Ontology-Enabled Metadata Schema Generator: The Design Approach

Jian Qin
School Of Information Studies
Syracuse University
Syracuse, NY

Xiaozhong Liu   Miao Chen
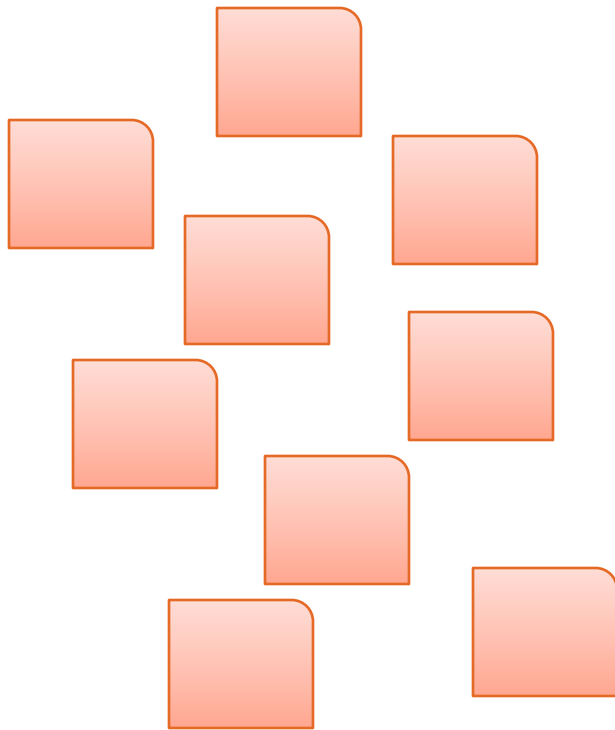School of Informatics and Computing
Indiana University
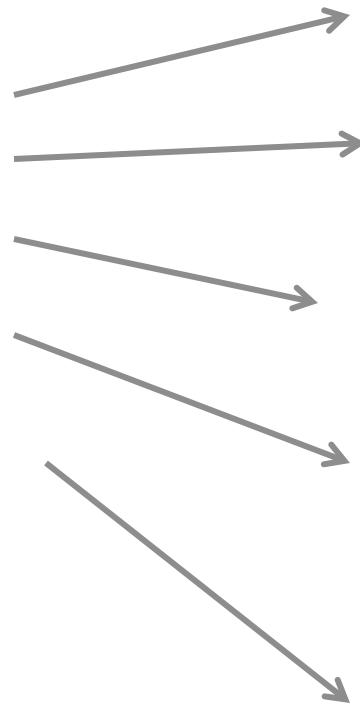Bloomington, IN

An infrastructure perspective for metadata

- Portable
- Customizable
- Extendable
- Reusable
- Easy to use

Portability is the key

Building blocks of metadata

Metadata generation output

Metadata for data citation
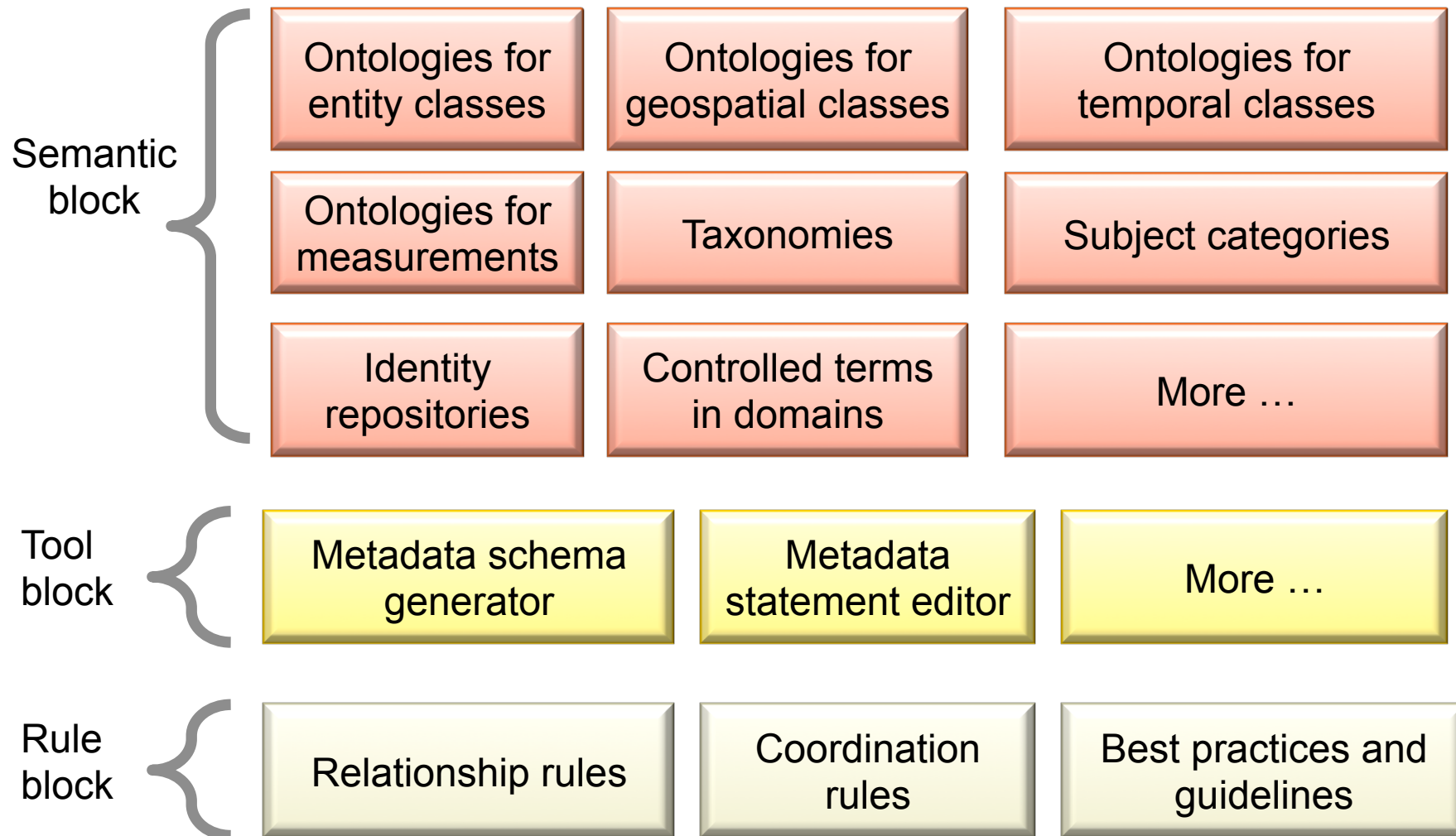
Metadata for data discovery

Metadata for data archiving
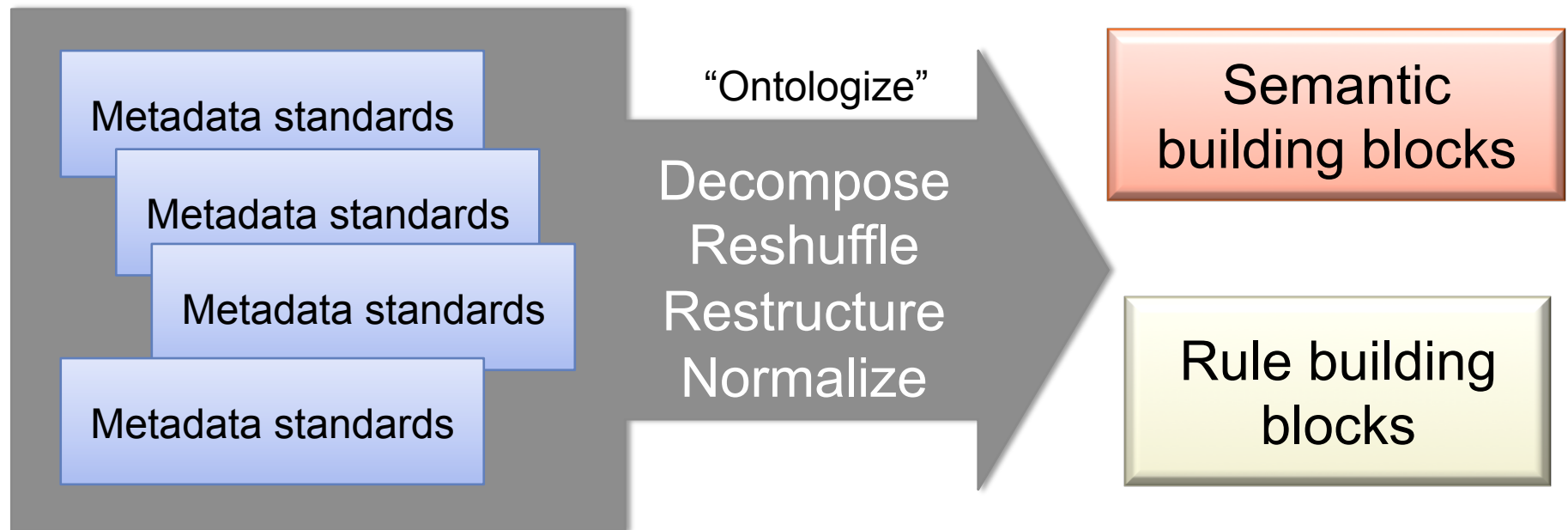
Metadata for data quality

Metadata for data provenance

Metadata for data management

What types of building blocks are there for a metadata infrastructure?

Semantic block

Ontologies for entity classes	Ontologies for geospatial classes	Ontologies for temporal classes
Ontologies for measurements	Taxonomies	Subject categories
Identity repositories	Controlled terms in domains	More …

Tool block

Metadata schema generator	Metadata statement editor	More …

Rule block

Relationship rules	Coordination rules	Best practices and guidelines

Metadata standards: an transition from one-covers-all to portable modules

Metadata standards

Metadata standards

Metadata standards

Metadata standards

"Ontologize"

Decompose
Reshuffle
Restructure
Normalize

Semantic building blocks

Rule building blocks

Envisioning the metadata infrastructure: ontology-enabled metadata schema generator

Backend ontologies

- ORCID
- FOAF
- VIVO
- Dublin Core
- Geographic metadata standards
- Context metadata elements

Ontology Interface

- Sci-Onto
  - Person
    - ID
    - Name
  - Organization
  - Dataset
    - ID
    - Name
    - Date
    - Status
  - Geospatial
    - Bounding coordinates
    - Place
  - Temporal
  - Context
    - Assessment
    - Bio-material
    - Human
    - Measurement

Drag and Select → Metadata schema generator

output → Schema in format

| XML | RDF | Others ? |

Envisioning the metadata infrastructure: ontology-enabled metadata generator
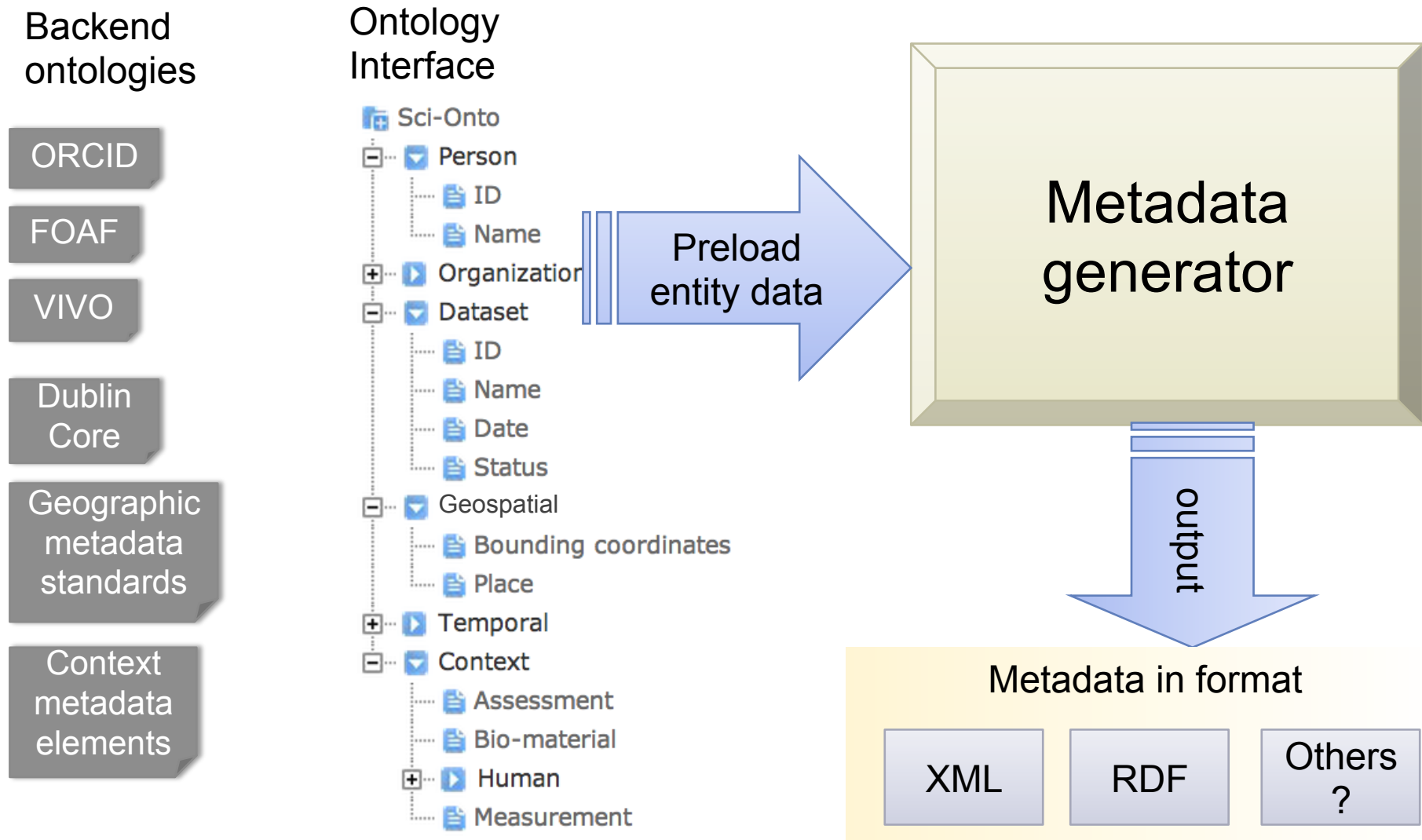
Questions to be addressed

Ontologies

- What ontologies?
- Who will create them and how can community consensus be reached about them?
- How big/small should ontologies be to best meet the schema generation and metadata generation needs?

Rules

- How will the assembling of schemas be regulated?
- Who will determine what rules for schema generation?

Plan for a test

Start from a small scale experiment

- Leveraging existing semantic and technical resources
- Utilizing the findings from our survey to metadata standards for scientific data for ontology development
- Prototyping the metadata schema generator
- Evaluate the result and scale up

Comments?

Thank you!