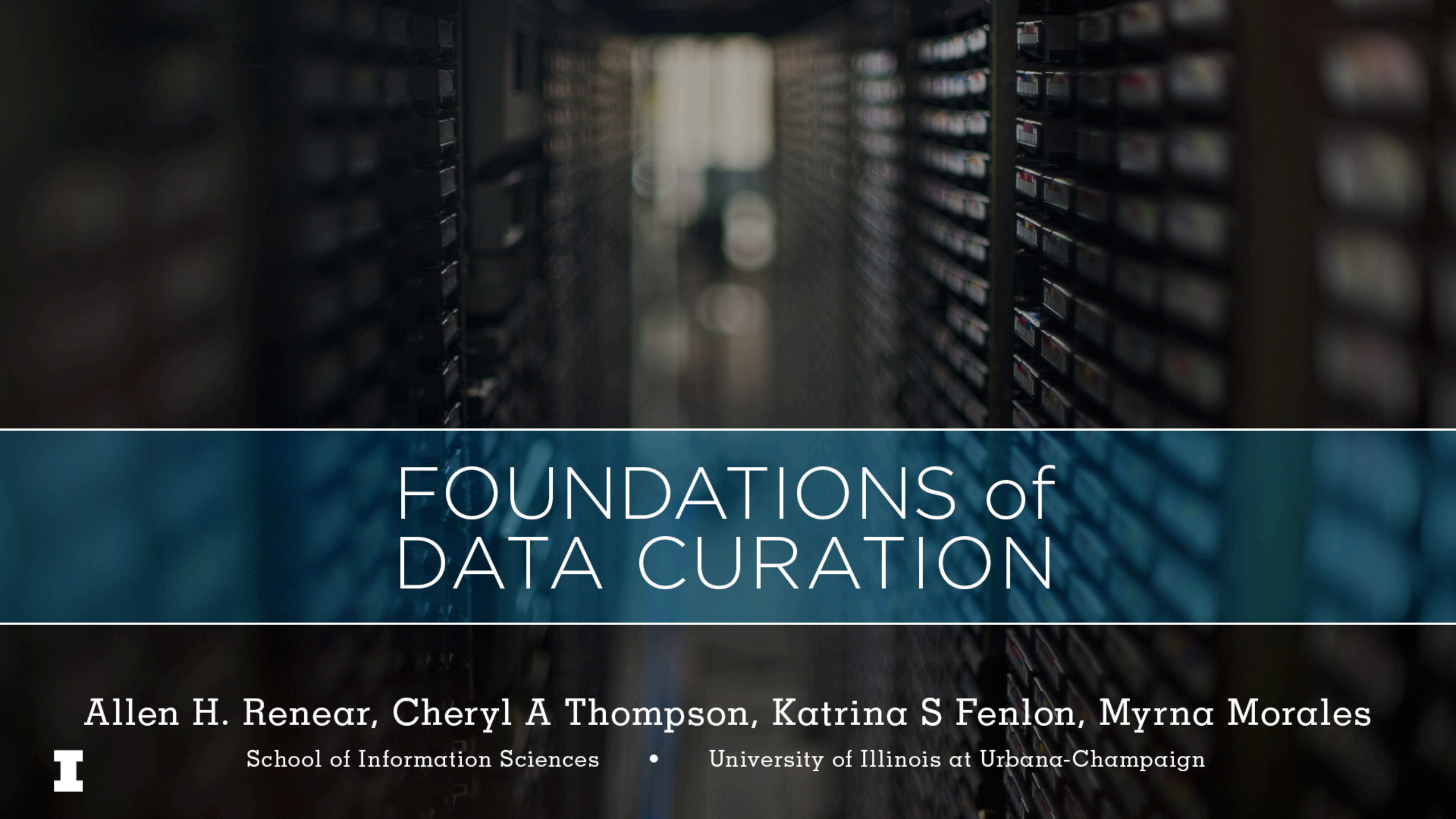
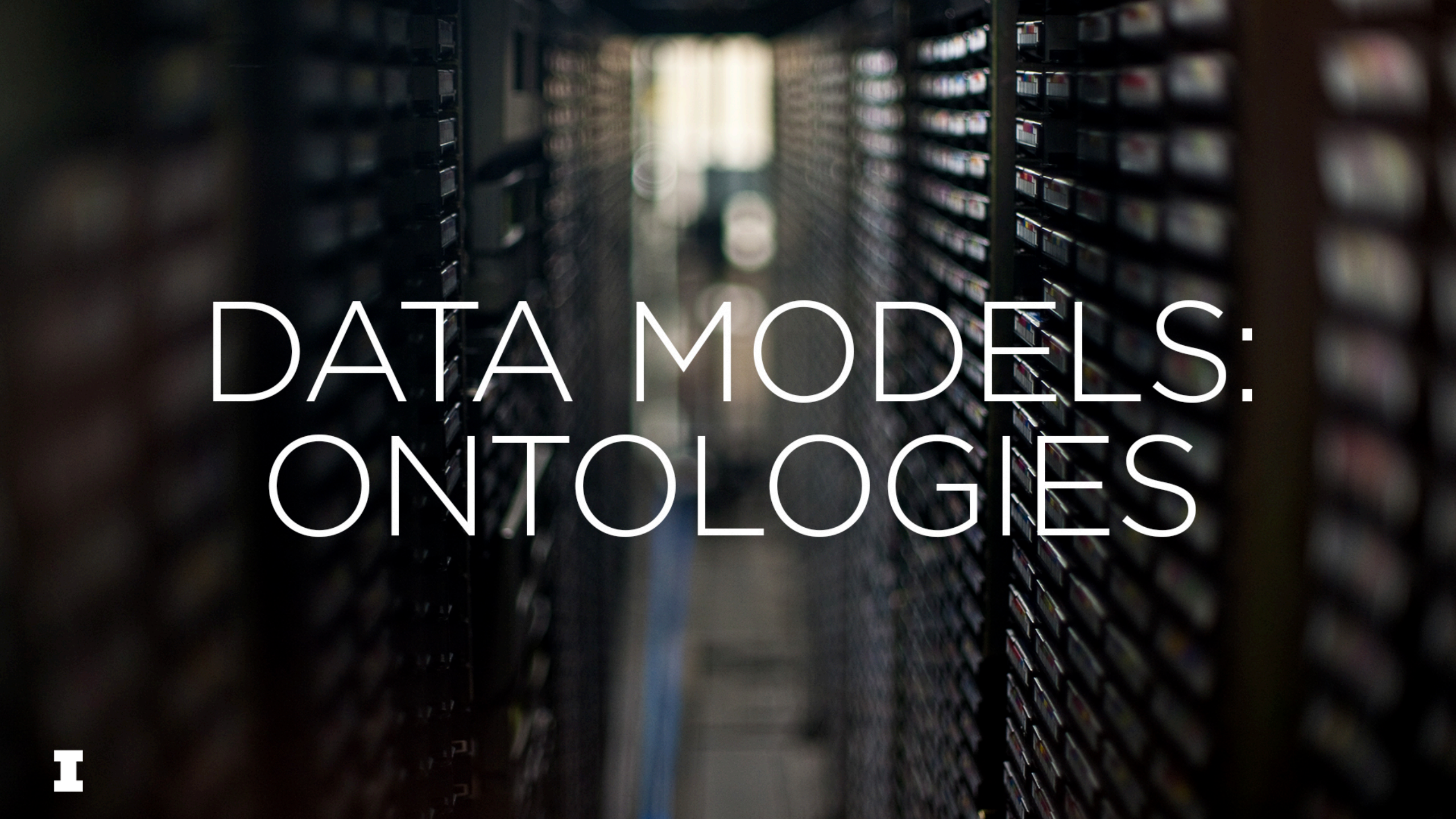
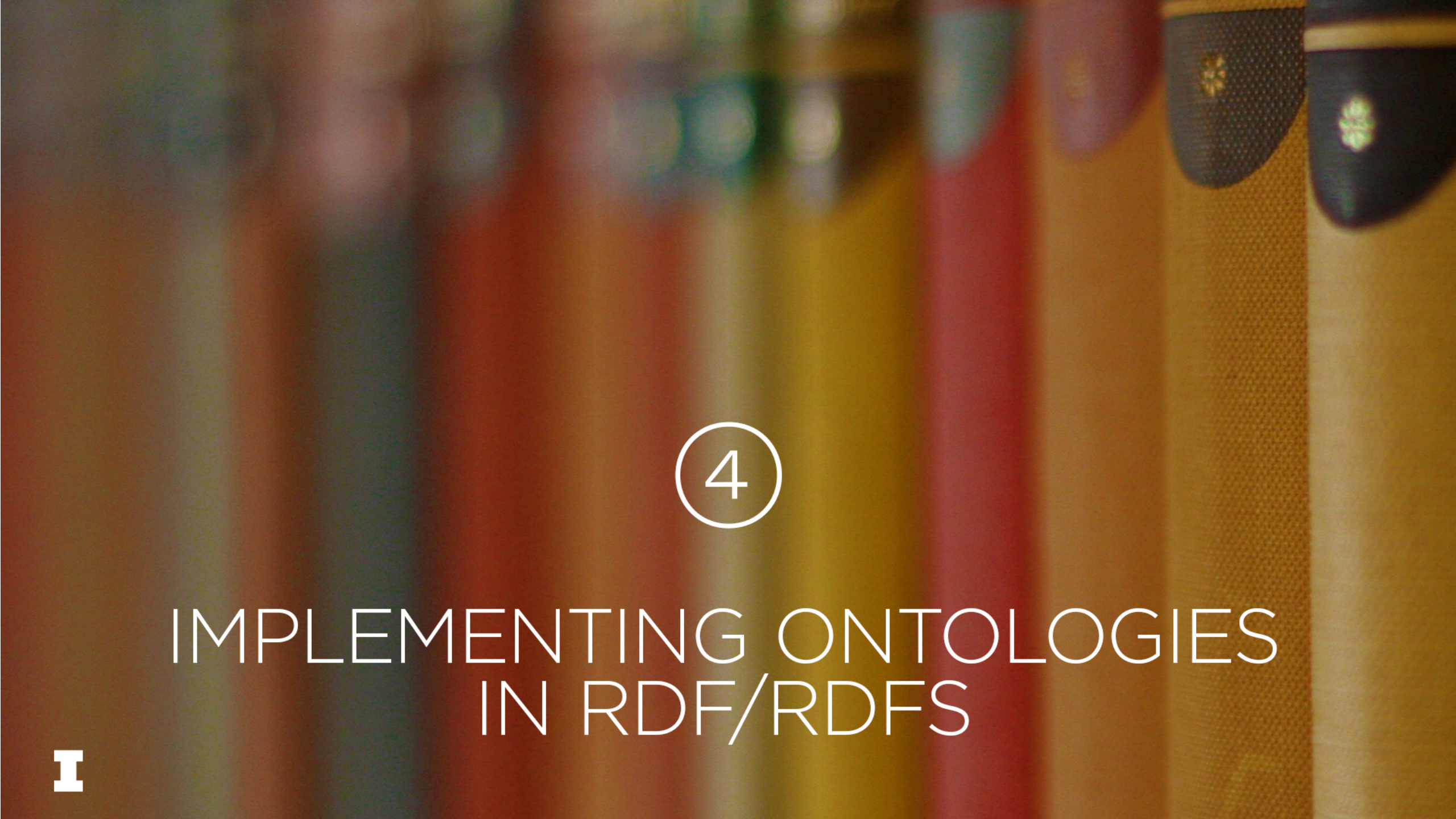


FOUNDATIONS of DATA CURATION

Allen H. Renear, Cheryl A Thompson, Katrina S Fenlon, Myrna Morales

School of Information Sciences • University of Illinois at Urbana-Champaign


DATA MODELS:
ONTOLOGIES


4
IMPLEMENTING ONTOLOGIES
IN RDF/RDFS

Implementing Ontologies In RDF/RDFS

- The RDF graph model for predication
- RDF Schema: for defining a simple ontology
- An example
- Our data model diagram is now completed!

RDF, RDFS, OWL

Today ontologies are often expressed in RDFS or OWL, theses are powerful logic-based languages designed to be read a processed by software, including inferencing systems.

RDF: Resource Description Framework

A simple model for predication; can be serialized in different languages

RDFS: RDF Schema

Basic schema concepts that can be used to define schemas for RDF instances

OWL: Web Ontology Language

General schema concepts for more advanced schemas

Resource Description Framework

- RDF gives a general model for describing things - by saying that some *thing* has a particular *property* with a particular *value*.

- These descriptions are intended to be processed by software and shared across applications and computing environments without loss of meaning.

- RDF can be expressed as a graph or serialized in various ways, including XML.

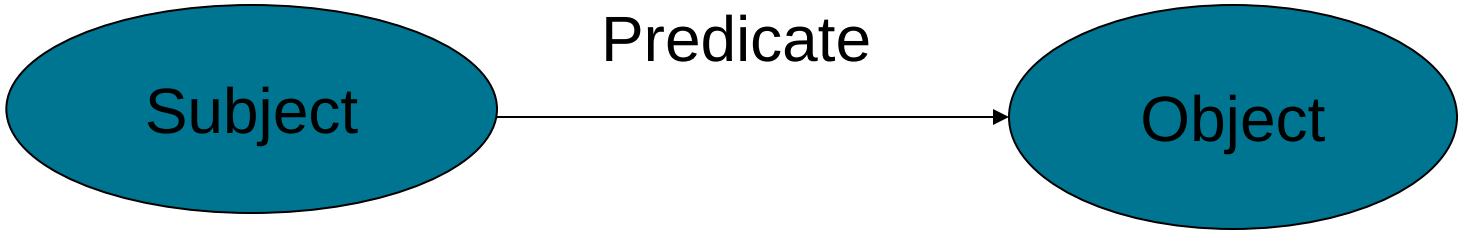
RDF Data Model

The basic abstract structure of a single descriptive assertion in RDF is a triple, an ordered set of these three elements:

- Subject (thing)
- Predicate (property)
- Object (value)

RDF Graph Languague

The representation of a triple in the RDF graph language:



An RDF graph is a collection of such triples

NB: Where ER has distinct constructs for both attributes and relationships, RDF uses a single construct, the predicate/property, for both. After all, under the hood they are both just two-place FOL predications R(x,y),as we have seen.

Example - To Express in RDF

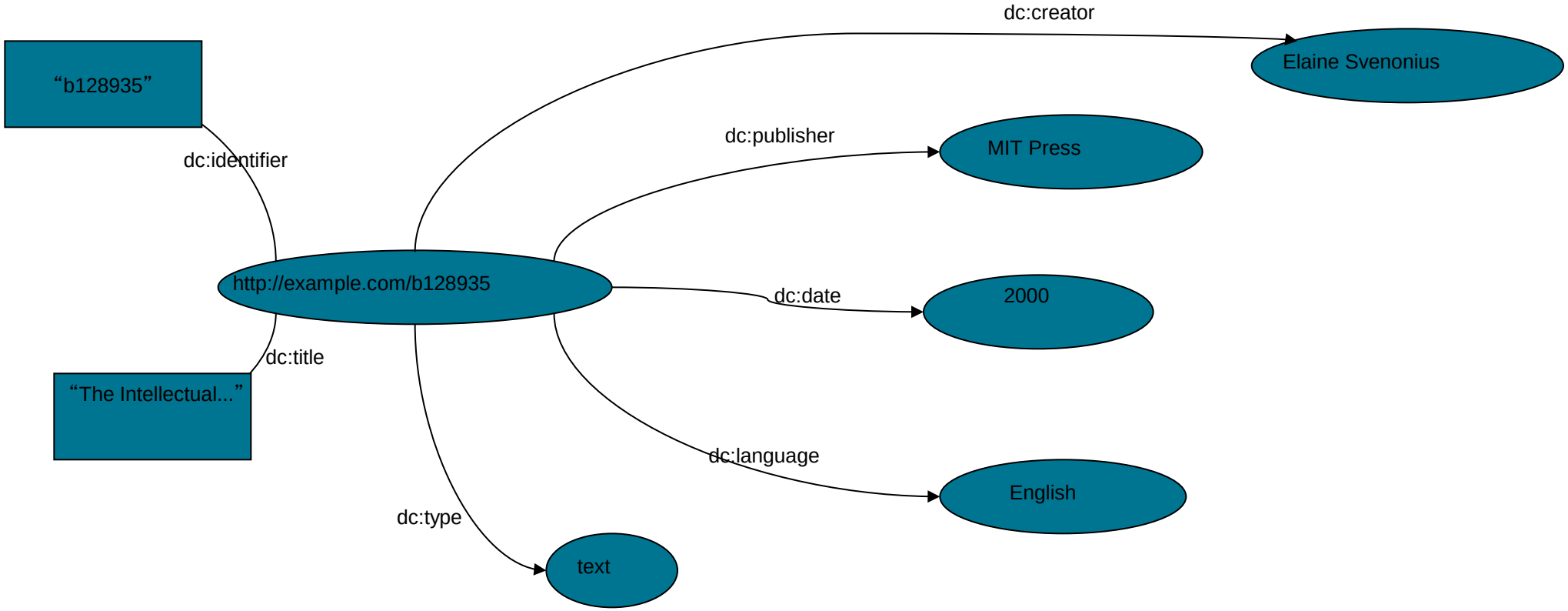
A simple bibliographic description using Dublin Core*:

- `identifier = b128935`
- `title = The Intellectual Foundation of Information Organization`
- `creator = Elaine Svenonius`
- `date = 2000`
- `publisher = MIT Press`
- `language = english`
- `type = text`

- *http://dublincore.org/documents/dces/

Expressing DC in RDF

The description as an RDF graph



NB: This graph represents a set of seven triples.

RDF Description as Triples

<b128935,dc:identifier,”b128935”>
<b128935,dc:title,”The Intellectual...”>
<b128935,dc:creator,Elaine Svenonius>
<b128935,dc:date,2000>
<b128935,dc:publisher,MIT Press>
<b128935,dc:language,english>
<b128935,dc:type,text>

But that picture is pre-FRBR

The RDF graph on the previous slide is, of course, not consistent with FRBR.

To adapt that description to the FRBR conceptual model, we need a system of related classes and properties.

That is, we need a schema, the sort of thing an ER diagram is:

- A schema language that interoperates with RDF
- That would be: RDF Schema

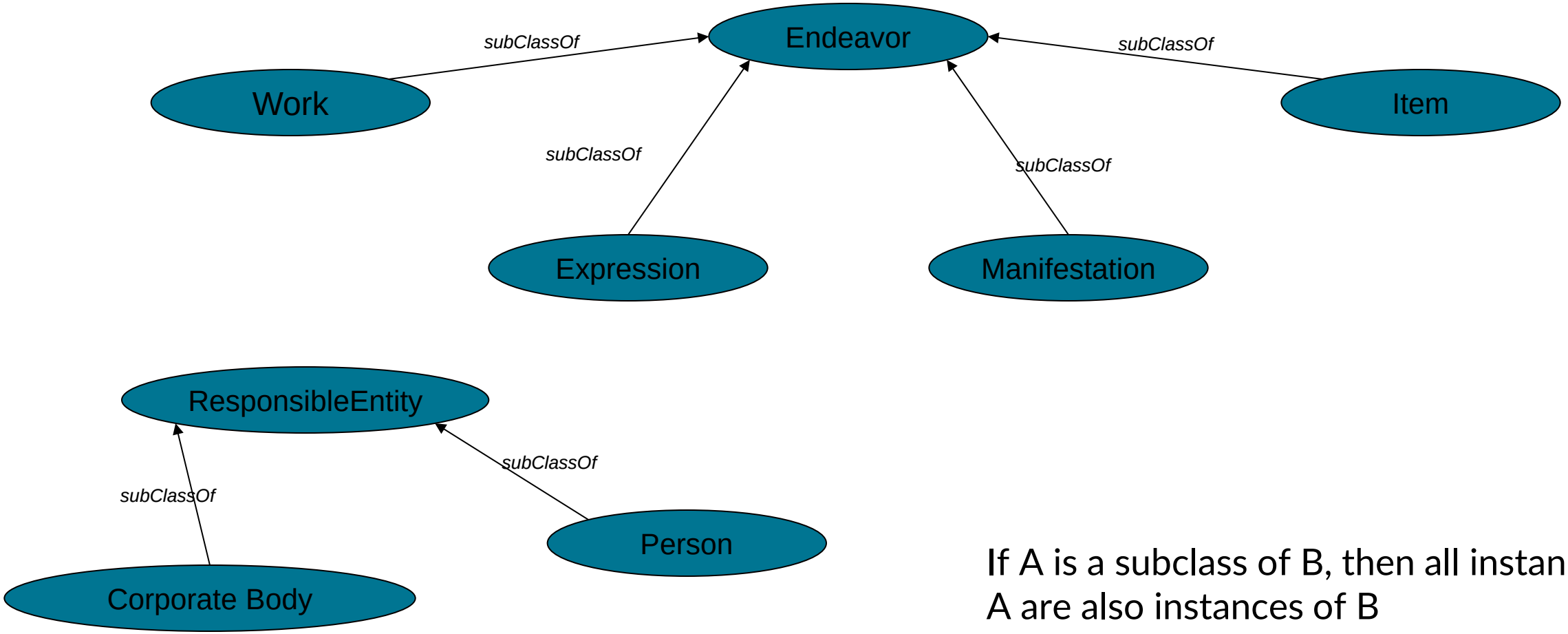
RDF Schema

RDF Schema is a schema language for defining RDF vocabularies:

- The RDF Schema language itself uses RDF triples as its syntax (!)
- It needs some vocabulary for its own. That is, it needs some fundamental properties and classes that it can use to define RDF vocabularies for specific domains. See the next slide.

What follows is largely based on the RDF vocabulary for FRBR published here: http://vocab.org/frbr/core

FRBR in RDFS: Classes

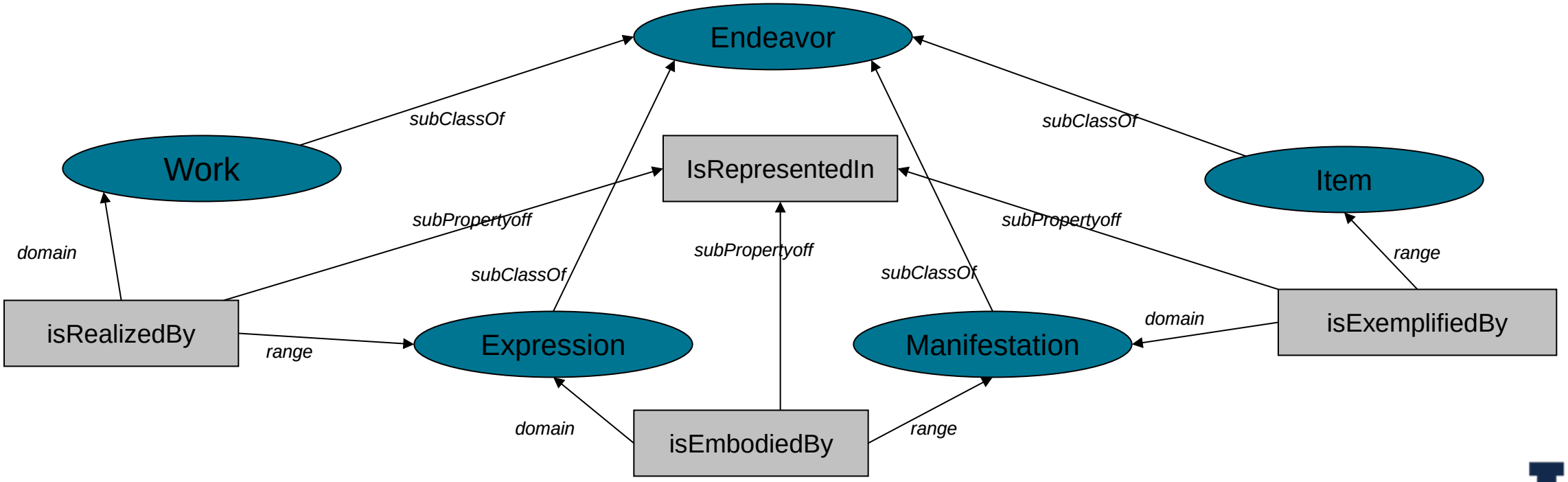


If A is a subclass of B, then all instances of A are also instances of B

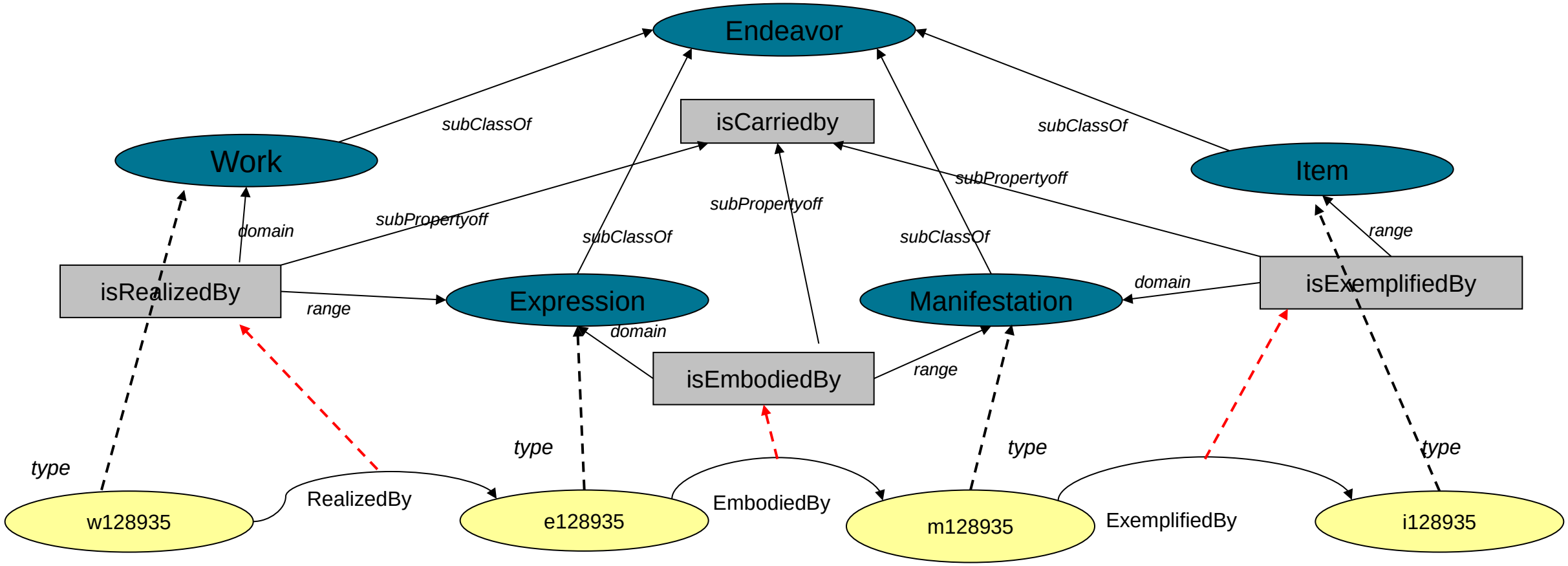
FRBR in RDFS: Properties

Properties have domains and ranges indicated by RDFS core properties domain and range.

Properties can also be arranged in a hierarchy just as classes can, indicated by the RDFS core property subproperty.



Connecting an Instance to the Schema: FRBR Entities

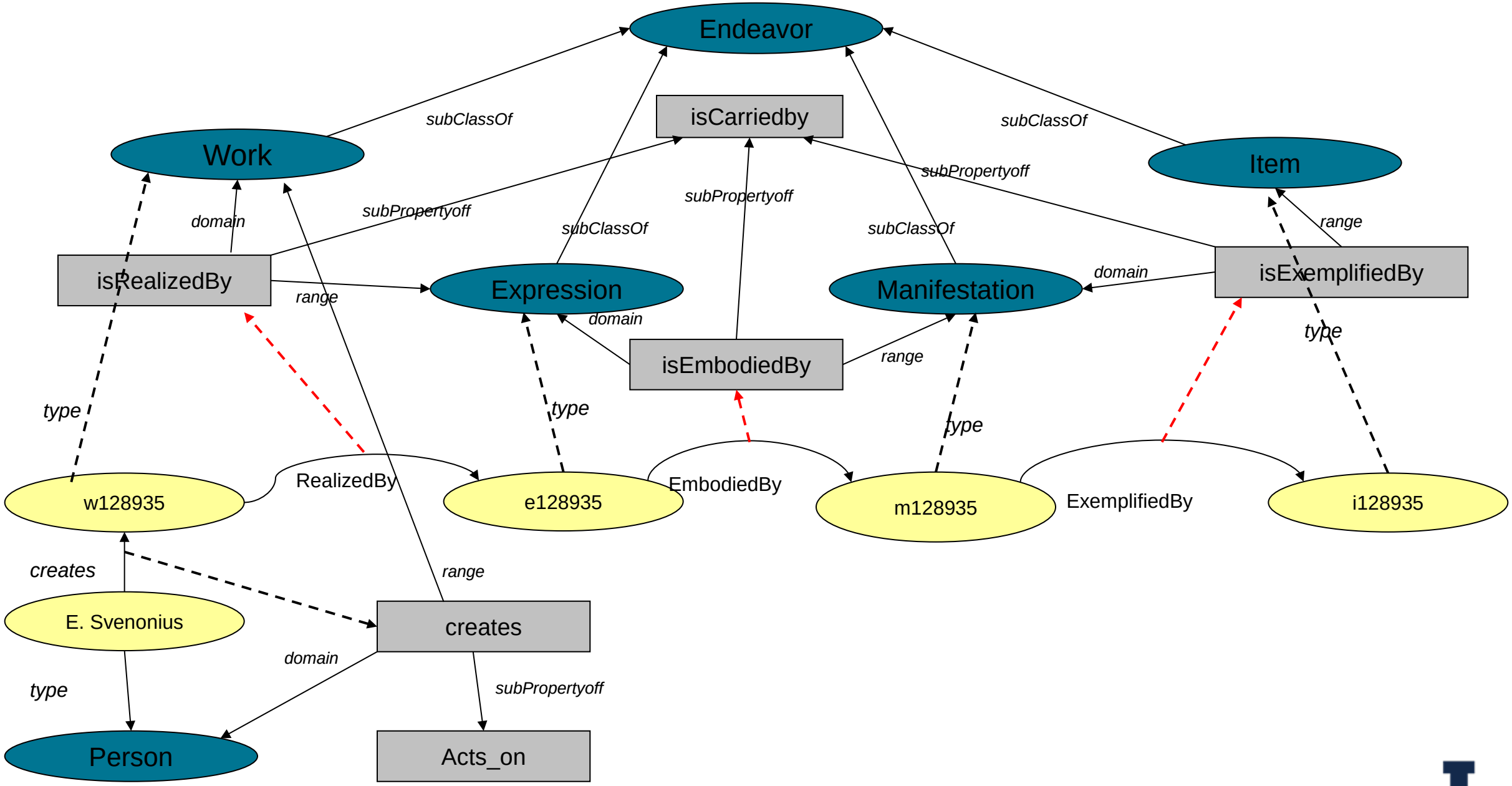


Connecting an Instance to the Schema: FRBR Entities

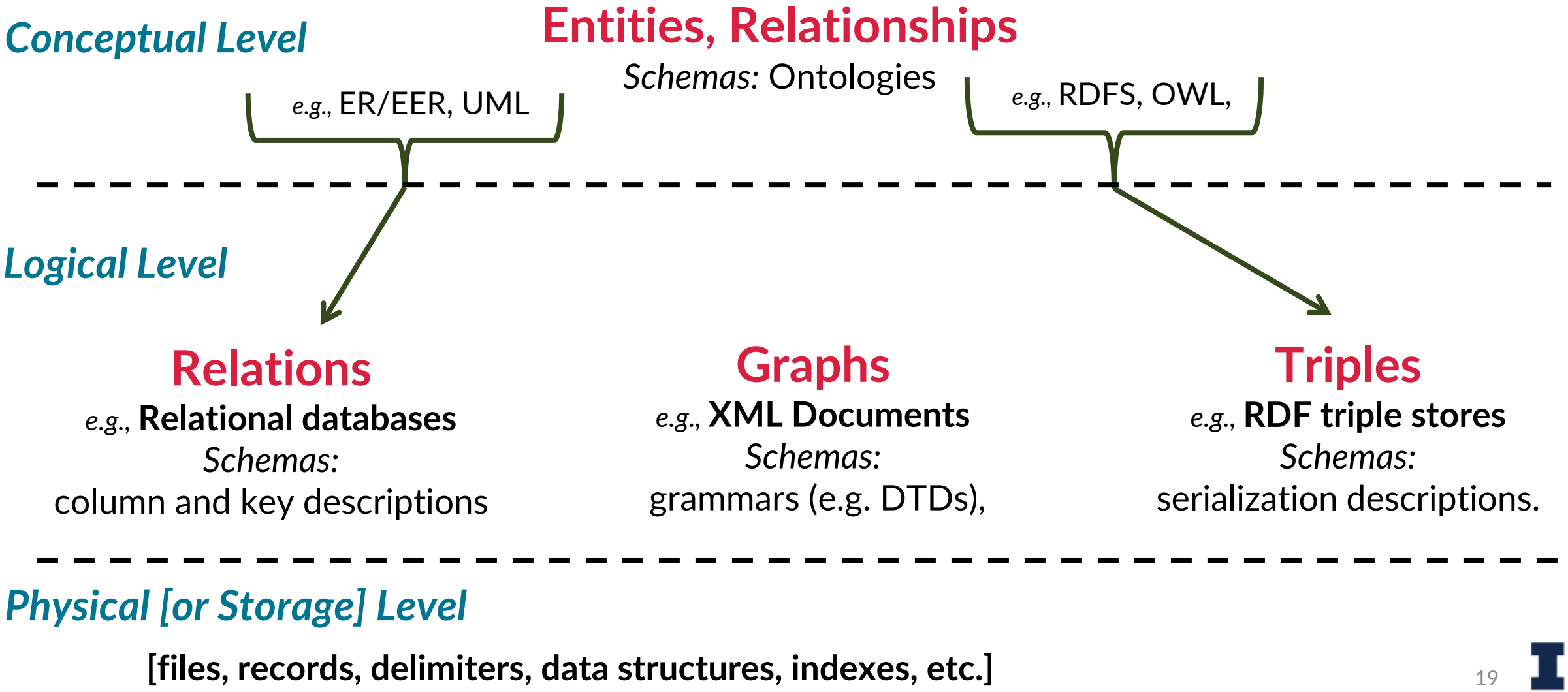
Here we add to the picture some “instance data,” four actual bibliographic entities: the work Intellectual Foundations, the text of the MIT Press 2000 edition, the edition itself, and an individual copy, all with their relevant inherent FRBR relationships asserted to hold between them.

All instance entities and relationships matched with the relevant schema level constructs

Notes: In RDFS the relationship between instance level property and the matching schema level property is identity, and so shown here with an unnamed red dotted arrow.


Endeavor
isCarriedby
Work
Item
subClassOf
subClassOf
subPropertyoff
subPropertyoff
subPropertyoff
domain
range
subClassOf
subClassOf
isRealizedBy
Expression
Manifestation
isExemplifiedBy
domain
range
domain
range
type
type
type
type
isEmbodiedBy
RealizedBy
EmbodiedBy
ExemplifiedBy
w128935
e128935
m128935
i128935
creates
range
E. Svenonius
creates
domain
type
subPropertyoff
Person
Acts_on

Data Model Relationships



FOUNDATIONS OF DATA CURATION (IS531)

Allen H. Renear, Cheryl A Thompson, Katrina S Fenlon, Myrna Morales
School of Information Sciences
University of Illinois at Urbana-Champaign

Includes material adapted from work by Carole Palmer, Melissa Cragin, David Dubin, Karen Wickett, Bertram Ludaescher, Ruth Duerr and Simone Sacchi.

Comments and corrections to: renear@illinois.edu.