

# National GeoPlatform – Object Editor

ISO Metadata Summit

May 24, 2017



# Topics

- What is the Object Editor?
- Why is it needed?
- What's new? A different approach...
- Use cases
- Demo of Object Editor
- Introduction to the GeoPlatform Profile for ISO-19115-3

# What is the Object Editor?

*A tool for “curating” Portfolios of Online Geospatial Resources...*

- Create and register portfolio objects (representations of resources):
  - Dataset
  - Service
  - Layer
  - Map
  - Organization
  - Contact
  - Concept (e.g., Keyword, Theme)
  - Concept Scheme (e.g., Thesaurus)
- Curate portfolio objects into linked collections (a web) of resources of interest
- Manage online resources as unique, consistent, and unambiguous objects
- Update / correct / augment information about objects. Specifically, elements essential for automated search and online access:
  - Identification
    - Citations
    - Responsible parties
    - Keywords/ topics/ themes
  - Distribution
    - Formats
    - Distributors
    - Digital transfers
    - Online resources
- A work-in-progress... with a roadmap for incremental rollout of new capabilities this year

# Why? So users can find what they need!

## ■ A-16 National Geospatial Data Assets (NGDA)

■ 177 datasets

## ■ Data.gov

■ 122,791 geospatial datasets

## ■ Data.gov CSW

(<https://catalog.data.gov/csw-all?REQUEST=GetCapabilities&SERVICE=CSW>)

■ 122,791 datasets and (and 32 services)

## ■ EPA Environmental Data Gateway (EDG)

■ 3,034 items

## ■ NOAA GEO-IDE UAF ERDDAP

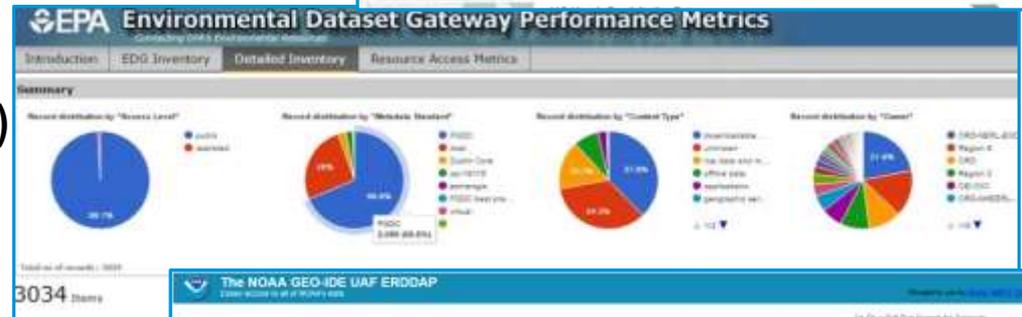
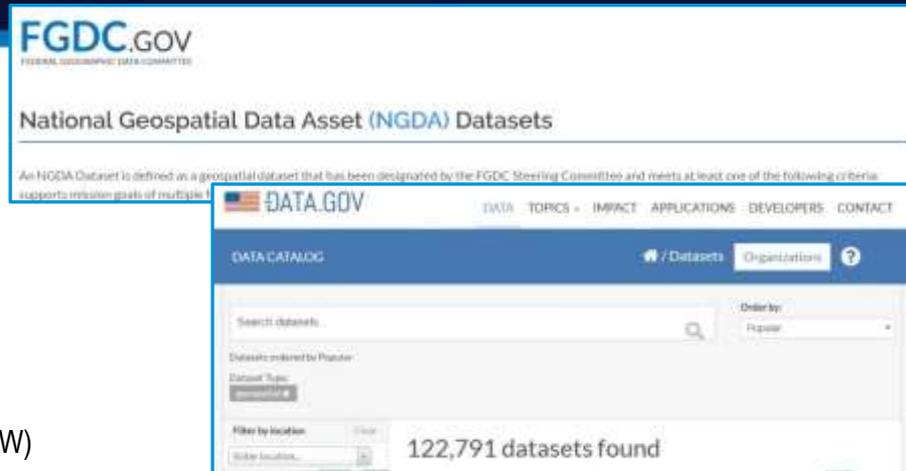
■ 8,985 datasets

## ■ NOAA CSW

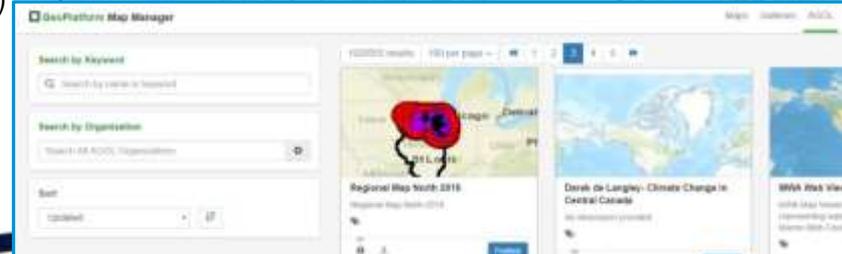
(<https://data.noaa.gov/csw?REQUEST=GetCapabilities&SERVICE=CSW>)

■ 45,654 datasets (and 65 services)

## ■ ArcGIS Online (1,020,570 maps)



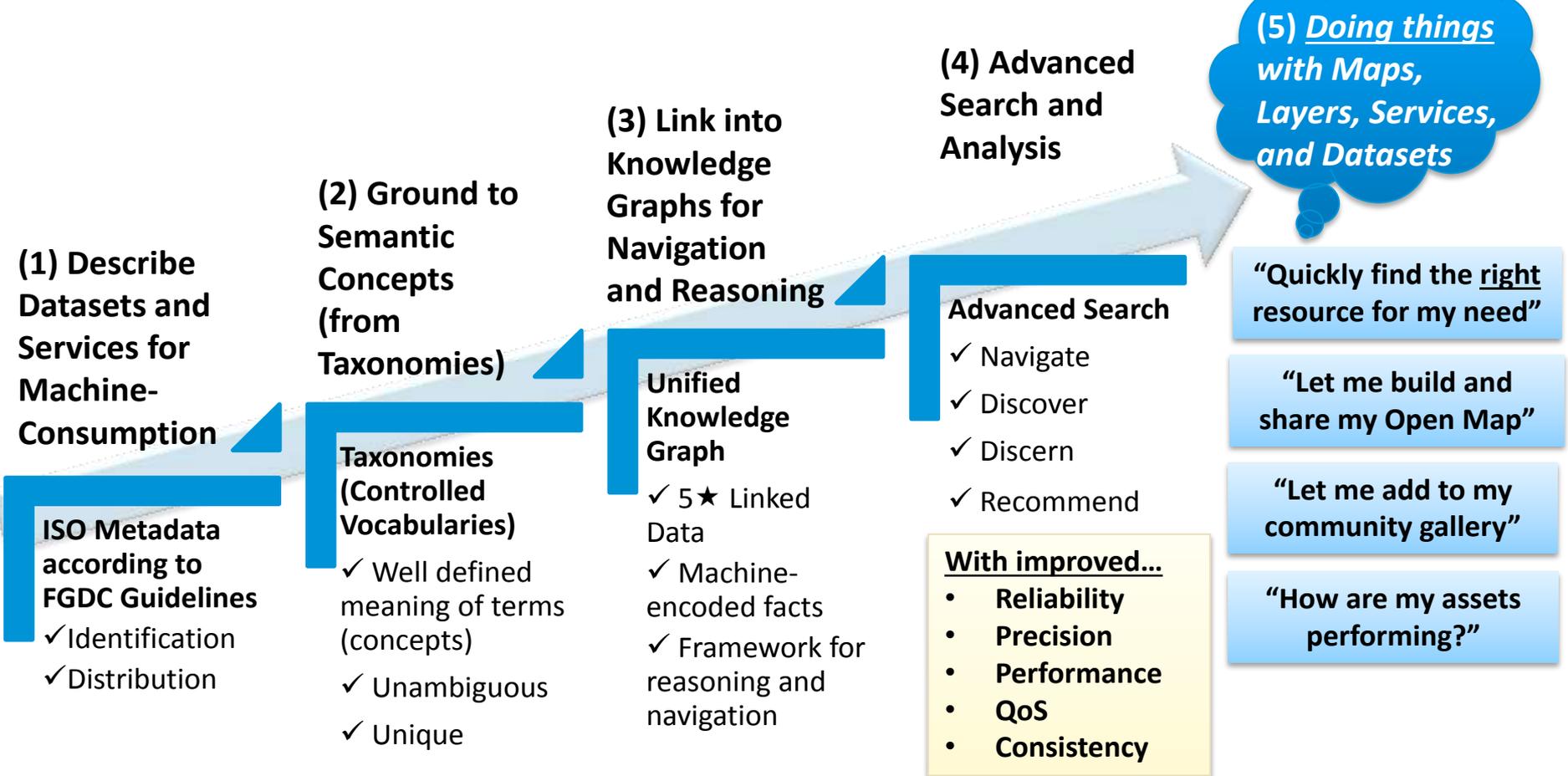
Dataset ID	Title	Date	Source
0001	NOAA List of Active Datasets in the ERDDAP	2018-01-01	NOAA
0002	NOAA Coastal Digital Elevation Model	2018-01-01	NOAA



# Why?

## Connecting user-needs to data and services

*Finding and remembering the good stuff, forgetting the bad stuff...*



# What's the Big Idea? *Linked Open Data (LOD)*

*Metcalf's Law... there's more valuable information when we link things together.*

- Data "on-the-Web" characterizes unlinked, hard to find, document-based data
- Data "in-the-Web" describes well-integrated, linked data providing value-added network effect benefits
- **Good news:** Geodata from US is generally...
  - Downloadable as files "on-the-Web"
  - Freely available under an open license
    - Data can be stored locally
    - Data can be used anyway you wish
    - Data can be shared with anyone
    - Data can be modified as you wish
- **Bad news:** still not good enough to find the right data at the right time...
  - To *automatically* get data *in-the-Web*, we need good identification and distribution metadata
- **Really good news!** *Basic metadata for Datasets and Services opens up the world of Linked Data, Knowledge Graphs and Advanced Search*

# Towards 5★ Open Data

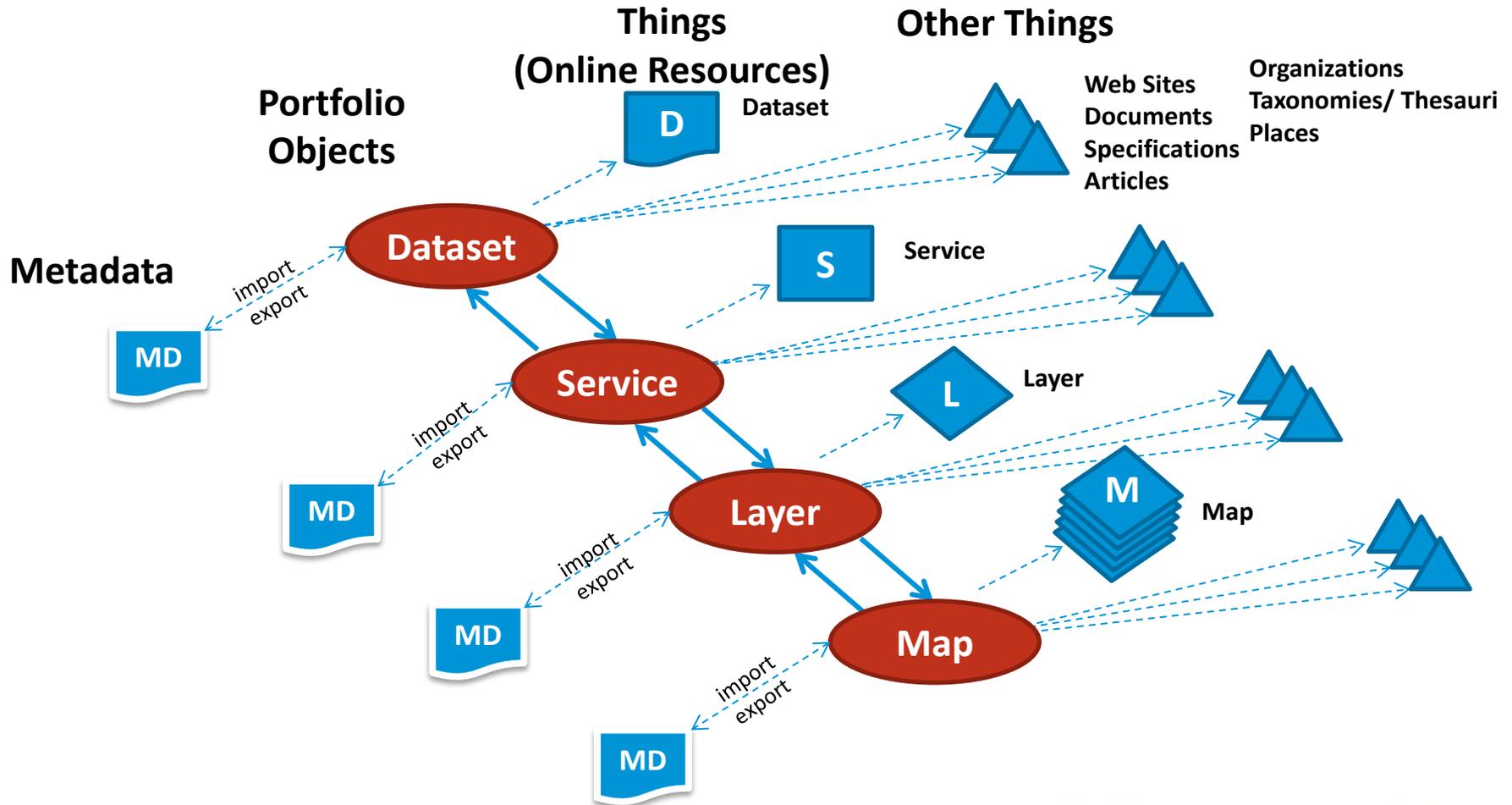
Towards a world of *unambiguous, semantically-grounded linked data* that adds rich context and meaning to shared data.... The last rung in the interoperability ladder.

Level of Openness	Description	Benefits
★	Make your stuff available on the Web (whatever format) under an open license	<u>OK</u> . It's great to have the data accessible on the Web under an open license (such as <a href="#">PDDL</a> , <a href="#">ODC-by</a> or <a href="#">CC0</a> ), however, the data is locked-up in a document. Other than writing a custom scraper, it's hard to get the data out of the document.
★★	Make it available as structured data (e.g., Excel instead of image scan of a table)	<u>Splendid!</u> The data is accessible on the Web in a structured way (that is, machine-readable), however, the data is still locked-up in a document. To get the data out of the document you depend on proprietary software.
★★★	<i>Use non-proprietary formats (e.g., CSV instead of Excel)</i>	<u>Excellent!</u> The data is not only available via the Web but now everyone can use the data easily. On the other hand, it's still data on the Web and not <a href="#">data in the Web</a> .
★★★★	<i>Use URIs to denote things, so that people can point at your stuff</i>	<u>Wonderful!</u> Now it's data in the Web. The (most important) data items have a URI and can be shared on the Web. A native way to represent the data is using RDF, however other formats such as Atom can be converted/mapped, if required.
★★★★★	<i>Link your data to other data to provide context</i>	<u>Brilliant!</u> Now it's data, in the Web <b>linked to</b> other data. Both the consumer and the publisher benefit from the <a href="#">network effect</a> .

[Tim Berners-Lee](#), the inventor of the Web and Linked Data initiator, suggested this [5 star deployment scheme](#) for Open Data

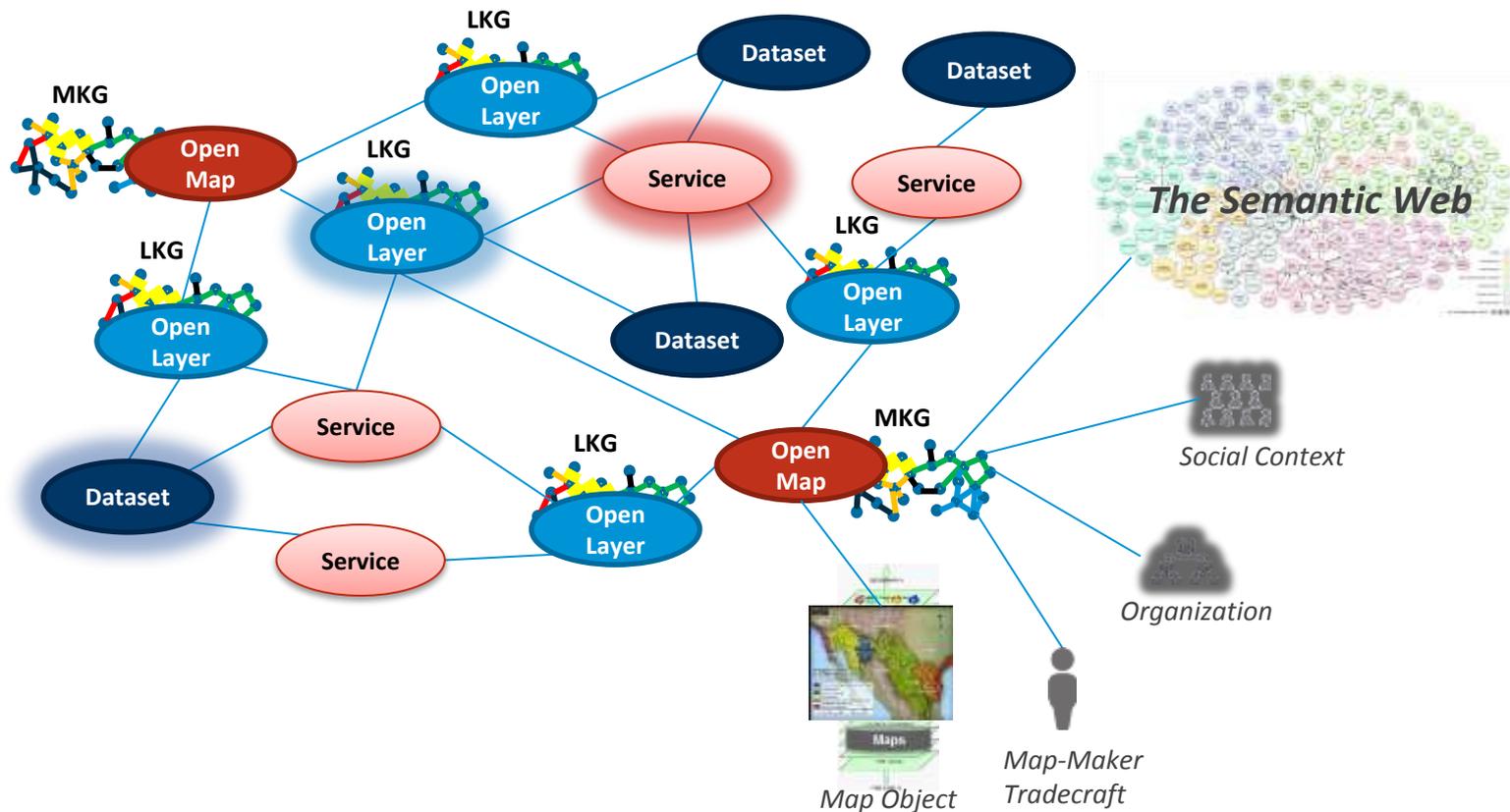
# GeoPlatform.gov: Objects and Internet Things

*Creating Curated Portfolios of Objects...*



# The Unified Knowledge Graph

*A Network of Maps, Layers, Services, Datasets, People, Organizations....*



# ISO-19115 GeoPlatform Profile (Summary)

- Primarily concerned with properly supporting the semantic web and enabling autonomous (machine-machine) access to services
  - mapping relationships between Datasets, Services, Layers, and Maps by...
  - ensuring they have clear and concise identification and location information
  - Support Linked Open Data standards
- Uniform Resource Identifiers (URIs) are used to ensure each resource has a unique and persistent identifier (as required for LOD)
- Open Layer and Open Map identification classes are added at the same level as Dataset and Service identification so they can be linked in metadata for exchange
- Formal semantics are added via Knowledge Graphs that wrap each resource
  - Make use of controlled vocabularies so that all key concepts are unambiguous

Draft spec and schema online here: <https://www.geoplatform.gov/gp-profile>

# ISO-19115 GeoPlatform Profile (Specifics)

Concerned with properly supporting the semantic web and autonomous (machine-machine) access to services...

- Added **identifier** elements to the MD\_Identification class and the CI\_Party class, **ontologyURI** element to MD\_FeatureTypeInfo, and a **uri** element to the CI\_Citation class. (*unambiguously ID things*)
- The element **serviceDocumentation** of type CI\_OnlineResource has been added to the class SV\_ServiceIdentification (*a non-standard, self-describing service... e.g., a swagger document*)
- The elements **format**, **representationTechnique**, and **mimeType** have been added to the CI\_OnlineResource data type to allow better description of online resources (*more machine-consumable*)
- SV\_ServiceIdentification **operatesOn** domain is modified to allow any specified class of MD\_Identification (*Services operate on Maps and Layers too, not just Datasets*)
- Introduces a new element of the MD\_Keywords class, **concept**, which is defined by a new class, MD\_Concept with the following elements and attributes: **conceptIdentifier**, **preferredLabel**, **alternateLabel**, and **description** (*more general-purpose than tags/keywords/themes/places/etc*)
- When **concept** is used, **keyword** should be used to refer to the preferredLabel of a SKOS Concept, MD\_Keywords: **type** should refer to the concept type, and MD\_KeywordClass: **ontology** should be used for concept scheme (*so we can link keywords to grounded SKOS concepts and ontologies*)
- To enable better semantic search, the attribute **type** should use the appropriate code in the KeywordTypeCodes code list and thus, five new KeywordTypeCodes have been added: **audience**, **subject**, **community**, **function**, and **domain** (*more ways to link to grounded concepts for “tradecraft”*)



# Questions?

# Thank You!

For more information, click [here!](#)

Contact Info

