



ONTOLOGIES

Community of Practice

Mapping Ontologies - Vocabularies

Marie Angélique Laporte
Alliance Bioversity-CIAT



Alliance



Platform for
Big Data
in Agriculture

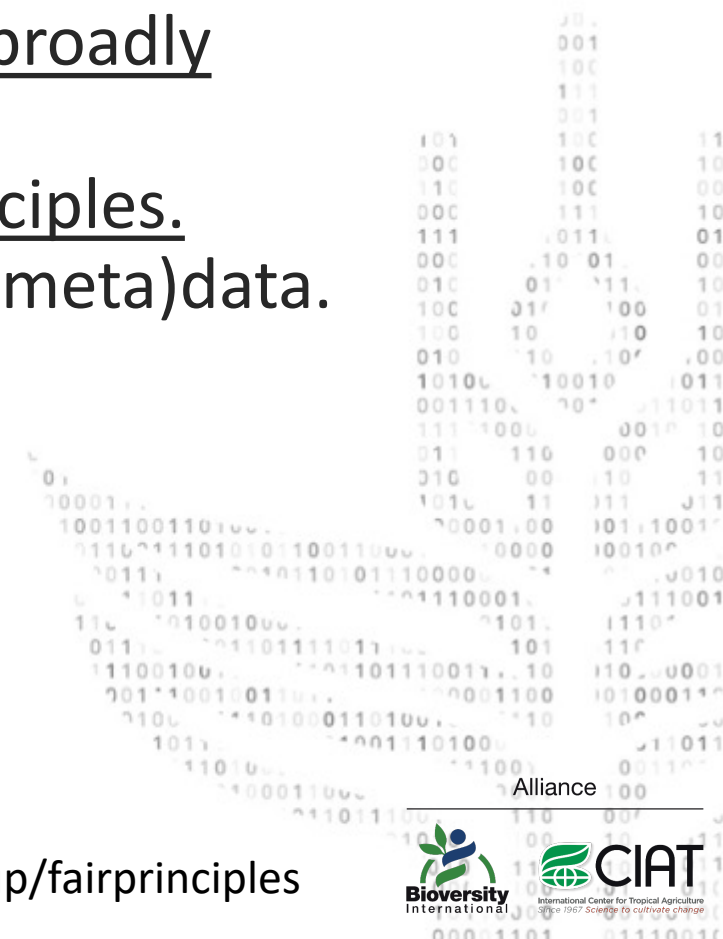
SciDataCon 2021

Ontologies, vocabularies, KOS,...???



FAIR principles

- **TO BE INTEROPERABLE:**
- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles.
- I3. (meta)data include qualified references to other (meta)data.



Knowledge Organization Systems

Different standards,
different use cases

Stronger semantics



Platform for
Big Data
in Agriculture



Ontologies

Taxonomies

Thesauri

Authority files

Glossaries

Weaker semantics

Adapted from the Semantic Spectrum



Alliance

Knowledge Organization Systems

Different standards,
different use cases

Stronger semantics



Platform for
Big Data
in Agriculture



Ontologies

Taxonomies

Thesauri

Harar = Arabica

Glossaries

Authority files

Weaker semantics

Adapted from the Semantic Spectrum



Knowledge Organization Systems

Different standards,
different use cases

Stronger semantics



Platform for
Big Data
in Agriculture



Ontologies

Taxonomies

Thesauri

Harar = Arabica

Glossaries

Authority files

Harar (coffee variety)
Is not the same as
Harar (City in Ethiopia)

Weaker semantics

Adapted from the Semantic Spectrum



Knowledge Organization Systems

Stronger semantics



Platform for
Big Data
in Agriculture



Ontologies

Different standards,
different use cases

Harar is a Coffee Variety
Coffee variety is a Variety
Harar (city) is part of Ethiopia

Taxonomies

Thesauri

Harar = Arabica

Glossaries

Authority files

Harar (coffee variety)
Is not the same as
Harar (City in Ethiopia)

Weaker semantics

Adapted from the Semantic Spectrum



Alliance

Knowledge Organization Systems

Stronger semantics



Platform for
Big Data
in Agriculture



Ontologies

Different standards,
different use cases

Harar is a Coffee Variety
Coffee variety is a Variety
Harar (city) is part of Ethiopia

Taxonomies

Thesauri

Harar is related to Harar (city)
Harar is related to coffee

Harar = Arabica

Glossaries

Authority files

Harar (coffee variety)
Is not the same as
Harar (City in Ethiopia)

Weaker semantics

Adapted from the Semantic Spectrum



Alliance

Knowledge Organization Systems

Stronger semantics



Platform for
Big Data
in Agriculture



Ontologies

Harar is a subclass of Coffee Variety
Coffee Variety is a subclass of Variety
Variety has exactly one Country of Origin
Coffee is derived from Harar

Different standards,
different use cases

Harar is a Coffee Variety
Coffee variety is a Variety
Harar (city) is part of Ethiopia

Taxonomies

Thesauri

Harar is related to Harar (city)
Harar is related to coffee

Harar = Arabica

Glossaries

Authority files

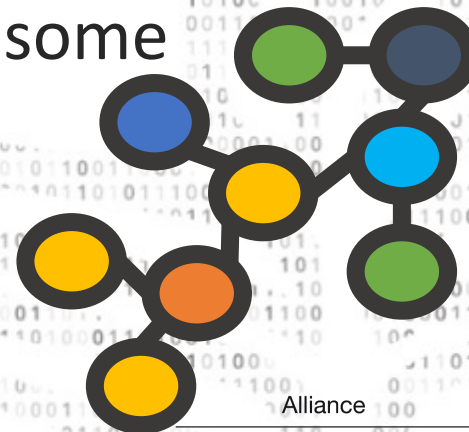
Harar (coffee variety)
Is not the same as
Harar (City in Ethiopia)

Weaker semantics

Adapted from the Semantic Spectrum

Knowledge Organization Systems

- provide a **shared vocabulary** for a domain (all the terms)
- provide **textual definitions** that describe the intended meaning of the terms in vocabularies
- provide **standard identifiers** for concepts describing a given domain
- **Machine readable** format. Ontologies provide **machine-readable axioms** and definitions that enable computational access to some aspects of the meaning of classes and relations – logical representation of human knowledge



Alliance



How to take advantage of existing standards?



Mapping standards

- Mappings are essential for bridging semantic standards (silos) and therefore use cases
- Ideal scenario: reuse of identifiers all the time
- Mapping creation:
 - Automatic mappings are efficient but not that accurate
 - Manual mappings done by curators/domain experts are very accurate but take time
- Mappings need to be maintained/redone every time the standards evolve (which can be very often)
- Rely on URIs/permanent identifiers



Mapping standards

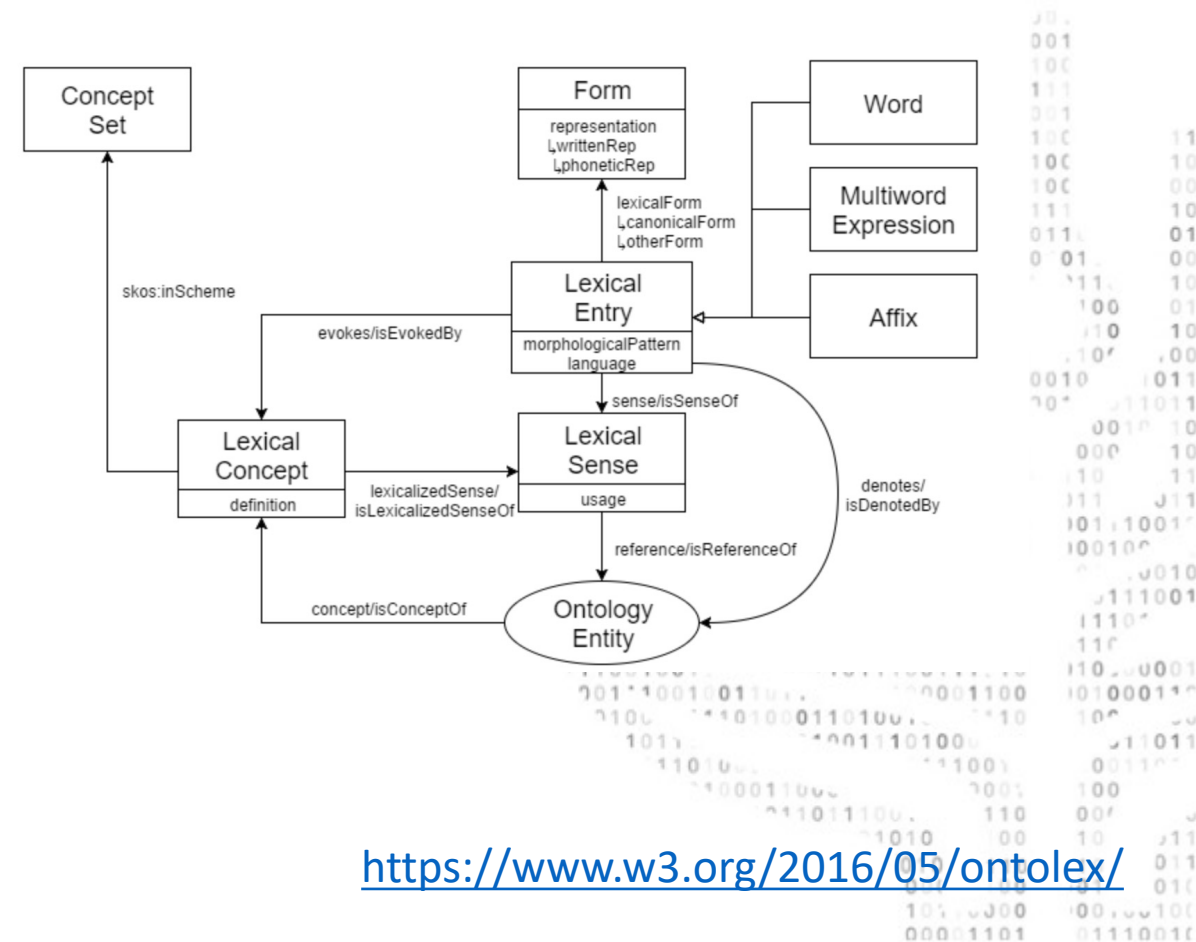
- Sharing ontology mappings:
 - SSSOM is a Simple Standard for Sharing Ontology Mappings, providing
 - a TSV-based representation for ontology term mappings
 - a comprehensive set of standard metadata elements to describe mappings and
 - a standard translation between the TSV and the Web Ontology Language (OWL).

<https://github.com/mapping-commons/sssom>

subject_id	predicate_id	object_id	match_type	subject_label	object_label
HP:0009124	skos:exactMatch	MP:0000003	Lexical	Abnormal adipose tissue morphology	abnormal adipose tissue morphology
HP:0008551	skos:exactMatch	MP:0000018	Lexical	Microtia	small ears
HP:0000411	skos:exactMatch	MP:0000021	Lexical	Protruding ear	prominent ears

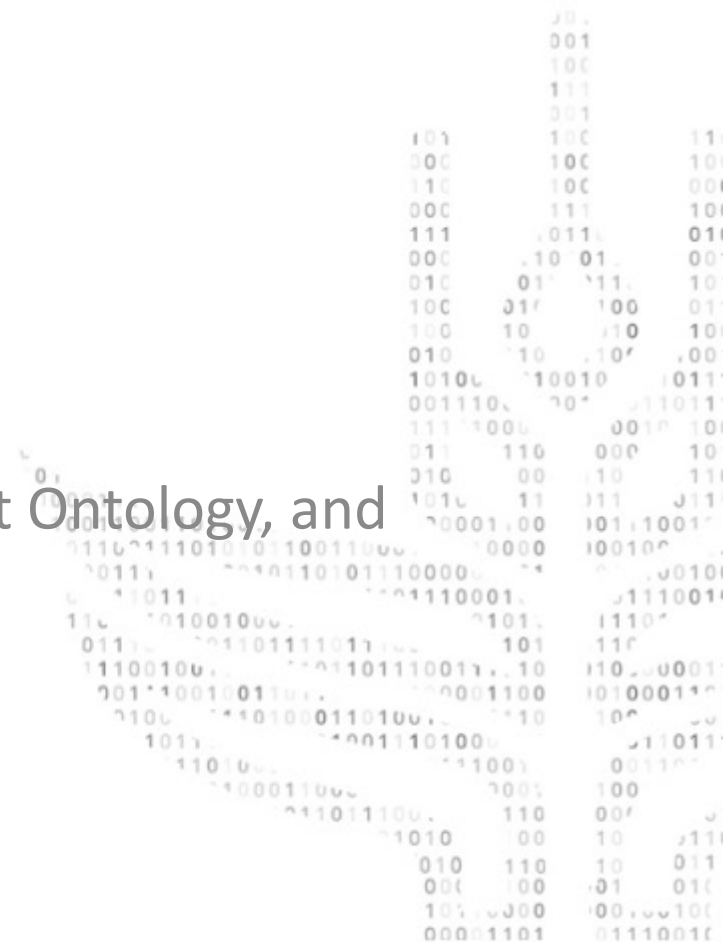
Mapping standards

- Lexicon Model for Ontologies
 - Aim: provide rich linguistic grounding for ontologies
 - providing a vocabulary that allows ontologies to be enriched with information about how the vocabulary elements described in them are realized linguistically, in particular in natural languages.



Mapping vocabularies to ontologies

Examples of mappings between the Crop Ontology and the Trait Ontology, and between the Agronomy Ontology and Agrovoc



Created in 2009 by the Integrated Breeding Platform for breeders' traits

To date

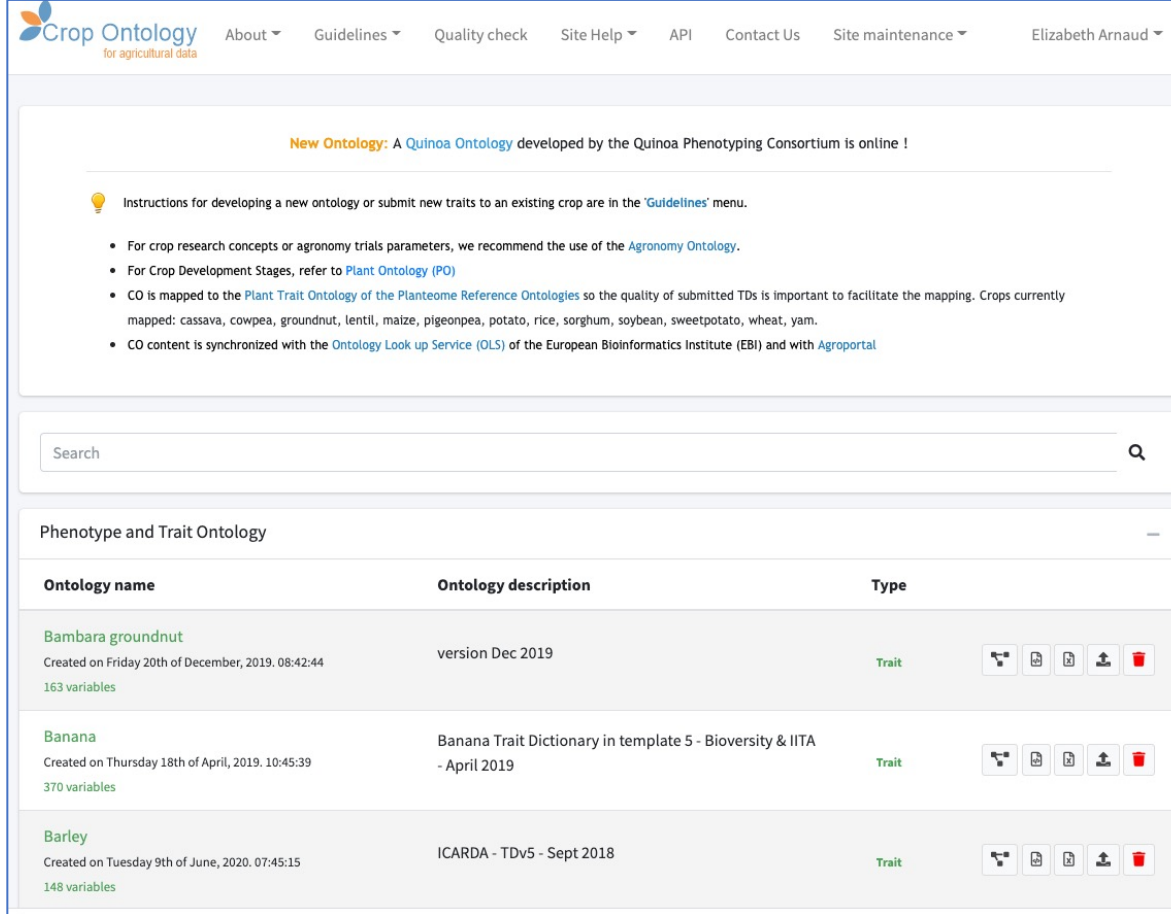
- 33 species
- > 4,300 traits
- > 6,300 variables



Last crops:

- **Quinoa** by the *Quinoa Phenotyping Consortium*
- **Coconut** by *CIRAD on behalf of COGENT*

CC-BY 4.0



The screenshot shows the Crop Ontology website. At the top is a navigation bar with links: About, Guidelines, Quality check, Site Help, API, Contact Us, Site maintenance, and Elizabeth Arnaud. Below the navigation bar is a banner for a new ontology: "New Ontology: A Quinoa Ontology developed by the Quinoa Phenotyping Consortium is online!". Below the banner is a section titled "Instructions for developing a new ontology or submit new traits to an existing crop are in the 'Guidelines' menu." followed by a list of instructions. Below this is a search bar. At the bottom is a table titled "Phenotype and Trait Ontology" with columns: Ontology name, Ontology description, and Type. The table lists three ontologies: Bambara groundnut, Banana, and Barley.

Ontology name	Ontology description	Type
Bambara groundnut Created on Friday 20th of December, 2019. 08:42:44 163 variables	version Dec 2019	Trait
Banana Created on Thursday 18th of April, 2019. 10:45:39 370 variables	Banana Trait Dictionary in template 5 - Bioversity & IITA - April 2019	Trait
Barley Created on Tuesday 9th of June, 2020. 07:45:15 148 variables	ICARDA - TDv5 - Sept 2018	Trait

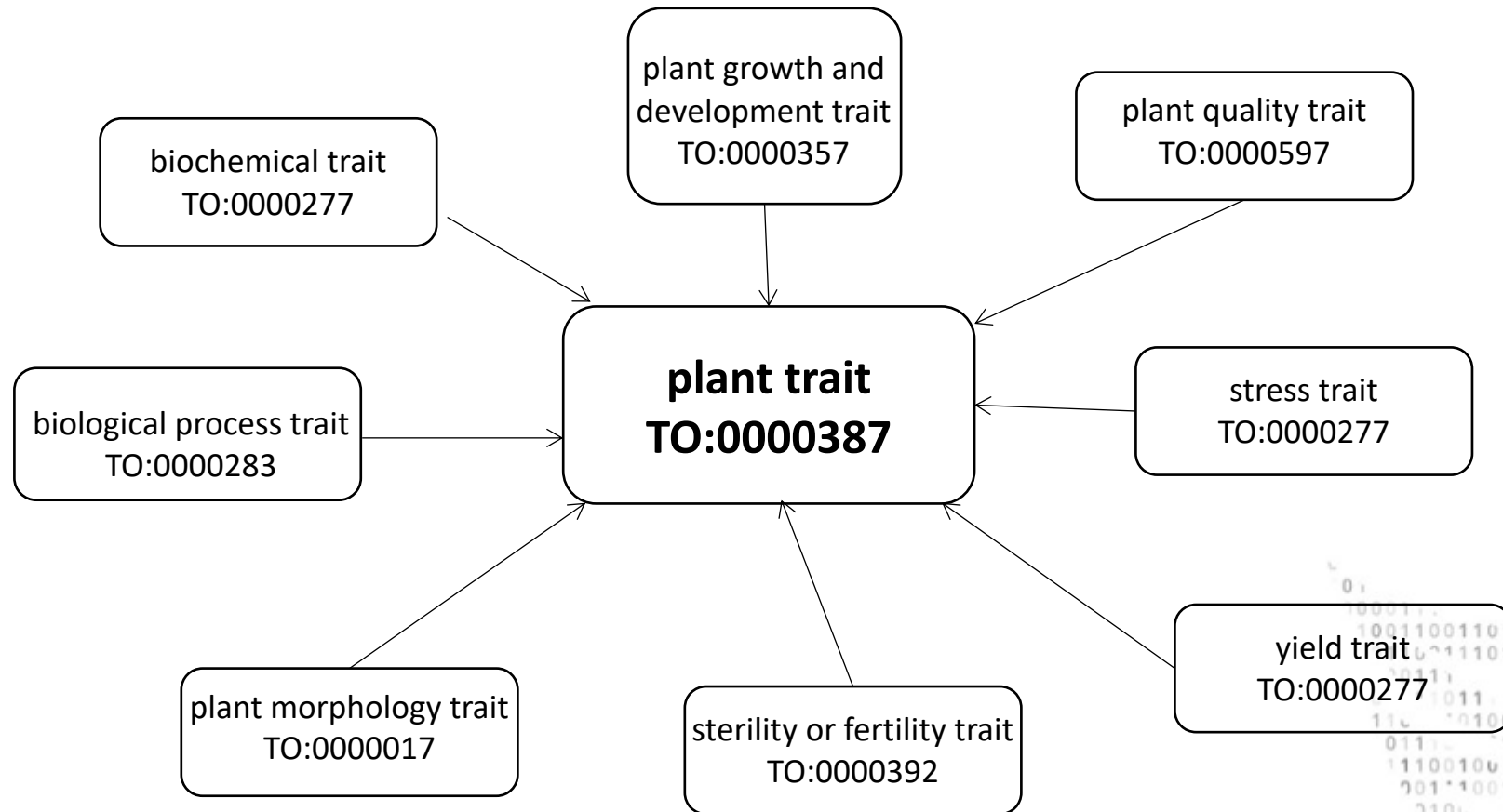
GUIDELINES FOR CREATING
CROP-SPECIFIC ONTOLOGIES
TO ANNOTATE
PHENOTYPIC DATA

Crop Ontology
for agricultural data

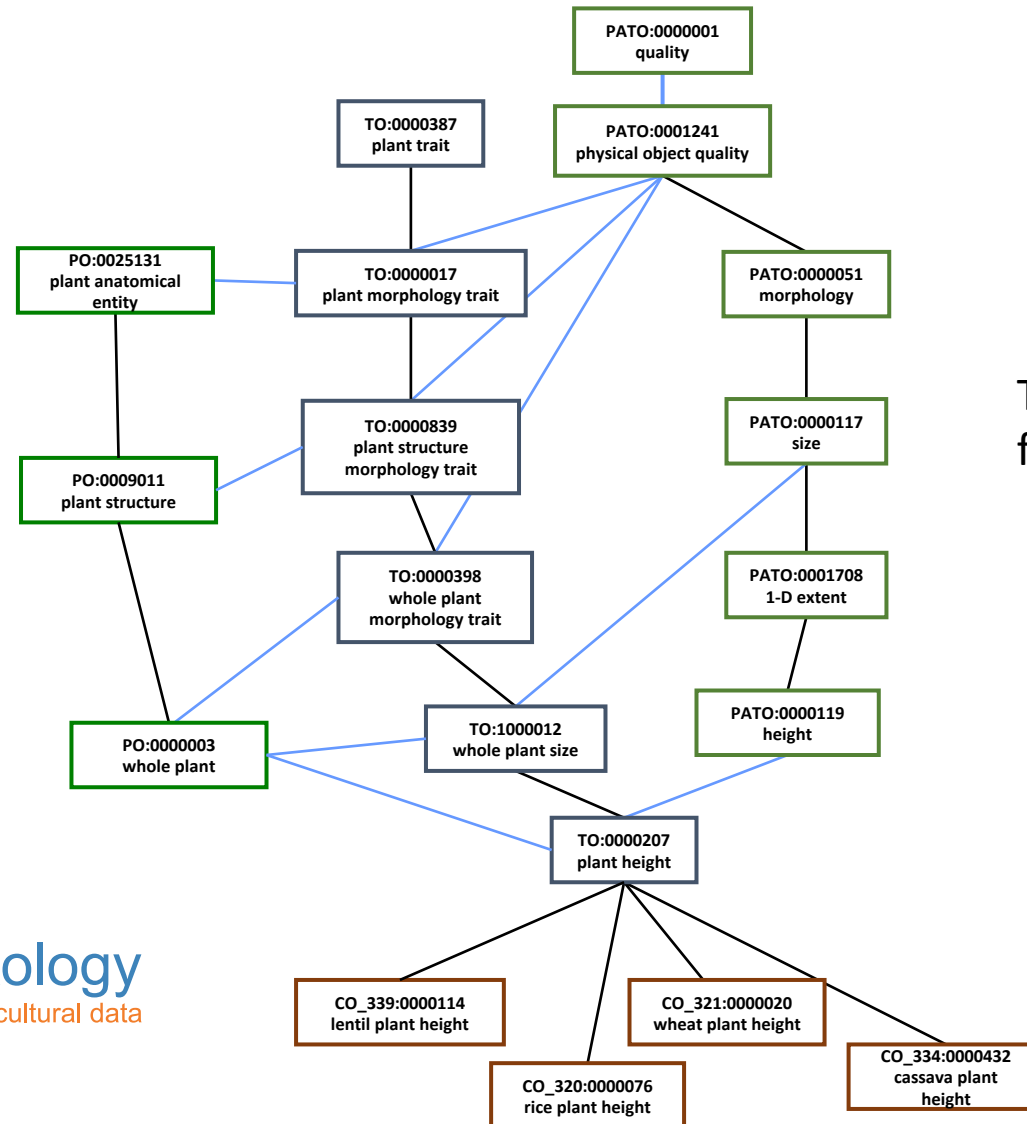




Plant Trait Ontology



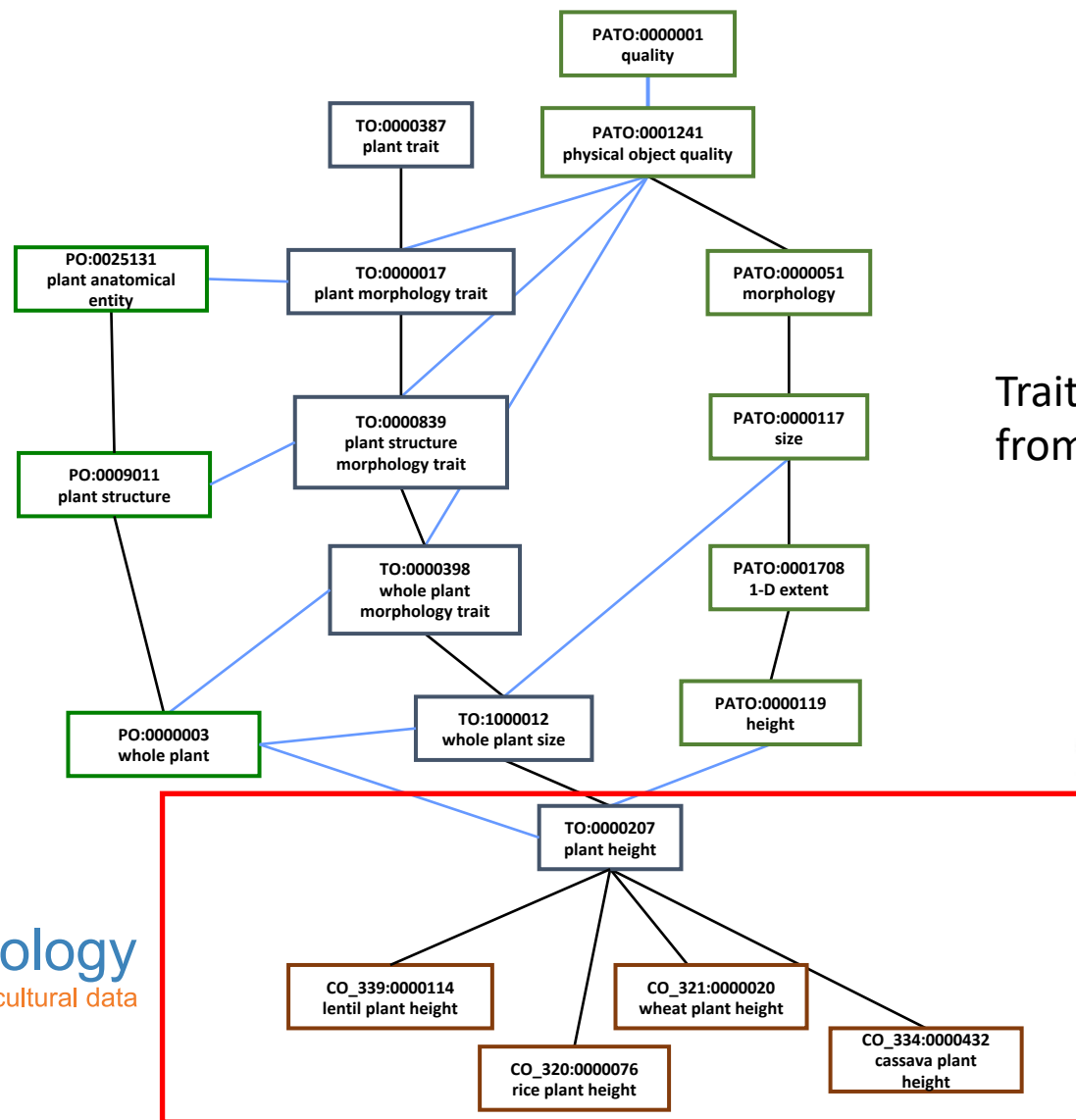
Entities from
Plant Ontology or
Gene Ontology



Trait qualities
from PATO



Entities from
Plant Ontology or
Gene Ontology



Trait qualities
from PATO



Agronomy Ontology

thinning

http://purl.obolibrary.org/obo/AGRO_00000027  Copy

A planned process in which selected parts of the plant are removed. The practice entails targeted removal of diseased, damaged, dead, non-productive, structurally unsound, or otherwise unwanted tissue from plants. This is done in order to shape the plant or tree, improve plant or tree health, reduce the risk from falling branches, prepare nursery specimens for transplanting, and increase the yield or quality of harvested products. [https://en.wikipedia.org/wiki/Pruning#Types_of_pruning http://aims.fao.org/aos/agrovoc/c_6274]

Synonyms: thinning process

Tree view

Term mappings

Term history

agricultural process

thinning

desuckering

detrashing

Graph view

Reset tree

Show all siblings

☒ Preferred root terms
 ☐ All terms

Term information

creator

<https://orcid.org/0000-0001-6284-4821>

description

Agricultural process during which removal of some plants, or parts of plants to make room for the growth of others but does not involve the cutting of the whole tree.

Agronomy Ontology

thinning

http://purl.obolibrary.org/obo/AGRO_00000027  Copy

A planned process in which selected parts of the plant are removed. The practice entails targeted removal of diseased, damaged, dead, non-productive, structurally unsound, or otherwise unwanted tissue from plants. This is done in order to shape the plant or tree, improve plant or tree health, reduce the risk from falling branches, prepare nursery specimens for transplanting, and increase the yield or quality of harvested products. [https://en.wikipedia.org/wiki/Pruning#Types_of_pruning

http://aims.fao.org/aos/agrovoc/c_6274]

Synonyms:

 Tree view

 Term mappings

 Term history

 agricultural process

 **thinning**

desuckering

detrashing

 Graph view

Reset tree

Show all siblings

☒ Preferred root terms

☐ All terms

Term information

creator

<https://orcid.org/0000-0001-6284-4821>

description

Agricultural process during which removal of some plants, or parts of plants to make room for the growth of others but does not involve the cutting of the whole tree.

AGROVOC

substances > soil amendments > mulches > organic mulches > bark mulches

PREFERRED TERM

① **bark mulches** 

DEFINITION

① A mulch where the outer corky bark layer of timber trees, shredded or not, are used to cover soil surface.
(en)

BROADER CONCEPT

[organic mulches](#) (en)

IN OTHER LANGUAGES

① نشارة اللحاء	Arabic
① 树皮覆盖层	Chinese
① 木屑土	
① kůrové mulče	Czech
① paillis d'écorces	French
① Rindenmulch	German
① Pacciamatura di corteccia	Italian
① древесная кора	Russian
① Mantillos de corteza	Spanish
① ağaç kabuğu malçı	Turkish

URI

http://aims.fao.org/aos/agrovoc/c_bc5f6400 

Download this concept:

[RDF/XML](#) [TURTLE](#) [JSON-LD](#)

Created 5/26/20, last modified 8/23/21

Download this concept:

[RDF/XML](#) [TURTLE](#) [JSON-LD](#)

Created 5/26/20, last modified 8/23/21



AGROVOC



Platform for
Big Data
in Agriculture



substances > soil amendments > mulches > organic mulches > bark mulches

PREFERRED TERM

bark mulches



DEFINITION

A mulch where the outer corky bark layer of timber trees,

Source: http://purl.obolibrary.org/obo/AGRO_00000104

Created: 5/26/20

BROADER CONCEPT

organic mulches (en)

IN OTHER LANGUAGES

نشارة اللحاء	Arabic
树皮覆盖层	Chinese
木屑土	
kůrové mulče	Czech
paillis d'écorces	French
Rindenmulch	German
Pacciamatura di corteccia	Italian
древесная кора	Russian
Mantillos de corteza	Spanish
ağaç kabuğu malçı	Turkish

URI

http://aims.fao.org/aos/agrovoc/c_bc5f6400

Download this concept:

[RDF/XML](#) [TURTLE](#) [JSON-LD](#)

Created 5/26/20, last modified 8/23/21

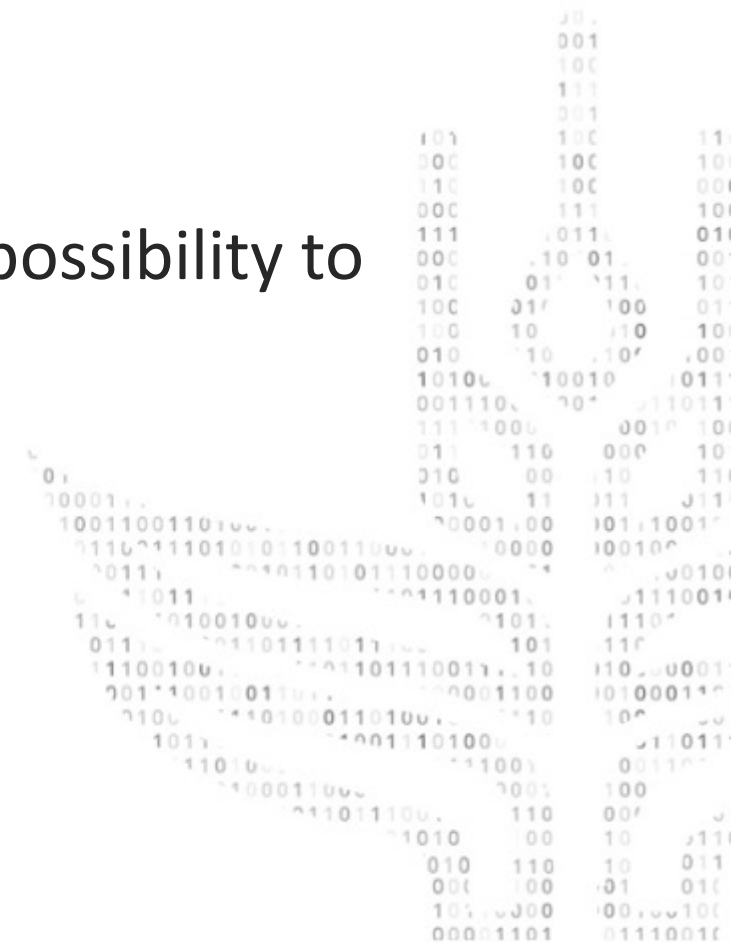


Conclusion



Conclusion

- Persistent Unique Identifiers
- Level of semantics needed according to use case
- Mappings have a cost
- Contribution to existing standards is always better: possibility to extract terms that are needed
- Long term vision: One computer, one dataset



Thank you



Platform for
Big Data
in Agriculture



ONTOLOGIES
Community of Practice

Alliance

