/pub/nrcan_rncan/vector/geobase_nrn_rrn/doc/GeoBase_ConceptualDataModel_SegmentedView_NRN_2_0_1_EN.pdf
ISO 14825:2011 Intelligent Transport Systems (geo)	https://www.iso.org/standard/54610.html

Dataset Tags/Keywords (for publishing)	
Dataset Tags	transportation, location, street centreline, centerline, road network, roads, Single Line Road Network, SLRN, roadway, infrastructure, road segments, right-of-way, highway, regional road, ally, laneway, route

Dataset Formats	
Original:	ESRI Shp file, Geodatabase
Published formats:	Shp, Json,

	GeoJson, GML, KML				
Dataset Metadata					
Geospatial Data - Reference ISO 19115 North American Profile http://www.nrcan.gc.ca/earth-sciences/geomatics/canadas-spatial-data-infrastructure/standards-policies/8912					
Data Dictionary of Attributes					
Attribute Name	Field Name		Description		
Road Name	ROAD_NAME		Road name designated by the jurisdiction		
Jurisdiction	ROAD_JUR		Name of the jurisdiction responsible for the road		
Road Type	ROAD_TYPE		Classification of road type (e.g. arterial, highway)		
Road Number	ROAD_NUM		Road number assigned by jurisdiction		
Alias Road Name - Last	ALIAS_1		First alias name of road, if available		
Alias Road Name - Next to Last	ALIAS_2		Second alias name of road, if available		
Road Segment Idenitifcation Number	ROADSEG_ID		Unique identification for each road segment		
Number of Lanes	LANES		Total number of lanes for this segment		
Speed Limit	SPEED_KPH		Speed Limit for road segment		
Road Segment Direction	ROAD_DIR		Direction of road segment from start point to end point		
From Address Left	LADD_F		Left address value on left side relative to direction		
To Address Left	LADD_T		High address value on left side relative to direction		
From Address Right	RADD_F		Low address value on right side relative to direction		
To Address Right	RADD_T		High address value on right side relative to direction		
Source	SOURCE		Source where road segment was created		
Source Date	SOURCE_DATE		Date the road segment was created		
Road Condition Rating	CONDITION		Rating for the road segment from road maintenance group.		

Last Maintained	MAINTAIN_DATE		Date that road was last resurfaced or rebuilt.
Public Access	ROAD_ACCESS		Define whether road has been assumed by municipality
From Node Lat & Long	FROM_LAT_LONG		The latitude and longitude of the "from" endpoint node.
To Node Lat & Long	TO_LAT_LONG		The latitude and longitude of the "to" endpoint node.

Transit

Open Aggregate Dataset Definition: Transit					
Dataset Name: Transit					
Revision History					
Date (YYYY/MM/DD per ISO 8601)	Version	Updates	Contact		
2018-05-07	0.1	Initial Draft	info@opennorth.ca		
Dataset Description					
Short Version:	This definition represents an aggregate dataset of multiple datasets. The individual dataset definitions will be found in future individual dataset definitions.				
Complete Description:					
Dataset Catgeory:	Transportation				
Related Standards					
Dataset Tags/Keywords (for publishing)					
Dataset Tags	Transportation, transit, public transit, transit schedule,routes, bus routes, transit routes, subway lines, bus schedules, transit schedules, current bus locations, live transit				
Dataset Formats					

Original:	multiple				
Published formats:	multiple				
Dataset Metadata					
Geospatial Data - Reference ISO 19115 North American Profile http://www.nrcan.gc.ca/earth-sciences/geomatics/canadas-spatial-data-infrastructure/standards-policies/8912					
Dublin Core					
Individual Dataset Names					
Dataset Name			Description		
Transit stop			Point location file of public transit stops		
Transit Route			Transit route identification and location connecting stops		
Transit Schedule			Static file of transit service schedules		
Transit Live (GTFS)			Next arrival time for bus API		

Zoning (GIS)

Open Dataset Definition - Zoning (GIS)					
Dataset Name: Zoning (GIS)					
Revision History					
Date (YYYY/MM/DD per ISO 8601)	Version	Updates	Contact		
2018-05-07	0.1	Initial Draft	info@opennorth.ca		
Dataset Description					
Short Version:	This dataset provides the visual identification of land use zoning areas in the municipality.				
Complete Description:	Municipal land use is defined by the "Official Plan" and one or more "Secondary Plans" together with written land use zoning by-laws. This dataset provides an overview of the land use zoning with linkages to the specific zoning by-laws.				
Dataset Category:	Land Development				

Related Standards					
Municipal Official Plans - Provincial Guide	http://www.mah.gov.on.ca/AssetFactory.aspx?did=11149				
Dataset Tags/Keywords (for publishing)					
Dataset Tags	Land Development, zoning, by-law, land use, committee of adjustment, parcels, planning				
Dataset Formats					
Original:	ESRI Shp file, Geodatabase				
Published formats:	Shp, Json, GeoJson, GML, KML				
Dataset Metadata					
Geospatial Data - Reference ISO 19115 North American Profile http://www.nrcan.gc.ca/earth-sciences/geomatics/canadas-spatial-data-infrastructure/standards-policies/8912					
Data Dictionary of Attributes					
Attribute Name	Field Name		Description		
Land Use Zoning Boundary Polygons	LANDZONE_POLY		The polygons that defines each of the different land use zoning areas		
Land Use Zoning Identification	LANDZONE_ID		The land use zoning designation for each polygon		
Land Use Zone Description	LANDZONE_DESC		Brief description/title of the land use zoning		
Land Use Bylaw	LANDZONE_BYLAW		The land use zoning bylaw associated with each polygon		

Appendix 3: Stakeholder Feedback

The following table is a summary of the comments received from the participating municipalities in the pilot project. The table is followed by comments received from the session's participants at the MISA Ontario Annual conference on Tuesday June 5th.

Section of Report	Feedback	Action Taken
Open Dataset Definition - Road Network	There may be some value in breaking out street name, type & direction	Propose review in developing version 1.0. Some is already included
Open Dataset Definition - Road Network	location as a tag is too broad	No action. Any geospatial should have location as a tag.
Open Dataset Definition - Road Network	Location is a very broad category and relates to all spatial datasets	Agreed - No action. See above.
Open Dataset Definition - Road Network	Ownership - in jurisdictions where there are multiple possible owners of a road (e.g. Regional, county, municipal, private, etc) this attribute is imperative	Jurisdiction attribute is to identify responsibility implying "ownership".
Open Dataset Definition - Road Network	Direction of traffic flow - to indicate 2-way or 1-way streets. 1-way can include a "+" or "-" to indicate which direction the traffic moves (e.g. "1-way +" indicates traffic moves in one direction and in the same direction as the line segment)	Propose review in developing version 1.0.
Open Dataset Definition - Road Network	Roadside Environment - e.g. Urban, Semi-Urban, Rural, etc	Propose review in developing version 1.0.
Open Dataset Definition - Road Network	Surface type - high level surface classification e.g. Paved, Gravel, etc	Propose review in developing version 1.0. The surface type and other elements may wish to be considered as part of a road asset definition.
Open Dataset Defintion - Address Points	Unit Number - the unit number. e.g. for townhouse, condo developments, apartments	Included in version 0.1. As Civic Address Unit
Open Dataset Defintion - Address Points	Additional attribute: Street name - the street name	Amended version 0.1 with this attribute name
Open Dataset Defintion - Address Points	Additional attribute:Civic Number - the civic address number	Included in version 0.1
Open Dataset Defintion - Address Points	Additional attribute: Road segment identifier - the ID of the road segment upon which this address	Propose review in developing version 1.0.

	lies	This may be considered redundant as address range defined in Road Network.
Open Dataset Definition - Road Construction	Additional attribute: Current activities - description of current work activities taking place	There is a project description and status in version 0.1
Open Dataset Definition - Road Construction	Additional attribute: Municipal contact email - municipal project manager email address	Added to version 0.1
Open Dataset Definition - Road Construction	Additional attribute: Municipal contact phone - municipal project manager phone	Added to version 0.1
Open Dataset Definition - Road Construction	Additional attribute: Municipal contact name - municipal project manager name	Included in version 0.1
Open Dataset Definition - Road Construction	Additional attribute: Emergency contact phone	Propose review in developing version 1.0.
Open Dataset Definition - Road Construction	Additional attribute: Emergency contact name	Propose review in developing version 1.0.
Open Dataset Definition - Road Construction	Additional attribute: Traffic impact description - e.g. reduced to one lane	Propose review in developing version 1.0.
Aggregate Open Dataset Definition - Public Facilities - Parks & Recreation	Tennis Courts: maybe should be an attribute of Parks?	Propose review in developing version 1.0. The approach to aggregate datasets should be decided and then individual datasets
Open Dataset Definition - Zoning	Additional attribute: Parcel ID - if parcel-based zoning	Propose review in developing version 1.0.
Open Dataset Definition - Zoning	Would be difficult and would not accurately reflect our zoning if it was rolled up to this level	Propose review in developing version 1.0. Zoning can be very complex after Official and Secondary plans and varies among municipalities
Initial Release Notes for Dataset Definitions - May 8, 2018	**Add:** *"Recreational Program Information Tenders/Procurement"* as new datasets	Propose review in developing version 1.0.
Initial Release Notes for Dataset Definitions - May 8, 2018	Re: Address Points - you can have different use/type in the same building or on the parcel so could be useful at the address level.	Propose review in developing version 1.0. Consider which dataset may be best suited to describe land use (e.g. address points vs. parcel dataset)
Initial Release Notes for Dataset Definitions - May 8, 2018	Re: Zoning. Have discussed this several times with the PSOD group. Would love to see a	Post pilot discussions would benefit issue of

	discussion as to how we might move this forward - but realize is outside the scope of this project - but it just keeps surfacing.	Teranet parcel data usage by municipalities in open data context.
Initial Release Notes for Dataset Definitions - May 8, 2018	Re: Election Results. Just a comment that some additional region's as of this fall will have election results - for elected CHAIR	Added Regional Chair in Candidate list description
Initial Release Notes for Dataset Definitions - May 8, 2018	Re: Road Network. Null attributes would exist as various attributes are maintained at different levels (i.e. municipal vs regional)	Propose review in developing version 1.0.
Initial Release Notes for Dataset Definitions - May 8, 2018	Re: Address Points. - *What is the address standard that should be used? There is a different between Canada Post vs. Bell (for 911) – how could an open data set capture both options?* - *Should the standard show both options (for urban addresses this is usually not an issue, but in the rural areas, this is the difference between RR#2 City/Town and 123 RoadName Having an open standard for this would be ideal*	Propose review in developing version 1.0. There is also a new NG 9-1-1- standard being imposed by CRTC and should be considered in this discussion
Initial Release Notes for Dataset Definitions - May 8, 2018	Re: Road Network: I'm surprised there is that kind of inconsistency but the work we are doing could help them overall if not just for open data. I was hoping we could use data standards from other areas to apply to open data. It may also be the reverse case where data standards developed for open data help define standards for various municipal services that don't have a standard.That said, if you compare what we do for MBN Canada, we don't aggregate statistics into the MBN Can database because it is tough to get good measures that we can all agree to do the same way. At least having basic standards would be useful.	Propose review in developing version 1.0. The pilot project has provided insight and the outcome has led to recommendations taking the work beyond just the top 10 datasets. It needs to be developed in a holistic manner with the necessary stakeholders around the table.
Initial Release Notes for Dataset Definitions - May 8, 2018	Re: Public facilities: I agree public facilities should be a super set. In reviewing the 3 items under that category, was there further definition on the identified datasets (ie police station location, etc.)	Propose review in developing version 1.0. The development of dataset definitions for all of the public facilities related datasets is beyond the scope of the pilot project but should be pursued.
Initial Release Notes for Dataset Definitions - May 8, 2018	Re: Address points: Is civic address one aggregated field for the entire address? Do we want a breakdown? Or is that too complicated.	Propose review in developing version 1.0. Street name has been added to version 0.1
Initial Release Notes for Dataset	Re: Budget - Operating. The annual operating	Propose review in

Definitions - May 8, 2018	budget format is good but I'm not sure if every City would have that kind of detail such as unique processes etc. but could probably squeeze into what is there. I like keeping the number of items to a low number as described.	developing version 1.0. Operating budget can vary dramatically among municipalities. There is value in communicating both with the municipal finance officers association and the provincial government regarding budget detail standardization.
Initial Release Notes for Dataset Definitions - May 8, 2018	Re: Budget - Capital I like the capital budget dataset. Seems to match with my municipality. Funding source could be challenging because those can come from as many as 3 or more sources for large projects.	Propose review in developing version 1.0.
Initial Release Notes for Dataset Definitions - May 8, 2018	Overall I like the breakdown. My only concern is that I think to really get better feedback for you, I would have had to pass each dataset to the respective areas for comments which would take much more time than was allowed here.	Propose review in developing version 1.0. Expanded stakeholder feedback would benefit developing version 1.0
Open Dataset Definition - Election Results	Should reflect election year either as an attribute or in the name of the dataset. Historical election results are useful as well	Added to version 0.1 It should be noted that historical election results can become confusing when ward and /or municipal boundaries change.

Feedback from MISA Conference Presentation Roundtable Discussions

The following is a summary of the key points defined in a roundtable discussion at the MISA Ontario 2018 annual conference. The discussion followed the release of the project's Executive Summary to attendees and a presentation at the session.

Question 1: What's key for adoption of open data standards?

Key points from roundtable discussion:

- Update standards research to ensure we leverage existing stakeholder initiatives;
- Tools such as a DIY will help, also templates for FME etc;
- Need to articulate WHY we need to do this - the quick elevator pitch - be able to answer if Council asks why do we need to do this;
- Demonstrate this meets the users needs and validate through citizen engagement;
- Get it out there
- Continue to engage the right people (OGRA/Ministry/other Open Data groups) to consolidate information;

- Timing has created the appropriate climate for adoption of a toolkit as many are just looking for someone to tell them what to do/use.

Question 2: What are the project's priority recommendations to follow-up next?

Key points from roundtable discussion:

- Develop Use cases to tell the Why story and demonstrate value;
- Watch for direction from Province or other groups (e.g. Stats Canada) and demonstrate partnerships that contributed to success;
- Identify tools, include vendor neutral transformation tools;
- Development of the taxonomy of Open Data to give them a menu of sort to use when classifying open datasets;
- Create a common language/definition for the technical and plain language aspects of Open Data;
- "Get it out there" so that practitioners can start using the framework;
- Not just the standards but language / data dictionaries, etc. (e.g. Public Facilities - hard to define because some Municipalities references community centres where others reference garden sheds.)

Question 3: What services should MISA Ontario provide its membership in the data management and standards area?

Key points from roundtable discussion:

- Neutral vendor agnostic tools;
- Publish best practices (e.g. Policies and Standards);
- Leverage relationship with other organizations, URISA, ARMA
- Create a municipal federated portal (e.g. use API from portals and transform into federated portal and become host for small municipalities);
- Take the lead in the Open Data space
- Build a toolset that would assist in helping / getting municipalities started in this space
- Consolidate information as it's difficult to find information in one place (Information Portal to assist discovery)
- Training "101" in Open Data. There seems to be a lack of knowledge in this space.
- There is a lack of connecting Open Data with Data Management/Information Management
- Answer the question of: What are Municipalities responsibilities in the Open Data space?
- Develop Governance Models. What makes a dataset valuable? Why release a specific dataset? etc.