



SCHOOL OF
INFORMATION
UNIVERSITY OF MICHIGAN

Open Data in Detroit: Transforming Residents' Relationships with their City



Marley Kalt and Lauren Seroka
University of Michigan School of Information

Problem Statement

The Detroit Open Data Portal¹ was created in 2015, aiming to transform the relationship between city residents and their government through the publication of the city's data. Data is collected by individual city departments, but is curated, cleaned, and ingested into the portal by a small team at the city's Department of Innovation and Technology (DoIT).

Our task was to test structured metadata schemas to standardize dataset descriptions and to develop a workflow that will allow the DoIT team to fill-in missing fields in existing and future datasets. Our goal was to make city data more accessible to users and foster further applications of city open data.

Metadata on the Open Data Portal

Number of Views	High	Medium	Low
Number of Datasets	10	37	53
Description			
Column Description			
Data Dictionary			
Tags			

Key: High: >6000 views
Medium: 500 - 6000 views
Low views: < 500 views

Metadata type for >75% of datasets
Metadata type for 25-75% of datasets
Metadata type for < 25% of datasets

Landscape Review

An early step was to investigate how open data is described in other environments. Our landscape review ultimately revealed that there are no metadata standards that are widely adopted across three primary types of open data providers.

Other City Portals

City portals hosted on Socrata and ESRI's ArcGIS also lack complete, structured metadata.

Federal Government

Data.gov uses the Project Open Data Metadata Schema. However, the National Park Service and United States Geological Survey have adopted other standards, indicating a lack of metadata standardization among federal agencies.

Academic Institutions

Stanford's Earthworks² shows that institutions adopt a variety of metadata standards for their geographic data.

Metadata Standards

	Pro	Con
ISO 19139	<ul style="list-style-type: none"> Supported by ESRI Implementation can be automated 	<ul style="list-style-type: none"> Needs to be used with another ISO standard (19115)
ISO 19115	<ul style="list-style-type: none"> Able to crosswalk from portal with existing metadata Can encode in XML 	<ul style="list-style-type: none"> Difficult to understand mandatory vs. optional fields Has to be purchased
Content Standard for Digital Geospatial Metadata	<ul style="list-style-type: none"> Created by federal government Supported by ESRI Well-documented 	<ul style="list-style-type: none"> No recent updates or maintenance Federal gov. encourages use of ISO standards
Project Open Data Schema	<ul style="list-style-type: none"> Created by federal government Used for datasets on Data.gov Well-documented Easy to understand 	<ul style="list-style-type: none"> Not supported by ESRI Not widely adopted Structured in json, ESRI only supports XML
MODS	<ul style="list-style-type: none"> Widely used Able to crosswalk from ISO 19139 Interoperability 	<ul style="list-style-type: none"> Not supported by ESRI

Metadata Recommendation

- The DoIT team should adopt ISO 19139 and ISO 19115
 - ISO 19115 standardizes geographic metadata
 - ISO 19139 is an XML encoding schema for several of the ISO geographic standards (including ISO 19115)
- The DoIT team should create MODS records for all of the datasets on the Open Data Portal
 - Pre-existing XSLT code can transform ISO 19139 to MODS

Interviews

We interviewed 8 people working in 6 departments at the City of Detroit, to understand what data they produce, how metadata is documented, and how they collaborate with DoIT:

- Many employees cite "the public" as being the intended users of their data, and understand the need for transparency in government
- Metadata is not documented internally in individual departments
- Data producers do not want to continuously maintain and update metadata

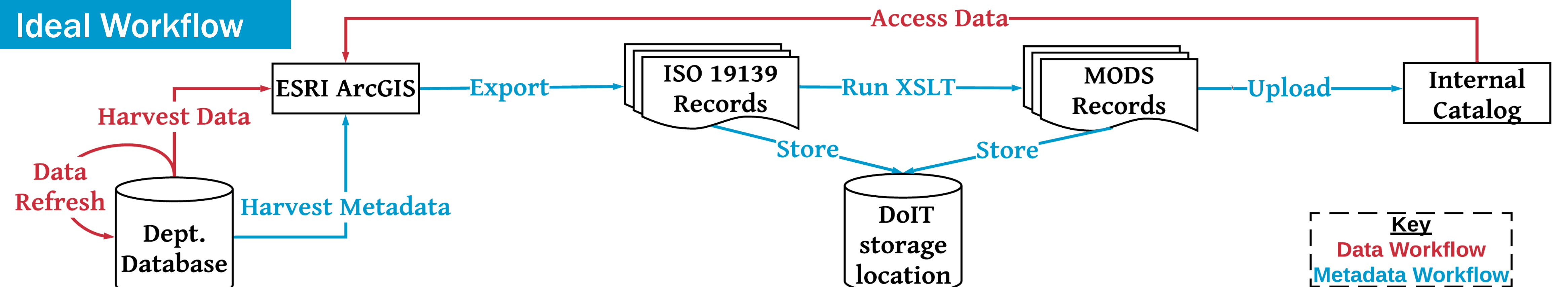
We recommend DoIT work one-on-one with department liaisons to create mutual understandings of metadata standards.

Looking Forward

The Detroit Open Data Portal needs robust metadata to help users understand what each dataset contains and how they can use open data to better understand city operations. The Portal must support a wide range of possible applications, as published datasets are used

by Detroit residents, journalists, researchers, and city employees. While our solution requires a significant initial time investment, we believe a strong foundation of documented metadata is necessary to support the future of the Detroit Open Data Portal.

Ideal Workflow



Resources

- <https://data.detroitmi.gov/>
- <https://earthworks.stanford.edu/>

Acknowledgements

We would like to thank the University of Michigan School of Information and the City of Detroit Department of Innovation and Technology for their support throughout this project.

Contact

Marley: marleyk@umich.edu
Lauren: lseroka@umich.edu