

#### Open Geospatial Consortium: Open Standards for Open Geospatial Data

March 6, 2019 Josh Lieberman Director of Innovation Programs



Copyright © 2018 Geospatial Consortium

## OGC Mission and Open Data

- "To advance the development and use of international standards and supporting services that promote geospatial interoperability."
- Geospatial information systems interoperate by exchanging geospatial information effectively (interfacial, syntactic, geometric, semantic levels) using interface, model, and encoding standards general to data and specific to geodata.
- Standards facilitate bilateral interoperability, but are absolutely necessary for multilateral / unexpected interoperability applications.
- Open data access and open use licensing complement open standards to maximize breadth and variety of data sharing, leading ideally to maximum data value.



#### **Open Data Ecosystems**



# Spatial Data on the Web Best Practices

- Joint product of OGC and W3C collaboration
- Joint document can be found here: <u>https://www.w3.org/TR/sdw-bp/</u>
- "Spatial data, like any other data, should be published on the Web. By this we
  mean more than providing spatial data file downloads or services; for data to be on
  the Web, the resources it describes need to be identified using HTTP URIs, be
  published in such a way that they are indexable by search engines, and be
  connected, or linked, to other resources."
- Examples of best practices:
- Best Practice 1: Use globally unique persistent HTTP URIs for Spatial Things
- Best Practice 2: Make your spatial data indexable by search engines
- Best Practice 3: Link resources together to create the Web of data
- Best Practice 12: Expose spatial data through 'convenience APIs'
- The rest of the best practices provide more detail on specific aspects of publishing spatial data on the Web, such as metadata, geometries, CRS information, versioned data, and so on.





## NYC Open Data Portal

- Open Data Portal is a wonderful initiative that could be even better
- Example of unexpected use: OGC CityGML model built from NYC open data leading to visualization & modeling possibilities for city and community
- Major value of open data is "unexpected use" enabled by open standards, facilitated by consumer feedback on successes and failures.
- Portal API's based on Socrata use OData and other Web-friendly approaches; useability could be improved by following SDW BP recommendations.



# NYC CityGML Model from NYC Open Data



The 3D CityGML model is Open Data! Download: www.gis.bgu.tum.de/en/projects/new-york-city-3d/

- > 1,000,000 buildings
- > 866,000 land lots
- > 149,000 streets
- > 16,000 parks
- > 9,500 water bodies
- > DTM with 1m resolution
- fully-automatically generated from the 2D geodata published in the NYC Open Data Portal
- semantic and geometric transformations
- · all objects have 3D geometry
- rich semantic information (5 - 75 attributes per object resulting from combining different NYC datasets)
- integrated within 1 dataset!
- only 2D and 2.5D data given  $\rightarrow$  generation of 3D geometries
  - volumetric building and tree models
  - all other feature types mapped onto the terrain
  - special treatment of road geometries to include different height levels
- · data heterogeneity
  - different coordinate reference systems
  - different exchange formats (Shapefiles, ESRI GeoDB, Excel etc.)
  - no standardized semantic data model / ontology (each department defines their own data structures)
  - 1:1, 1:n, and n:m mappings required
- · huge data volume
  - large area with > 1 million buildings; big DTM; in total about 4 million objects

## **Additional Thoughts**

- Federated model (physical and/or virtual) can improve currency, enforce standards conformance.
- Storage fit for use
  - Filesystems for datasets
  - RDB's for granular API access
  - Graph stores for linked data
- Linking can balance between integration for usability and segmentation for security
- Privacy controls such as differential privacy budgeting can widen the scope of publishable data
- Explainability computations can increase usability
- Spatial spatial spatial

Concurrent use by city agencies could improve capacity,
 OGC<sup>B</sup>ring more outside ideas into city operations

## **Concluding Thoughts**

- Geography vs GIS
- Ease of use vs There be (MAUP) Dragons
- VGI opportunities and challenges (noise vs bias)
- Unexpected use may be inappropriate use, but may also motivate better data.
- Does it matter and is it possible to quantify the economic benefits of open data in order to support and fund it

