The Open Community Data Exchange (OCDX)

Advancing Data Sharing and Discovery in Open Online Community Science

Link, Germonprez, Goggins, Hemsley, Rand, Squire
OpenSym 2016, Berlin, Germany, Aug 17-19, 2016

Agenda

- Motivation

- Goal

- OCDX

- Current Partners

- Invite

National Science Foundation

- Funded through a grant from the National Science Foundation's Science Across Virtual Institutes Program SAVI/NSF #144920

- Pending new proposal

Motivation

- Online activities increase

- Social scientists collect much digital data

- To date conducted piecemeal

- Behavioral datasets must be described consistently

  - Discoverability

  - Comparability

  - Distribution

Goal

- Support research – valid and coherent comparisons (time and place)

- Metadata standard – describe dataset

- Tools – manage metadata

- Infrastructure – search and find

- Science of Science – changes

METADATA

OCDX Metadata Specification

##OCDX Manifest - (Required/Not Repeatable)
ocdx_manifest:id
      Unique identifier for manifest -- is required; is not repeatable
ocdx_manifest:creator
      Name of person creating manifest -- is required; is not repeatable

##OCDX Dataset - (Required/Not Repeatable)
ocdx_dataset:title
      One sentence title for the dataset -- is required; is not repeatable
ocdx_dataset:abstract
      Summary of the dataset -- is required; is not repeatable
ocdx_dataset:provenance_narrative
      Workflow of collecting, filtering, or cleaning the data -- is not required; is not repeatable

##OCDX Dataset Creator - (Required/Not Repeatable)
ocdx_creator:name
      Person or organization with role in producing the dataset -- is required; is repeatable

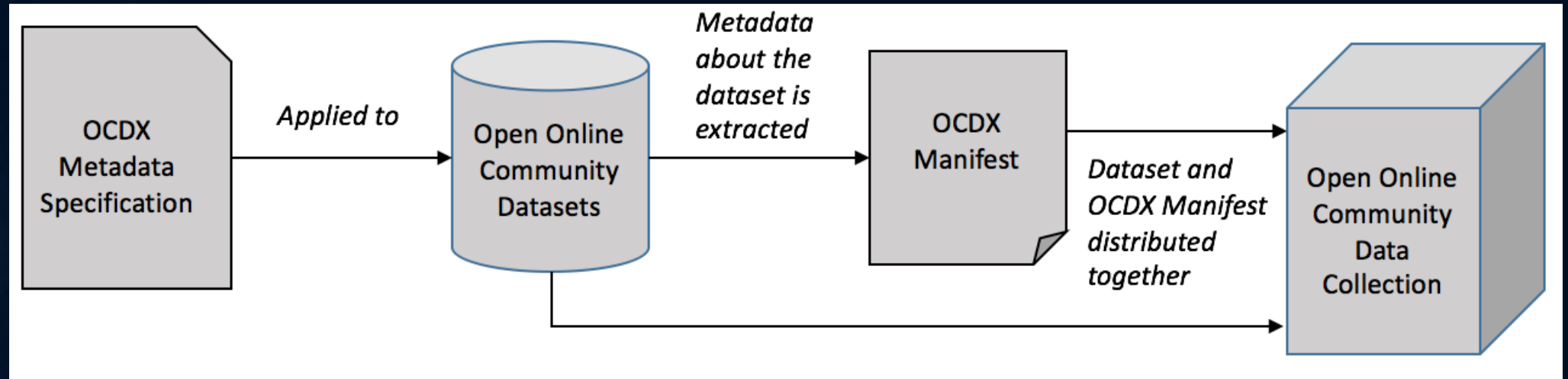##OCDX Dataset Files - (Required/Repeatable)
ocdx_file:name
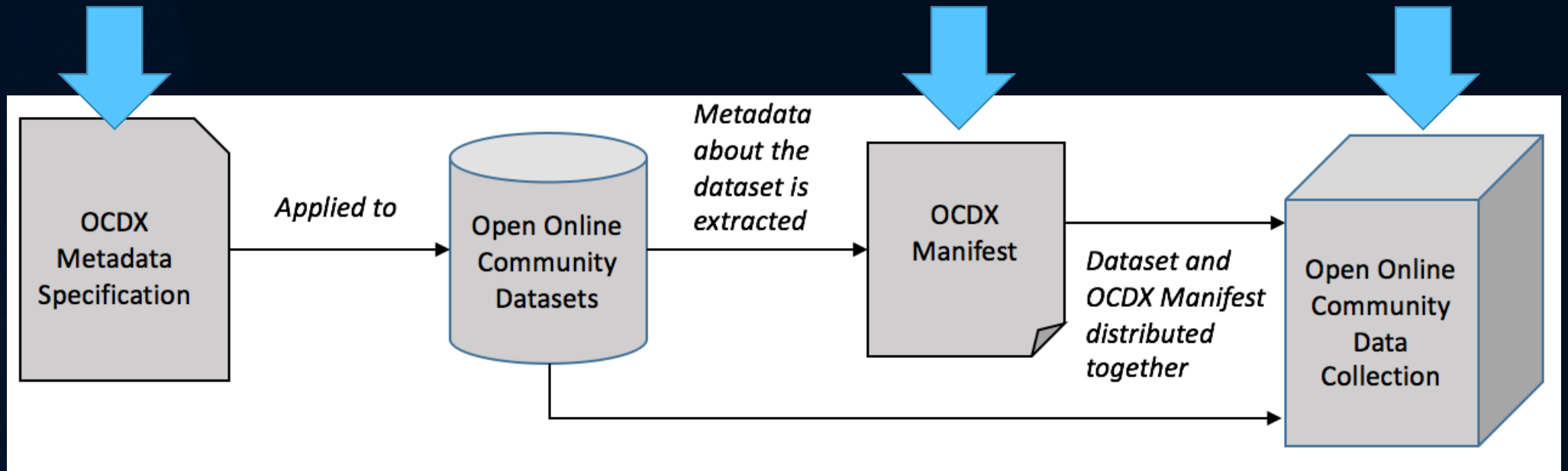   Name of dataset file -- is required; is not repeatable
ocdx_file:permissions
      Notices of rights/obligations that define use of the dataset file -- is not required; is not repeatable

https://github.com/OCDX/OCDX-Specification

OCDX

OCDX

Current Partners

GitHub for sourcing real-time datasets from communities

Wikimedia Foundation for building infrastructure in the production, distribution, and sharing of OCDX manifests

Invitation

- Apply OCDX to your datasets

- Find other datasets

- Share your insights

- Contact Matt Germonprez or Sean Goggins for participating

    - ICIS in Dublin

    - ACM Group in FL
        http://www.ocdx.io/?p=1133
        http://oldwww.acm.org/conferences/group/conferences/group16/

    - ACM CSCW in Spring
        http://cscw.acm.org/2017/

http://OCDX.io