A USE CASE OF ONTOLOGY ADAPTATION

the VIVO ONTOLOGY adapted to the French National Institute for Agricultural Research (INRA)

Sophie Aubin(1), Esther Dzalé Yeumo Kaboré(1), Thomas Francart(2), Sylvie Cocaud(1), Fanny Dedet(1), Pascal Aventurier(1), Linlu Li(1) and Mathieu Andro(1)

(1) Institut National de la Recherche Agronomique

(2) Sparna

1st European Ontology Network (EUON) Workshop, Amsterdam, The Netherlands 25 / 09 / 2014

INRA
SCIENCE & IMPACT

~10,000 people
- ✓researchers
- ✓ engineers
- ✓technicians

Food and nutrition

Sustainable agriculture

Genetics

Plant health

Global food balances

Global food security

Economics and society

Mechanisms of living organisms

Plant biology

Natural resources and environments

Animal biology

Global warming

Economics and social sciences

Diet and health

Biotechnologies

Agricultural systems

Green chemistry

Agroecology

Animal health

Biomass

http://www.inra.fr/

LINKED OPEN PUBLICATIONS PROJECT GOALS

✓ Become familiar with conceptual and technical issues: modelling, ontology reuse, data transformation, enrichment, publication…

✓ Demonstrate the power of Linked data to information professionals, IT specialists, scientists, and policy-makers

✓ Identify pitfalls, limits

✓ Experiment Open data ★★★★★

Make it real !

PROJECT DATA SOURCES

Institutional repository

people & structures directory

What we got

- ✓ Siloed data
- ✓ Application oriented
- ✓ XML via web services

ProdINRA

Published & grey sc. literature
198,487 records
10,1% Open Access full-text
+25 000 new entries/year
56% English, 43% French

HR

16,272 persons
+ structures
authority list

INRA keywords

59 000 entries
11 000 synonyms
19 000 English eq.

Activities

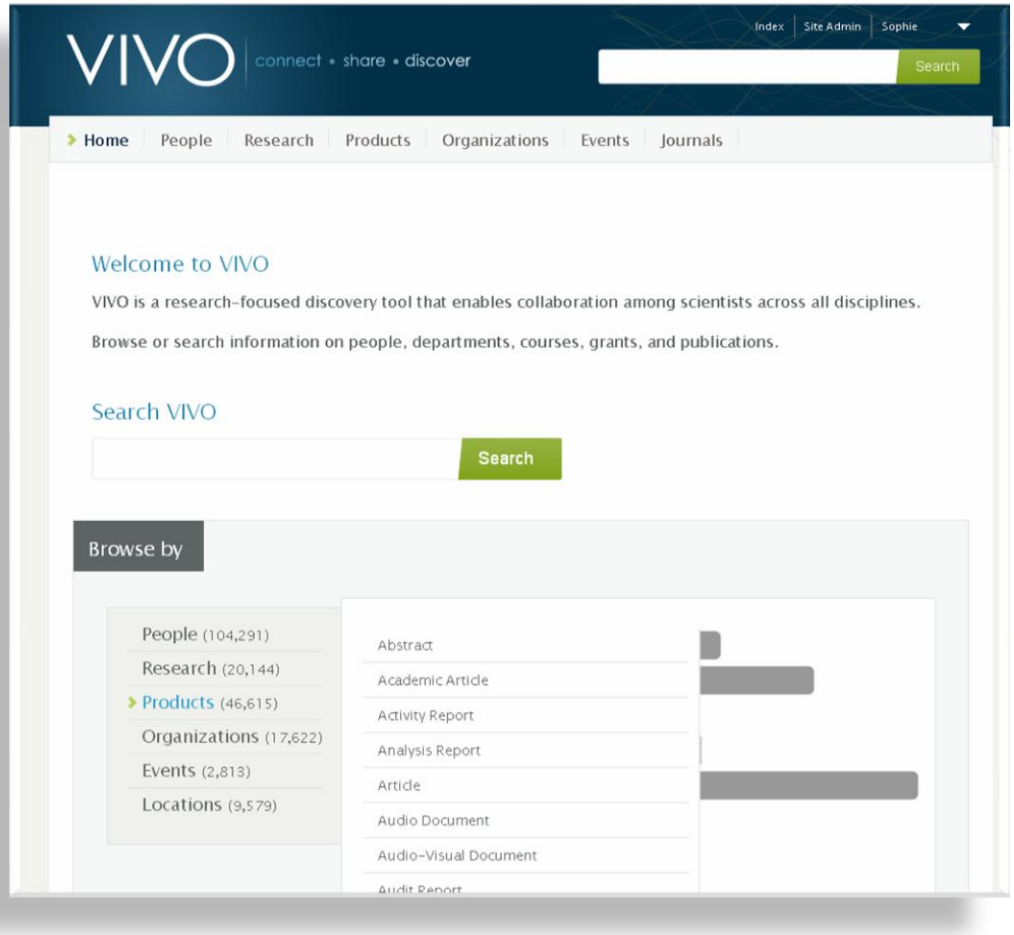
3,965 descriptions
1,426 in use

How VIVO answers our needs

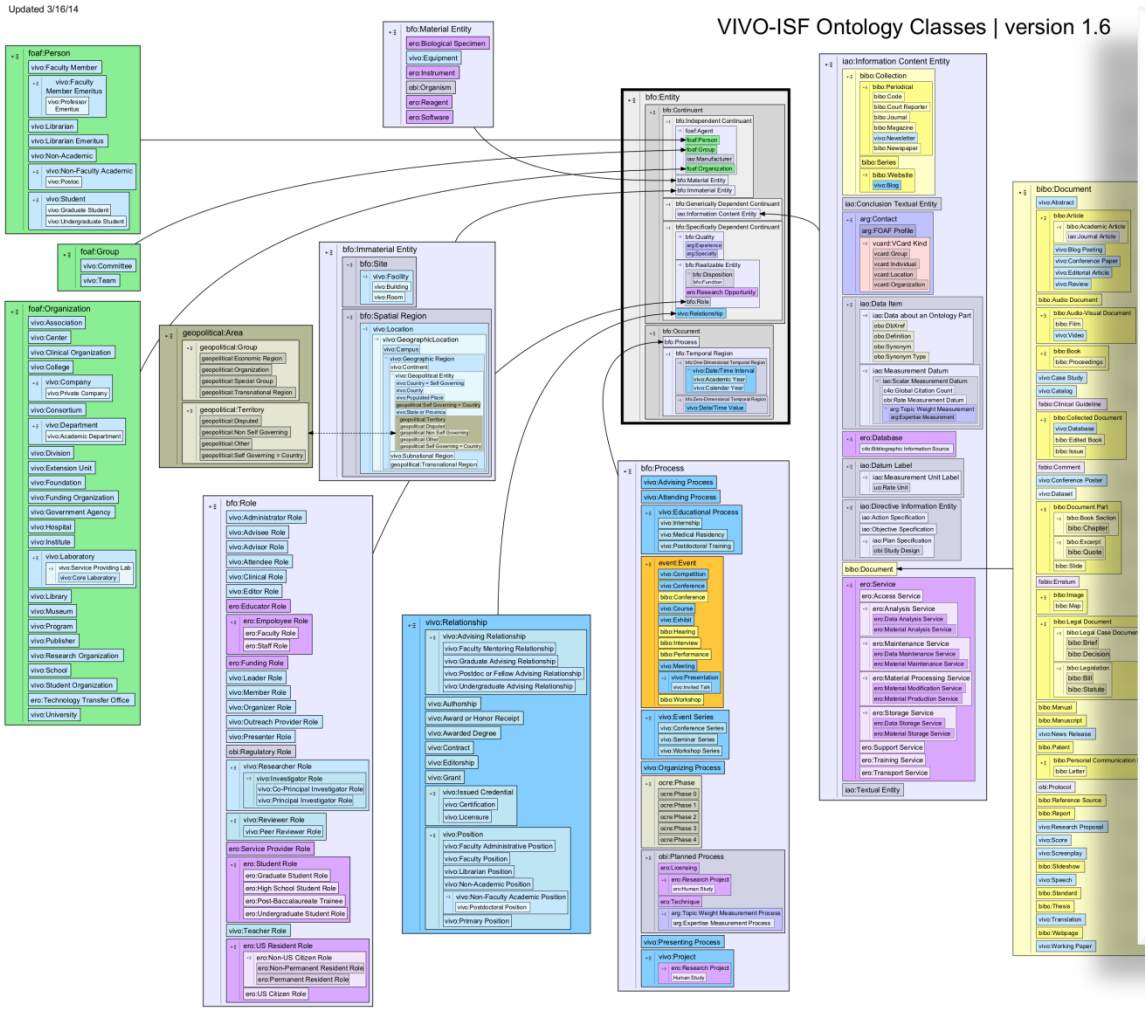
"VIVO is an open source semantic web application that enables the discovery of research and scholarship across disciplines at a particular institution and beyond."

http://www.vivoweb.org/about/faq/about-project

- ✓ **comparable services**
  - publication repository
  - person and structure directory
- ✓ **integrated data**
  - uniform representation
  - linked data
- ✓ **semantics**
  - clearly defined object
  - meaningful relations



The VIVO ontology



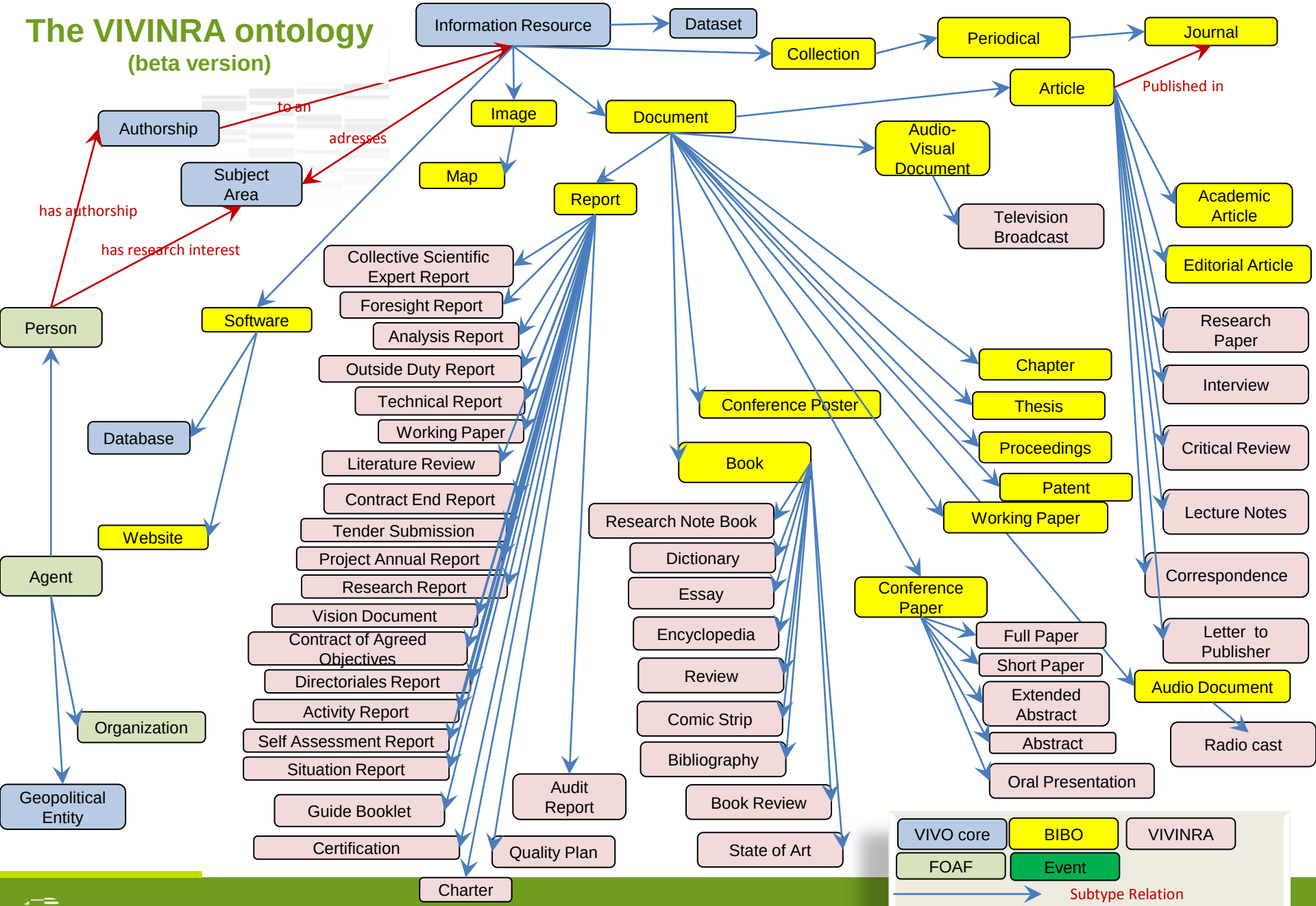
Vivo Ontology sources

VIVO
Bibliographic Ontology (BIBO)
Event Ontology
Friend of a Friend (FOAF)
Geopolitical.owl (FAO)
SKOS (Simple Knowledge Organization System)

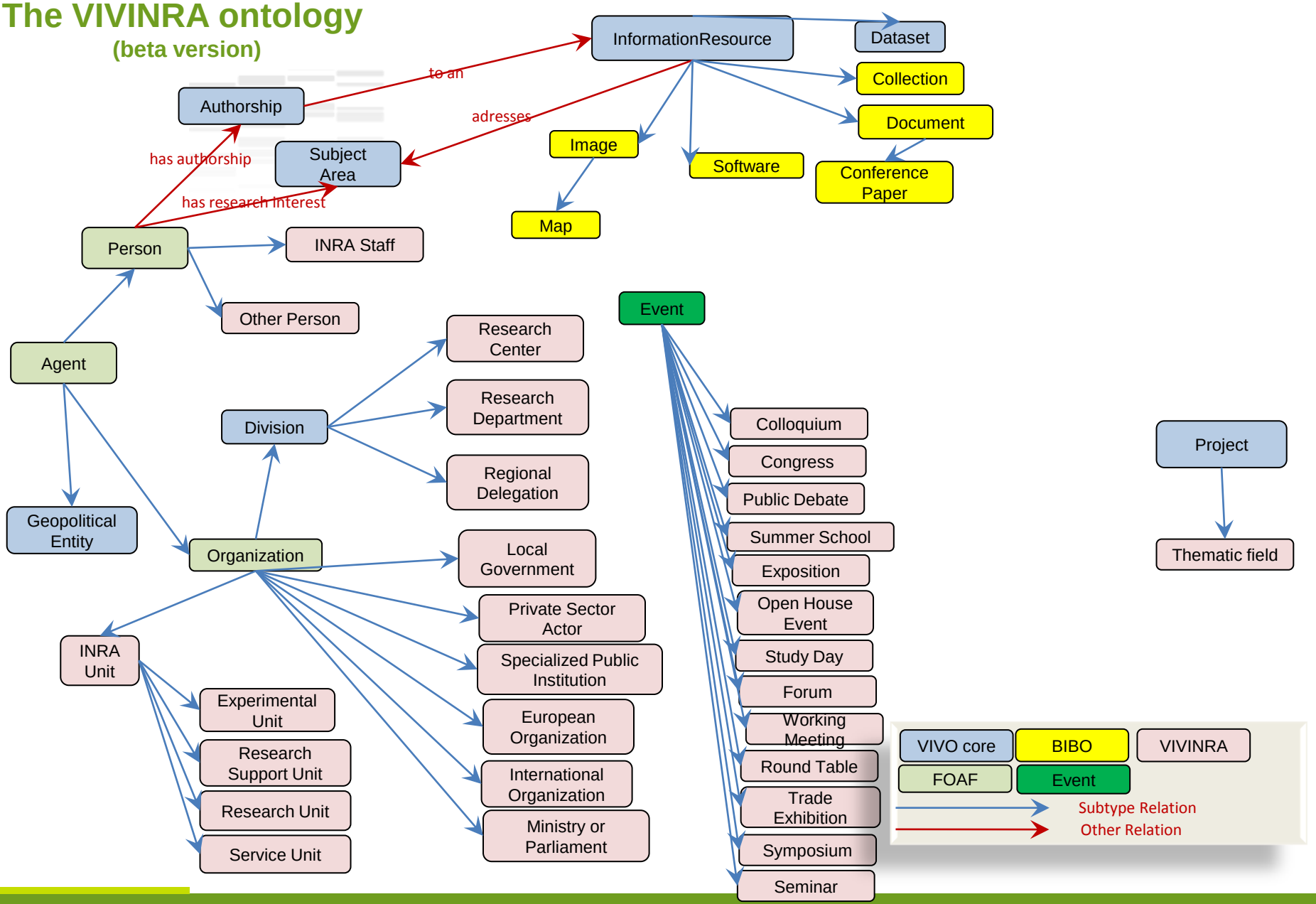
eagle-i Resource Ontology (ERO)
Basic Formal Ontology (BFO)
Cell Ontology (CL)
Gene Ontology (GO)
Information Artifact Ontology (IAO)
Ontology for Biomedical Investigations (OBI)
Ontology of Clinical Research (OCRe)
Reagent Ontology (ReO)
Relations Ontology (RO)
Software Ontology (SWO)
Sequence Ontology (SO)
Uberon (Uber anatomy ontology)
Vcard

Source: https://wiki.duraspace.org/display/VIVO/VIVO-ISF+Ontology+v1.6+Overview%3A+Classes

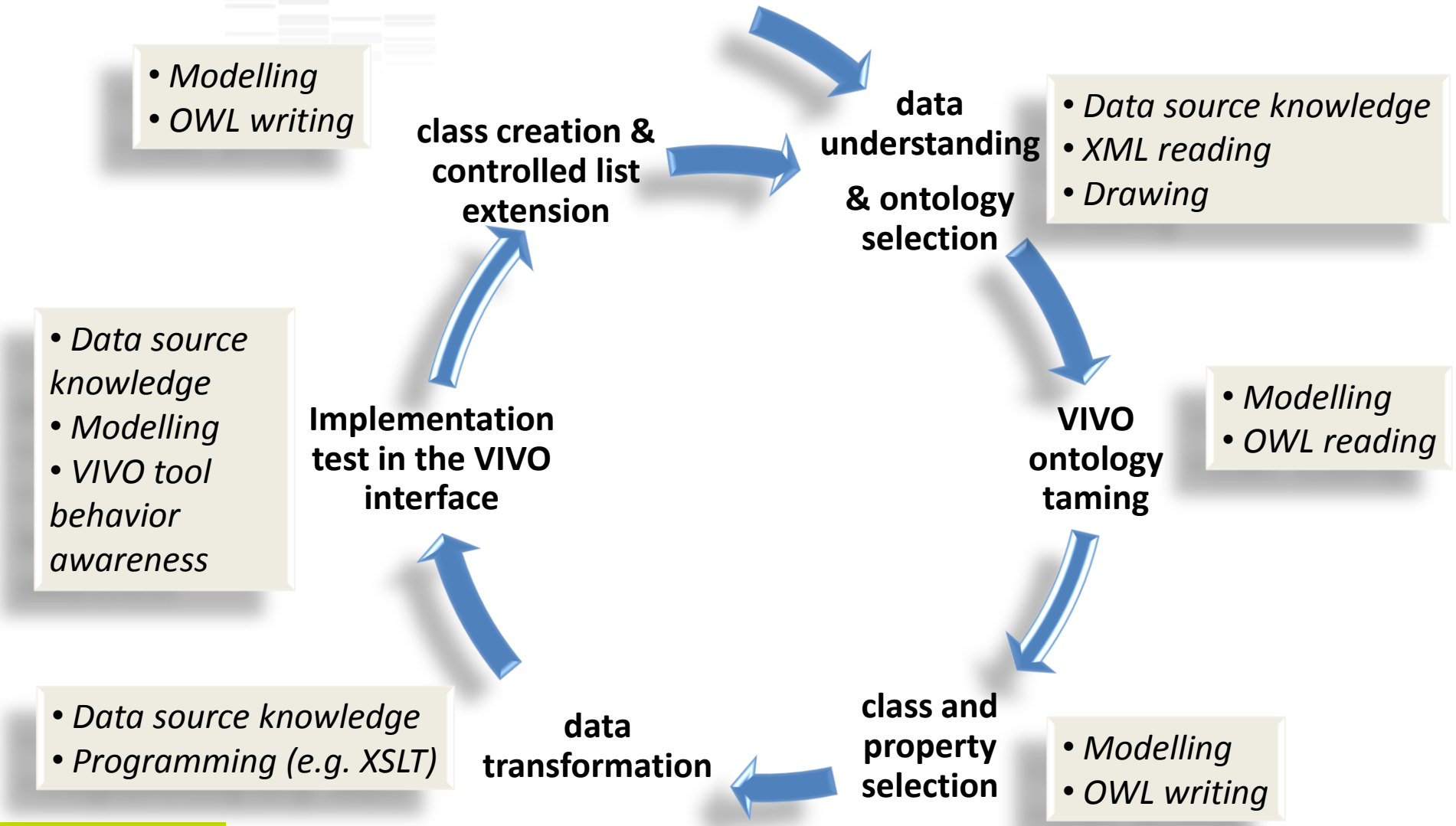
The VIVINRA ontology
(beta version)



The VIVINRA ontology
(beta version)



ONTOLOGY EXTENSION PROCESS + skills



SOME ISSUES

Is this piece of information worth modelling?

keep in mind the final application goal(s)

Did we respect the semantics?

visit other applications (ex: VIVO repos) that use the concept to get examples (ex: vivo:Private Company vs. vivinra:Private Sector Actor)

check that all the instances of a class refer to the same kind of objects in the real world

Should we create a sub-class OR a controlled value on a data property?

depends on how you want to see it in the VIVO interface (facets on classes only)

How do instances that are referenced coexist with instances that are not?

create a (potentially) underspecified instance

Should we create redundant information, e.g. relation shortcuts?

theoretically no, but you may have to in order to fit in templates

NEXT STEPS

✓Add data that were first discarded for complexity reasons

✓Test the ontology stability in another system with other usages

✓Reconsider using first [2]identified ontologies to enrich the model

✓Display the created data in another information system (e.g. Drupal) to experiment mash-up

✓Use text-mining techniques to enrich data and add links to external Linked Data

✓Use VIVO as a demonstrator to accompany the institutional policy towards open research data

THANK YOU!

VIVO Ontology: https://wiki.duraspace.org/display/VIVO/VIVO-ISF+Ontology

VIVO International Researcher Network: http://www.vivoweb.org/

VIVO Open Source Community: http://vivo.sourceforge.net/

Learn more about Inra open data policy:
http://www.ciard.net/community/interviews/towards-open-science-face-agriculture-challenges