Schema.org Dataset Issues

STEPHEN M. RICHARD

ESIP SUMMER 2018, TUCSON ARIZONA

Workflow

CINERGI

IEDA

Harvest metadata

Resource registration

Enhance

Database

ISO19139 xml

ISO19139 xml

Index/Search UI

presentation

HTML + Schema.org/ JSON-LD

Search UI
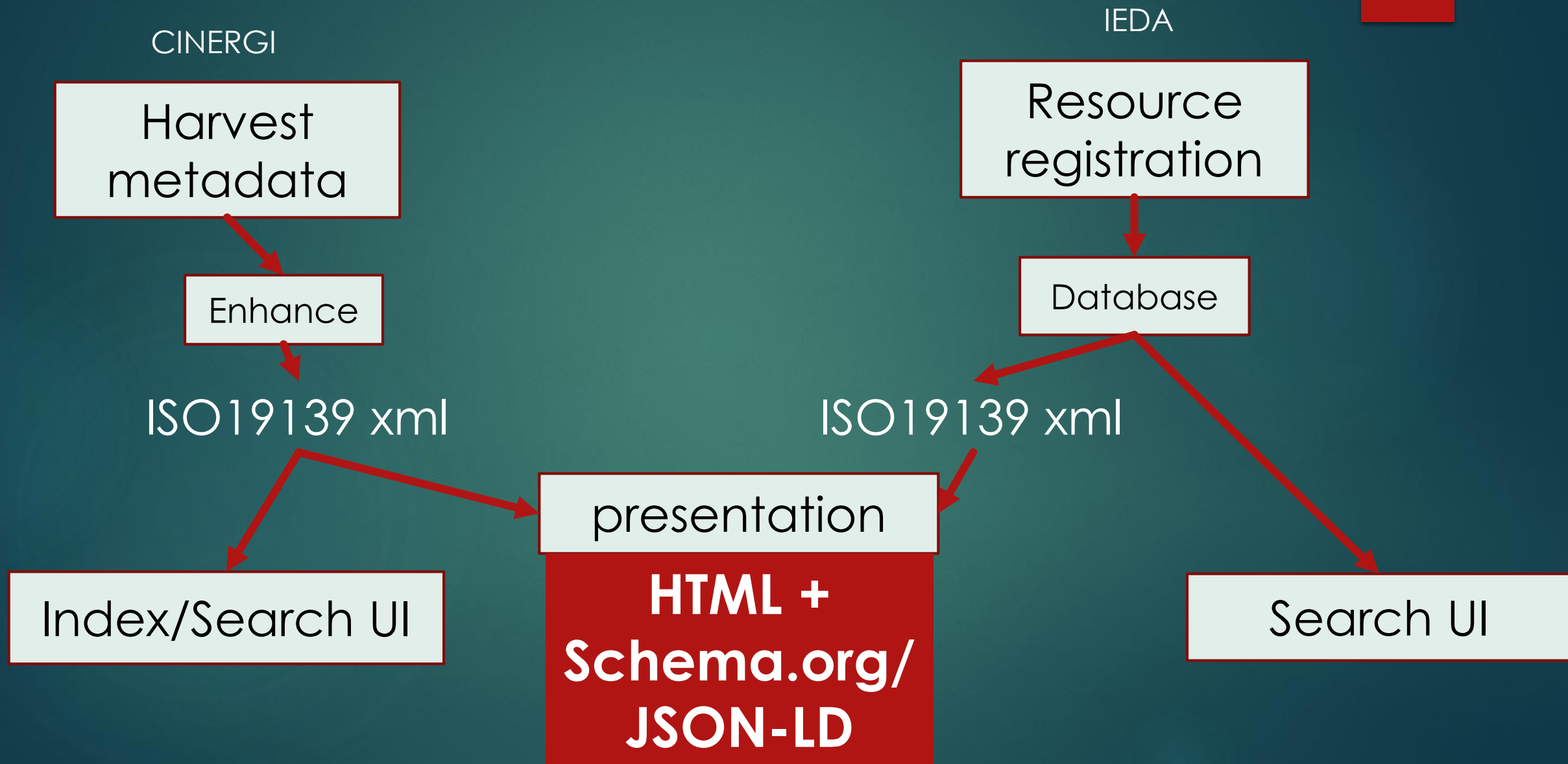
Multiple distributions for resource:

Resource distribution information
Digital Transfer Options

name   **ScienceBase Item Summary Page**
URL: https://www.sciencebase.gov/catalog/item/5032ab9de4b0d64661a77224
protocol   WWW:LINK-1.0-http--link
link function   information
Description   Link to the ScienceBase Item Summary
target application profile   Web Browser

name   **KML Service**
URL: https://www.sciencebase.gov/catalogMaps/mapping/ows/5032ab9fde4b0d64661a77224?mode=download&request=kml&service=wms&layers=WYSagegrouse_currentdistribution
protocol   WWW:LINK-1.0-http--link
link function   file download
Description   KML Download
target application profile   KML client

name   **ScienceBase WMS Service**
URL: https://www.sciencebase.gov/catalogMaps/mapping/ows/5032ab9de4b0d64661a77224?service=wms&request=getcapabilities&version=1.3.0
protocol   OGC:WMS
link function   service description
Description   OGC Service Capabilities URL
target application profile   OWS Client

name   **ScienceBase WFS Service**
URL: https://www.sciencebase.gov/catalogMaps/mapping/ows/5032ab9de4b0d64661a77224?service=wfs&request=getcapabilities&version=1.0.0
protocol   OGC:WFS
link function   service description
Description   OGC Service Capabilities URL
target application profile   OWS Client

name   **Download Attached Files**
URL: https://www.sciencebase.gov/catalog/file/get/5032ab9de4b0d64661a77224
protocol   WWW:LINK-1.0-http--link
link function   file download
Description   Download Link
target application profile   Web Browser
File format

Schema.org encoding

"**distribution**":
  "**@id**": "…",
  "**@type**": "DataDownload",
  "**additionalType**": "dcat:Distribution",
  "**dcat:accessURL**": "https://www.sciencebase.gov/catalog/…",
  "**url**": "https://www.sciencebase.gov/catalog/…",
"**name**": "**ScienceBase Item Summary Page**",

"description": "**Description**: 'ScienceBase Item Summary page' | **Service Protocol**: 'WWW:LINK-1.0-http--link' | **Application Profile**: 'Web Browser' | **Link Function**: 'information'"

"@id": "…",
"@type": "DataDownload", "additionalType": "dcat:Distribution",
"**dcat:downloadURL**": "https://www.sciencebase.gov/catalogMaps/m…",
"url": "https://www.sciencebase.gov/catalogMaps/m…",
  "**name**": "**KML Service**",
  "description": "**Description**: 'KML Download' | **Service Protocol**: 'WWW:LINK-1.0-http--link' | **Application Profile**: 'KML client' | **Link Function**: 'file download'"

"@id": "…",
"@type": "DataDownload", "additionalType": "dcat:Distribution",
"dcat:accessURL": "https://www.sciencebase.gov/catalogMaps/m…",
"url": "https://www.sciencebase.gov/catalogMaps/m…",
"**name**": "**ScienceBase WMS Service**",

"description": "**Description**: 'OGC Service Capabilities URL' | **Service Protocol**: 'OGC:WMS' | **Application Profile**: 'OWS Client' | **Link Function**: 'service description'"

"@id": "…",
"@type": "DataDownload", "additionalType": "dcat:Distribution",
"dcat:accessURL": "https://www.sciencebase.gov/catalogMaps/m…",
"url": "https://www.sciencebase.gov/catalogMaps/m…",
  "**name**": "**ScienceBase WFS Service**",

  "description": "**Description**: 'OGC Service Capabilities URL' | **Service Protocol**: 'OGC:WFS' | **Application Profile**: 'OWS Client' | **Link Function**: 'service description'"

"@id": "…",
"@type": "DataDownload", "additionalType": "dcat:Distribution",
  "**dcat:downloadURL**": "https://www.sciencebase.gov/catalog/…",
"url": "https://www.sciencebase.gov/catalog/…",
  "**name**": "**Download Attached Files**",

  "description": "**Description**: 'Download Link' | **Service Protocol**: 'WWW:LINK-1.0-http--link' | **Application Profile**: 'Web Browser' | **Link Function**: 'file download'"
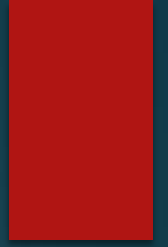
SDO testing tool doesn't like DCAT

Interoperability requires conventions

name   KML Service
URL: https://www.sciencebase.gov/catalogMaps/mapping/ows/5032ab9de4b0d64661a77224?mode=download&request=kml&service=wms&layers=WYSagegrouse_currentdistribution
protocol   CONTROLLED
link function   CONTROLLED
Description   KML Download
target application profile   CONTROLLED
File format   CONTROLLED

- Other considerations:
  - links to zipped packages (container type vs. content type)
  - URL templates
  - Size of target (bytes)
  - Existing href/rel property
  - Accommodate other required parameters

Funding acknowledgement

- Schema.org extension proposal:
  - See https://github.com/schemaorg/schemaorg/issues/383
  - https://webschemas.org/Grant
    - https://webschemas.org/MonetaryGrant
  - https://webschemas.org/Project
    - https://webschemas.org/FundingAgency
    - https://webschemas.org/FundingScheme
    - https://webschemas.org/ResearchProject
- Project is a kind of organization…

Related activities

- DCAT version 2:
  - Editors draft https://w3c.github.io/dxwg/dcat
  - Proposed DCAT-Schema.org alignment: https://github.com/w3c/dxwg/blob/gh-pages/dcat/rdf/schema.ttl
  - Adds better handling for service distributions
- OSGEO 'Cat-Interop'
  - Informal crowd sourcing for link property vocabularies
  - https://github.com/OSGeo/Cat-Interop
- Information model for 'hypermedia affordances' http://usgin.github.io/usginspecs/ContentModelForLinks.htm
- XSLT transforms DataCite to SDO, ISO19139 to SDO: https://github.com/usgin/metadataTransforms