



# Ontology mapping for wheat trait information management



Claire Nédellec, Sophie Aubin, Clara Sauvion, Liliana Ibanescu, Sonia Bravo, Jacques Le Gouis, Thierry C. Marcel, Cyril Pommier, Robert Bossy, Michaël Alaux

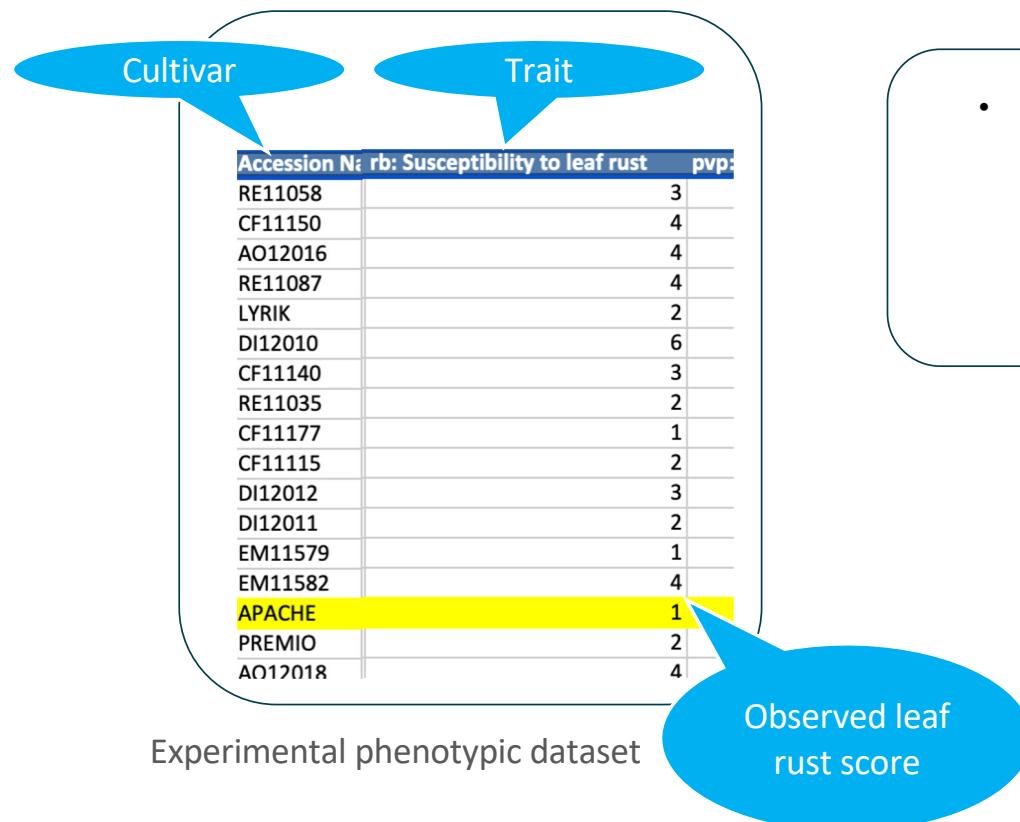
Séminaire *semantic linked data*,  
10-13 oct 2023

## Un exemple de valeur phénotypique : susceptibilité à la rouille des feuilles (*susceptible to leaf rust*)



Trait correspondant : resistance to *leaf rust*

# Experimental data and text - complementary sources of information



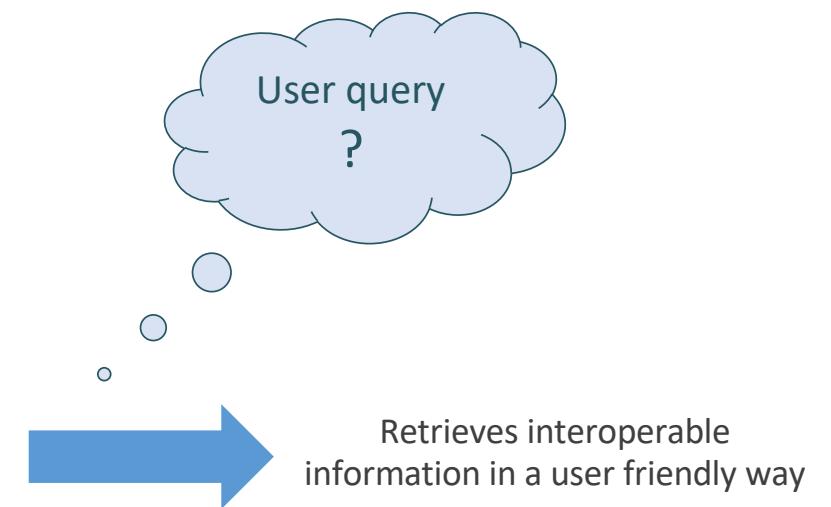
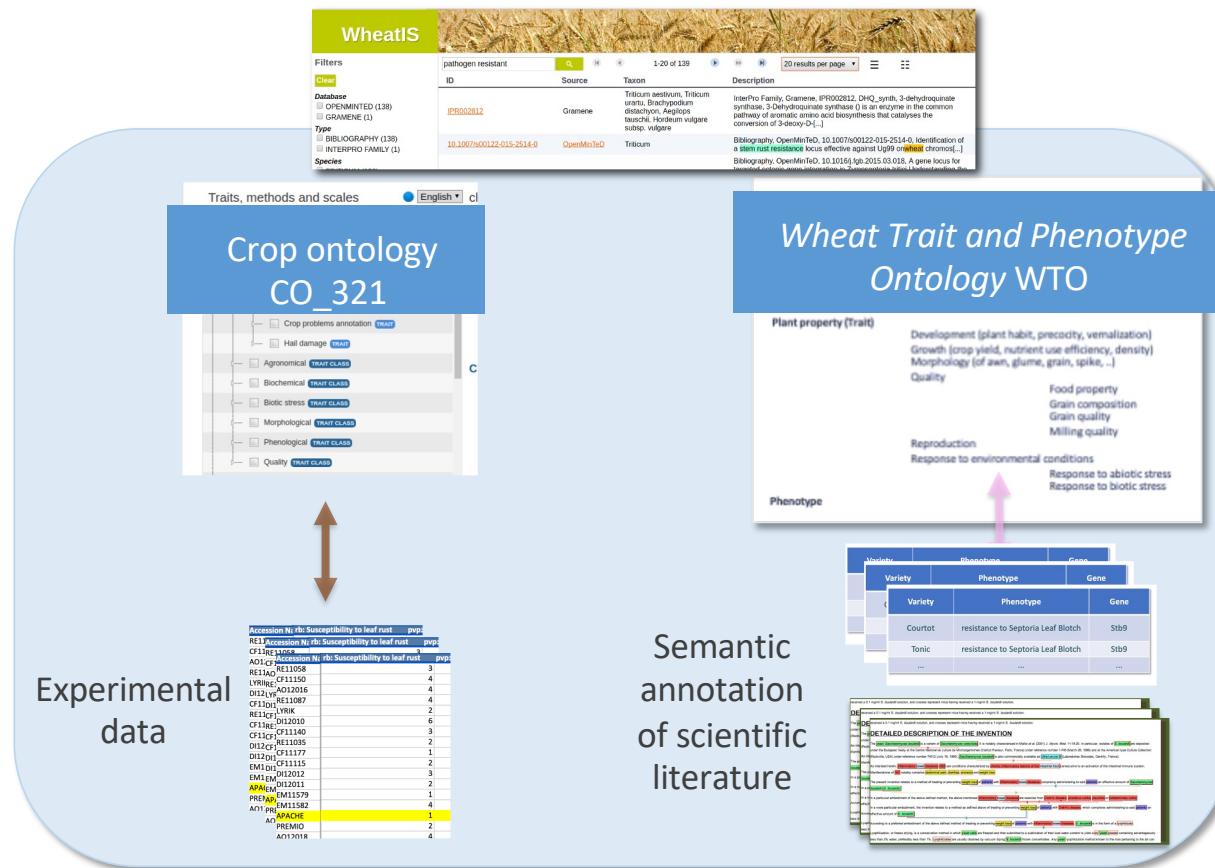
- These assays included inoculation with three different pathotypes on a segregating population originating from a cross between cv. *Apache* and cv. *Taldor*, a French cultivar susceptible to stripe rust
  - moderately tolerant to stripe rust
  - highly resistant to stripe rust
  - ...

Scientific articles



# Semantic web technologies for a unified Knowledge Base

Phenotypic data and literature information on bread wheat cultivars



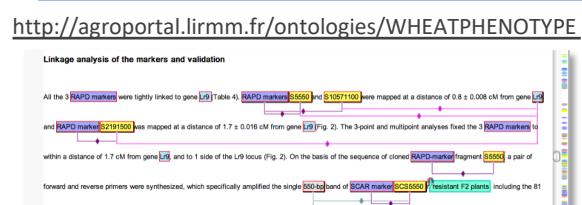
## Components

- (1) Experimental data indexing and lifting
- (2) Literature semantic annotation and lifting
- (3) Mapping of ontology classes and lifting
- (4) Knowledge model design

### (3) Trait and phenotype information, two complementary ontologies

#### WTO

- Wheat Trait&Phenotype Ontology
- 753 traits and phenotypes
- Deep structure: 9 levels



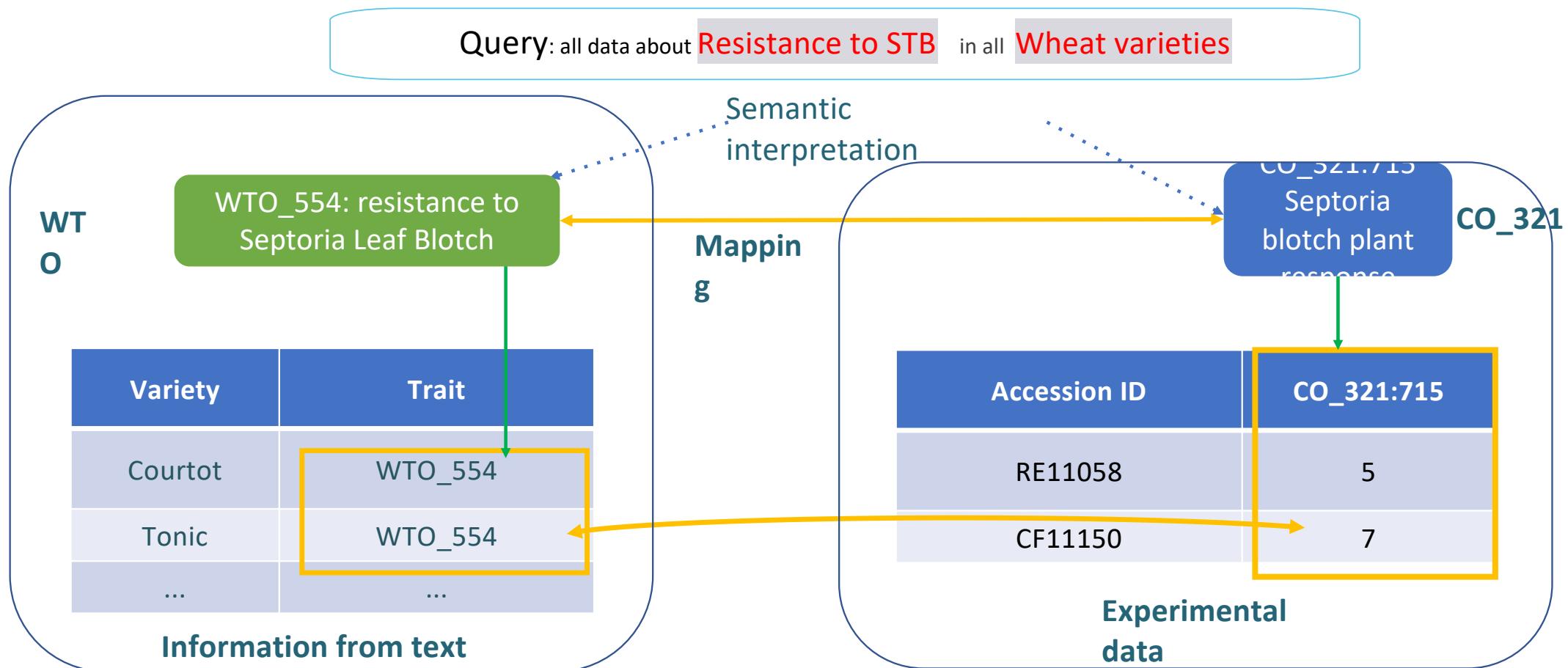
#### CO\_321

- Crop Ontology for wheat
- Observation variables
  - trait +method+scale
- 460 trait classes
- Flat structure: 3 levels



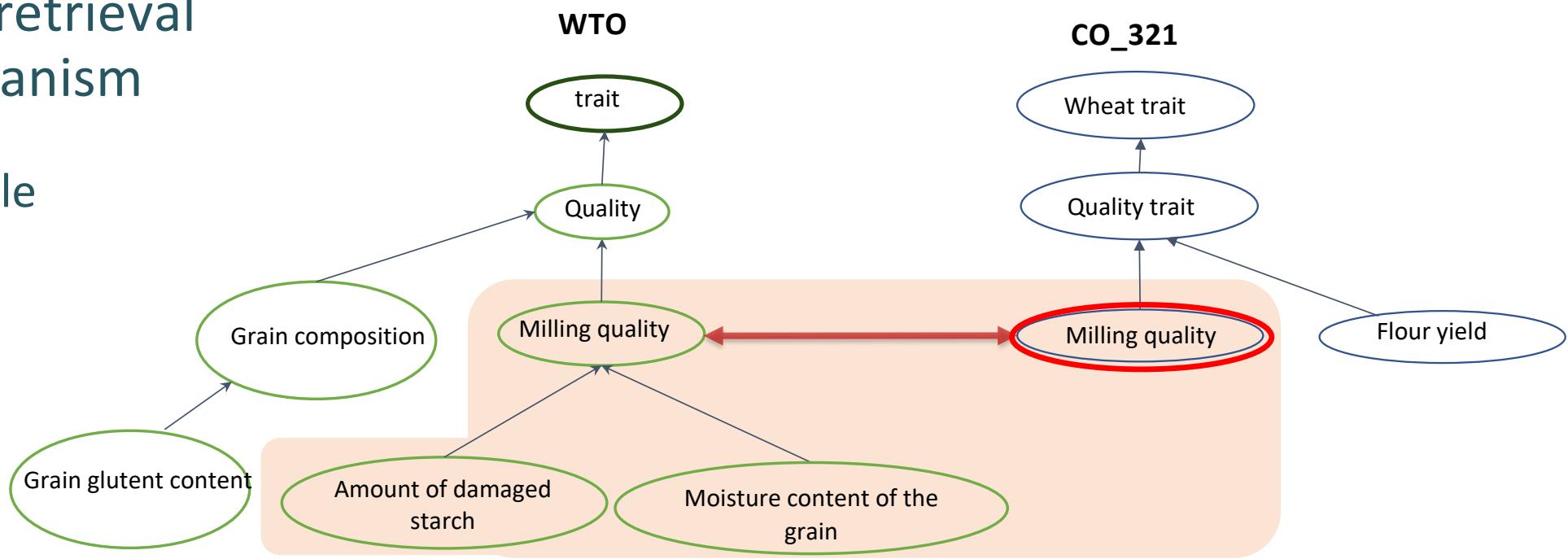
Crop Ontology for agricultural data  
Template and Guidelines :  
[www.cropontology.org](http://www.cropontology.org)

# Searching for variety traits in textual and experimental databases



# Data retrieval mechanism

## Example



Query: All data about « milling quality » of wheat varieties?

Data indexed by all subsumed classes:

WTO:Milling quality      CO:Milling quality

WTO:Moisture content of the grain

WTO:Amount of damaged starch

*more specific classes in WTO*

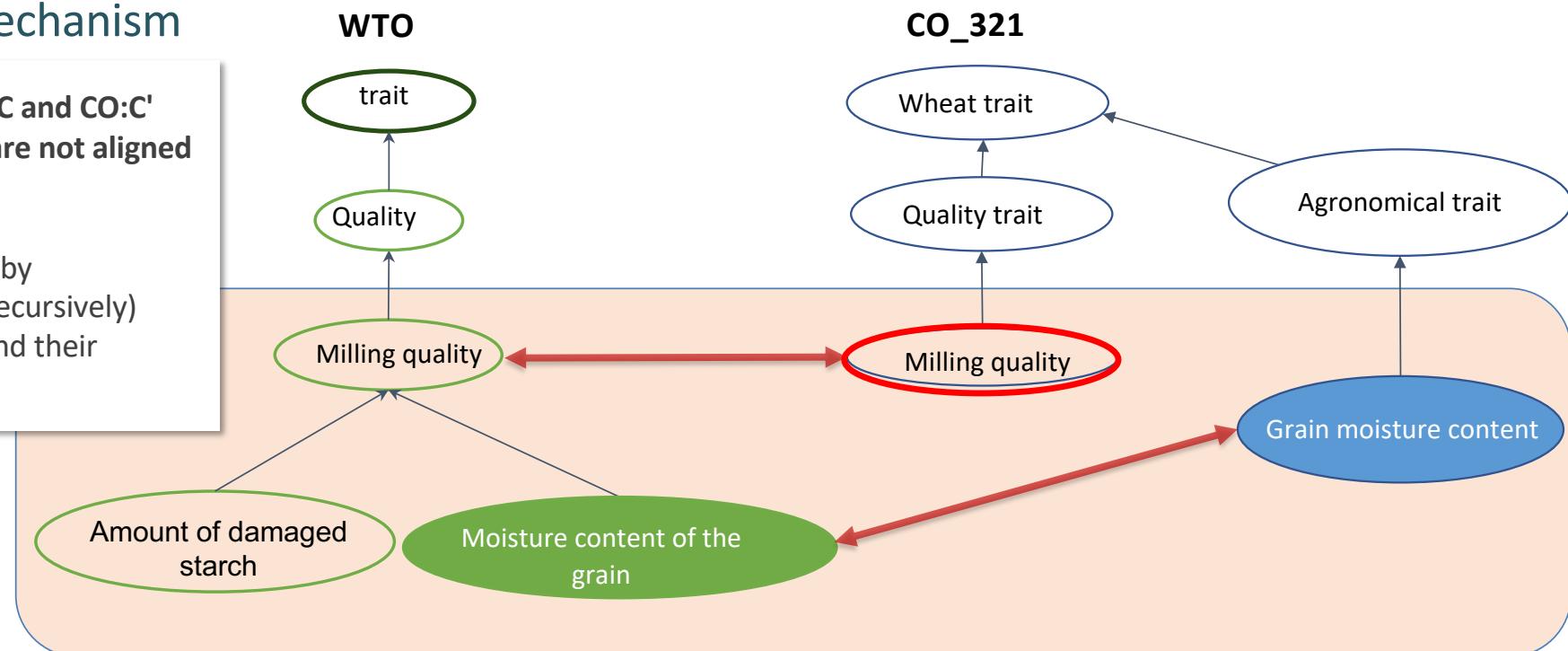
## Data retrieval mechanism

Two aligned classes WTO:C and CO:C' may have ancestors that are not aligned

### Strategy

Union of the data indexed by

- all subsumed classes (recursively)
- all equivalent classes and their subsumed classes



**Query: All data about « Milling quality » of wheat varieties?**

Data indexed by all subsumed classes= class subtrees:

WTO: Milling quality CO:Milling quality

more specific  
classes in  
WTO

WTO:Moisture content of the grain  
WTO:Amount of damaged starch

CO:Grain Moisture content  
in CO

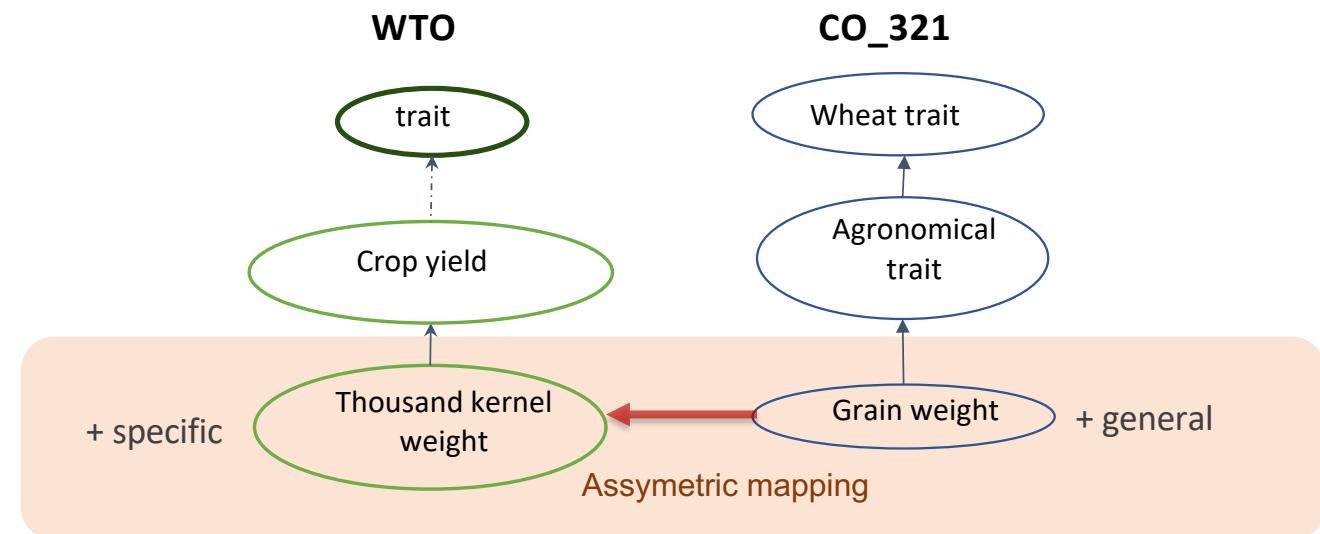
p. 11

# Data retrieval mechanism Asymmetric mapping relation

CO:class is more general than any WTO classes (or the reverse)

## Strategy

Asymmetry of the mapping  
Inference, but not equivalence



Query: All data about « Grain weight» of wheat varieties?

Retrieved data indexed by :  
WTO: Thousand kernel weight    CO: Grain weight

Query: All data about « Thousand kernel weight» of wheat varieties?

Retrieved data indexed by :  
CO: Thousand kernel weight



# Typology of mapping justifications

- 0** The classes are different
- 1** The classes are equivalent
  - 1.1** The classes have the same label or are slightly different, but have the same sense
  - 1.2** The classes share labels or synonyms
  - 1.3** The classes have different labels, but their definitions are similar, or the definition includes a label or synonym of the other class
  - 1.4** The classes are equivalent with respect to formal rules
  - 1.5** The mapping uses supplementary information (e.g., reference)
- 4,5,6** Complex mapping
- 7** The CO\_321 class is more general than the WTO class, and there is no relevant, more specific class
- 8** The WTO class is more general than the CO\_321 class, and there is no relevant, more specific class

## Examples of mapped classes

WTO:plant height      CO\_321:plant height      exact

WTO:plant habit      CO\_321:plant growth habit      exact

WTO:glutenin content      CO\_321:glutenin Glu-D3 composition      + gen.

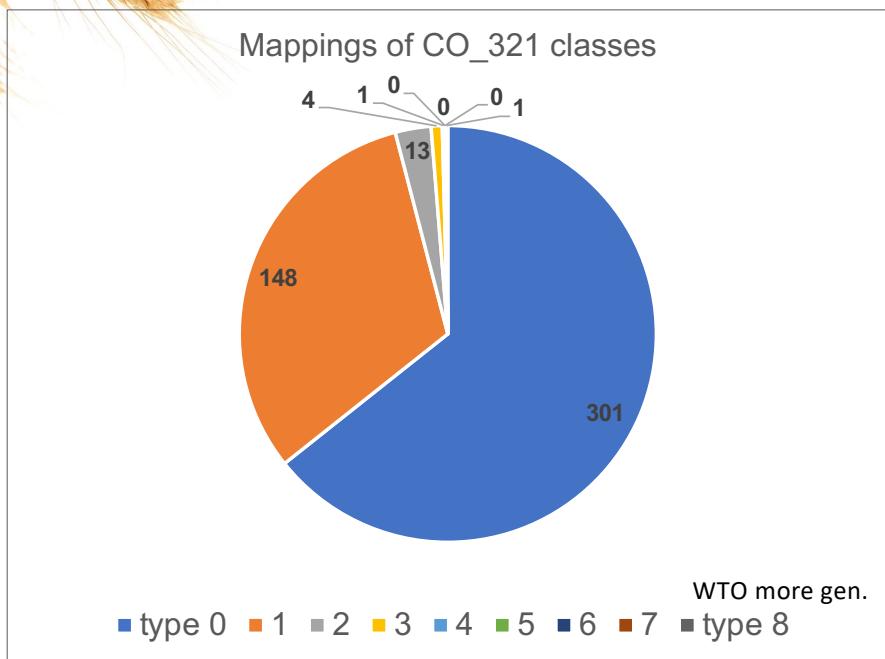
WTO:above-ground dry matter quantity      CO\_321:aboveground biomass      + spec.

# Method for mapping the classes

- Tests with **automatic tools** find obvious string-matched pairs → Manual mapping.
- **We reused the mappings** of WTO and WIPO (ancestor of CO\_321) from a previous project
- We defined general **rules** for class label rewriting
- **Discussion and review** (recorded as Gitlab issues)
  - More than 60 questions that required expertise in wheat breeding
- We recorded **justification** and references to external sources
- **Consistency checking**

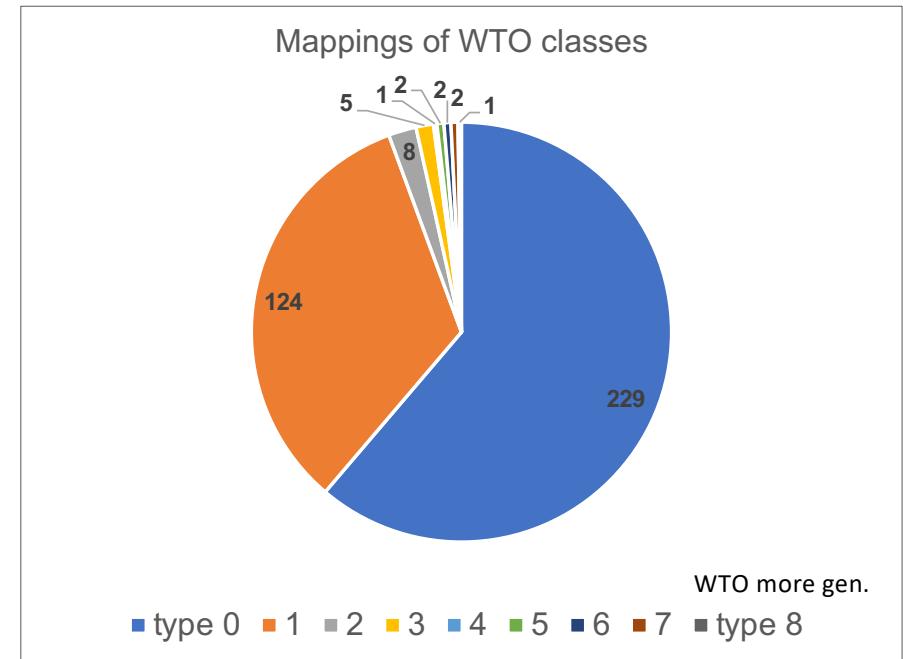


# Manual mapping of the classes of CO\_321 and WTO



## CO\_321 classes without mappings (type 0)

- topics absent from WTO (e.g. canopy traits)
- non trait classes (e.g. date, sample)



## WTO classes without mappings (type 0)

- topics absent from CO\_321 (e.g. root architecture)
- high-level classes (9 levels)

# Class mapping rules

Label mapping follows formal rules  
They formalize the rationale for the mappings  
with an objective of consistency.

## Examples of biotic stress trait class mapping rules

### Score

**resistance to + 'name of disease' (WTO)**  $\leftrightarrow$  'name of disease' + **score (CO\_321)**

Example: WTO: **resistance to** Septoria Leaf Blotch (WTO:554)  $\leftrightarrow$  CO\_321:Septoria **score** on leaf (CO\_321:1010036)

### Severity

**resistance to + 'name of disease' (WTO)**  $\leftrightarrow$  'name of disease'+ **severity (CO\_321)**

Example: WTO: **resistance to** Powdery Mildew (WTO:551)  $\leftrightarrow$  CO\_321:Powdery mildew **severity** (CO\_321:0000939)

### Incidence + part of plant (name of disease)

**resistance to + 'name of disease' (WTO)**  $\leftrightarrow$  'name of disease' + **part of plant +incidence (CO\_321)**

Example: WTO: **resistance to** Fusarium head blight (WTO:0483)  $\leftrightarrow$  CO\_321:Fusarium head blight **spikelet incidence** (CO\_321:0000711)

# Mapping information

WTO_ID:class_ID	WTO class name(PrefLabel)	Verification (Or)	Mappings	Sub ma	CO_321_ID:trait_ID	CO_321 trait name	Reason for alignment	Rules
	subject_label			predicate	predicate	object_id	object_label	comment
<a href="#">WTO_0000063</a>	above-ground dry matter quantity	Maria et al.			CO_321:0000006	Aboveground biomass		
<a href="#">WTO_0000121</a>	admixture	Jacques et al.		0			no appropriate mapping found	
<a href="#">WTO_0000205</a>	aluminium resistance	Jacques et al.		1	1,4 CO_321:0000079	Aluminum tolerance		<a href="#">Rule 4</a>
<a href="#">WTO_0000450</a>	aluminium toxicity	Jacques et al.		1	1,4 CO_321:0000079	Aluminum tolerance		<a href="#">Rule 3</a>
<a href="#">WTO_0000123</a>	amount of damaged starch	Maria et al.		0			no appropriate mapping found	
<a href="#">WTO_0000064</a>	anther color	Maria et al.		0			no appropriate mapping found	
<a href="#">WTO_0000065</a>	anther extrusion	coming from alic		1	1,1 CO_321:0000982	Anther extrusion		
<a href="#">WTO_0000296</a>	aphid resistance	AO1H project		1	1,4 CO_321:0000085	Aphid damage		<a href="#">Rule 2</a>
<a href="#">WTO_0000281</a>	resistance to root lesion nematode			1	1,4 CO_321:0000086	Armyworm damage		<a href="#">Rule 2</a>
<a href="#">WTO_0000451</a>	arsenic toxicity	Jacques et al.		4	CO_321:0500024	Grain Arsenic content	related to plant toxicity without	
<a href="#">WTO_0000066</a>	awn	Maria et al.		0			no appropriate mapping found	
<a href="#">WTO_0000125</a>	awn color	Maria et al.		1	1,1 CO_321:0000960	Awn color		
<a href="#">WTO_0000126</a>	awn length	Maria et al.		1	1,1 CO_321:0000026	Awn length		
<a href="#">WTO_0000452</a>	bacterial leaf blight resistance	Clara et al.		1	1,4 CO_321:0000932	Bacterial leaf blight se	disease would be deduced from	<a href="#">Rule 1, 1.3</a>
<a href="#">WTO_0000067</a>	beginning of anthesis	Maria et al.		1	1,3 CO_321:0000121	Anthesis time	development and has a	
<a href="#">WTO_0000302</a>	billbug resistance	Clara et al.		7	CO_321:0000089	Insect damage		<a href="#">Rule 2</a>
<a href="#">WTO_0000453</a>	boron toxicity	Clara et al.		1	1,4 CO_321:0000678	Boron tolerance		<a href="#">Rule 3</a>

# Les besoins pour des alignements réalisés manuellement

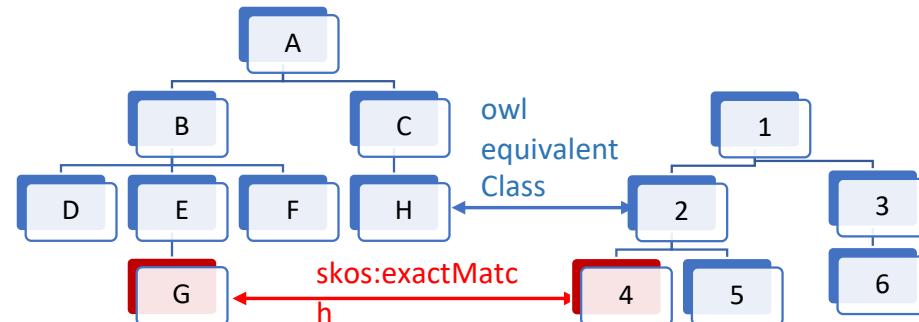
Ce qu'on souhaite représenter :

- Des informations sur les ontologies sources (Version, Identifiant...)
- Autour des alignement :
  - Spécifier des classes ou des instances (via les URI)
  - Une direction entre les deux objets alignés
  - La cardinalité (1:1, 1:n)
  - La nature de l'alignement (exact; broad ; narrow)
  - Des propriétés d'alignement issues de formalismes divers (owl; SKOS)
  - La méthode utilisée (outil, algorithme...)
  - *Restriction d'application du mapping*
  - Raison du mapping

**En gras, les besoins couverts par SSSOM**

# WTO-CO\_321 mapping

Will be shortly published on AgroPortal



**author ID, author label, reviewer ID, reviewer Label, mapping\_Date**  
automatically inferred from the metadata

## Skos mappings

### Equivalent

code 1 → skos:exactMatch

More general/more specific

code 7 → skos:broadMatch

code 8 → skos:narrowMatch

## OWL mappings

### Equivalent

code 1 → owl:equivalentClass

More general/specific

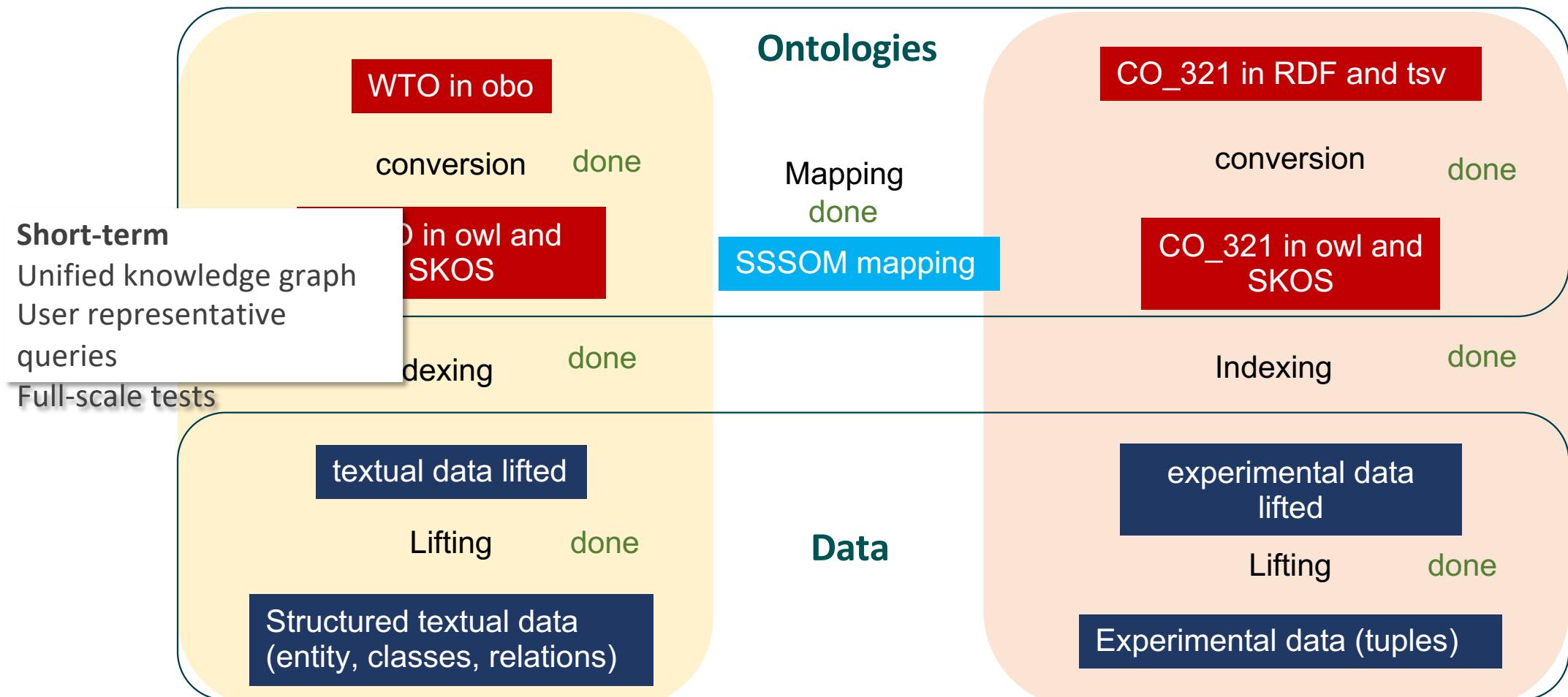
code 7, 8 → rdfs:subClassOf

## Evolution of SSOM wrt our needs

Formalization of the mapping rules (currently free text)  
Specification of the use scope

A	B	C	D	E	F	G	H	I	J
#mapping_set_id:"A COMPLETER :DOI/URI ?"									
#creator_id: "https://ror.org/02kvxyf05"									
#creator_label: "INRAE"									
#curie_map:									
# skos: "http://www.w3.org/2004/02/skos/core"									
#subject_source: "http://opendata.inrae.fr/wto"									
#subject_source_version: "3.1"									
#object_source: "https://cropontology.org/term/CO_321"									
#object_source_version: "Feb 2016"									
#license: "https://www.etalab.gouv.fr/licence-ouverte-open-licence/"									
subject_id	subject_label	predicate_id	object_id	object_label	mapping_justification	author_id	author_label	reviewer_id	reviewer_label
tp://purl.obolibrary.org/obo/WTO_000063	above-ground dry matter quantity	skos:broadMatch	ontology.org/rdf/CO_321:000006	Aboveground biomass	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_0000205	aluminium resistance	skos:closeMatch	ontology.org/rdf/CO_321:000079	Aluminum tolerance	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_0000450	aluminium toxicity	skos:closeMatch	ontology.org/rdf/CO_321:000079	Aluminum tolerance	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_000065	anther extrusion	skos:exactMatch	ontology.org/rdf/CO_321:0000982	Anther extrusion	semav:ManualMappingCuration	https://orcid.org/0000-0003-4805-8 Sophie Aubin Claire Nédellec Thomas Letellier	https://orcid.org/0000-0003-4805-8 Sophie Aubin Claire Nédellec Thomas Letellier	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_0000296	aphid resistance	skos:closeMatch	ontology.org/rdf/CO_321:0000985	Aphid damage	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_0000281	Armyworm resistance	skos:closeMatch	ontology.org/rdf/CO_321:0000986	Armyworm damage	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_0000125	awn color	skos:exactMatch	ontology.org/rdf/CO_321:0000960	Awn color	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_0000126	awn length	skos:exactMatch	ontology.org/rdf/CO_321:0000926	Awn length	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_0000452	bacterial leaf blight resistance	skos:closeMatch	ontology.org/rdf/CO_321:0000932	Bacterial leaf blight severity	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_000067	beginning of anthesis	skos:closeMatch	ontology.org/rdf/CO_321:0000121	Anthesis time	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_0000302	billbug resistance	skos:broadMatch	ontology.org/rdf/CO_321:0000089	Insect damage	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_0000453	boron toxicity	skos:closeMatch	ontology.org/rdf/CO_321:0000678	Boron tolerance	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_0000127	bread making quality	skos:exactMatch	ontology.org/rdf/CO_321:0500001	Baking quality	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_0000127	bread making quality	skos:narrowMatch	ontology.org/rdf/CO_321:1010150	bread score	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_0000127	bread making quality	skos:narrowMatch	ontology.org/rdf/CO_321:1010147	bread volume	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_000286	Cecidomyiidae resistance	skos:broadMatch	ontology.org/rdf/CO_321:0000089	Insect damage	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_000304	cold resistance	skos:closeMatch	ontology.org/rdf/CO_321:1010144	Frost susceptibility	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_000304	cold resistance	skos:closeMatch	ontology.org/rdf/CO_321:0000080	Cold tolerance	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_000305	cold tolerance	skos:exactMatch	ontology.org/rdf/CO_321:0000080	Cold tolerance	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_000070	coleoptile length	skos:exactMatch	ontology.org/rdf/CO_321:0000029	Coleoptile length	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_000071	crop yield	skos:closeMatch	ontology.org/rdf/CO_321:0000013	Grain yield	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_000215	crumb firmness	skos:broadMatch	ontology.org/rdf/CO_321:0000065	Bread crumb structure	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_000072	culm length	skos:exactMatch	ontology.org/rdf/CO_321:0000060	Stem length	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_0000011	development	skos:closeMatch	ontology.org/rdf/CO_321:0000016	Growth stage	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_000308	dough extensibility	skos:narrowMatch	ontology.org/rdf/CO_321:0000212	Dough alveograph value L	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_000309	dough resistance to shearing	skos:relatedMatch	ontology.org/rdf/CO_321:0000213	Dough alveograph value P	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_000310	dough strength	skos:closeMatch	ontology.org/rdf/CO_321:0000145	Dough alveograph P/L	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis
tp://purl.obolibrary.org/obo/WTO_000311	drought tolerance	skos:exactMatch	ontology.org/rdf/CO_321:0000131	Drought tolerance	semav:ManualMappingCuration	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	https://orcid.org/0000-0002-1110-8 Mariya Evtimova Claire Nédellec	Michael Alaux Cyr Jacques Legouis	Michael Alaux Cyr Jacques Legouis

# Final knowledge graph (on going)



# université PARIS-SACLAY



**INRAE MaiAGE**  
Claire Nédellec  
Clara Sauvion  
Robert Bossy  
Louise Deléger  
Mariya  
Borovikova  
Mouhamadou Ba



**INRAE URGI**  
Michaël Alaux  
Cyril Pommier



**INRAE BIOGER**  
Thierry Marcel



**INRAE GDEC**  
Jacques Legouis  
Pierre Sourdille



**INRAE**

SLD 2023  
Claire Nédellec et al.



**DIPSO**  
Sophie Aubin  
Sonia Bravo



**AgroPortal**  
Clément Jonquet

INRAe

