Ontological formalisation of mathematical equations for phenomic data exploitation

Felipe Vargas-Rojas

Supervisors: Danai Symeonidou, Llorenç Cabrera-Bosquet, François Tardieu

Seminar In-OVIVE

September, 2021

MOTIVATION

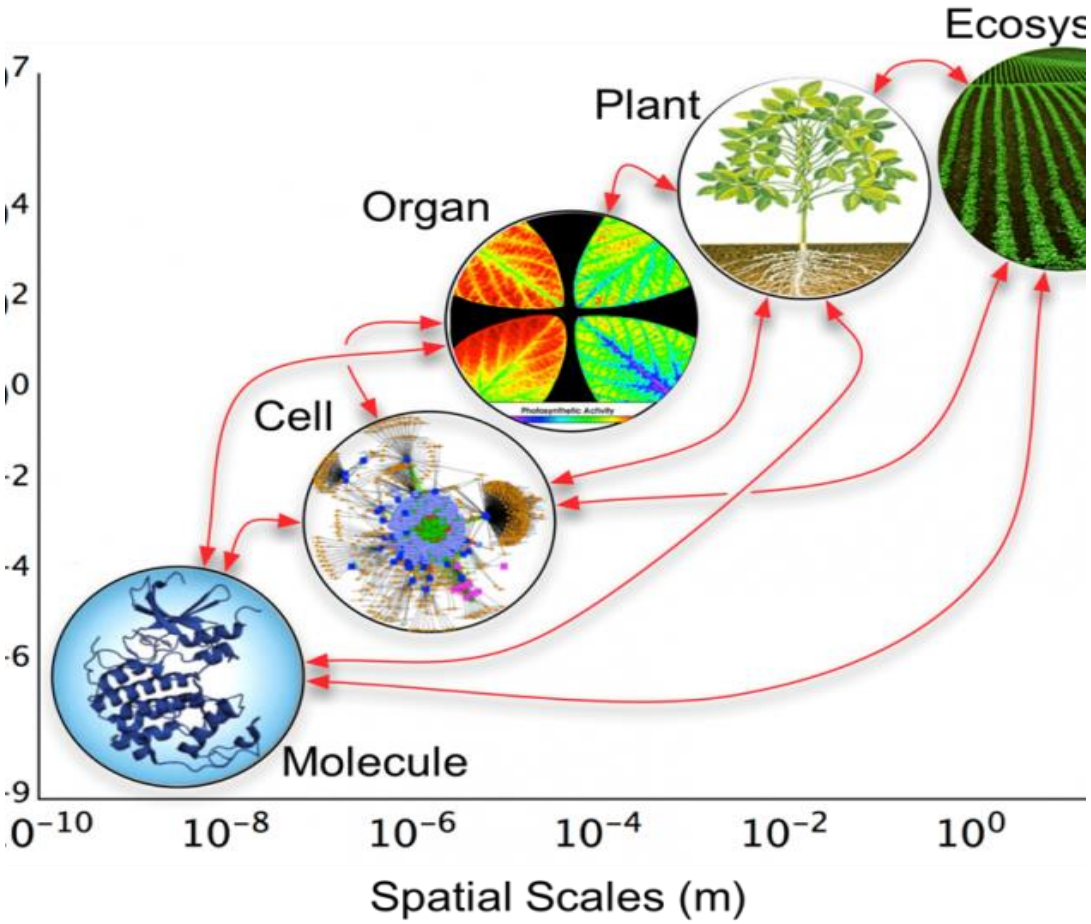
RELATED WORKS

METHODOLOGY

EVALUATION PLAN

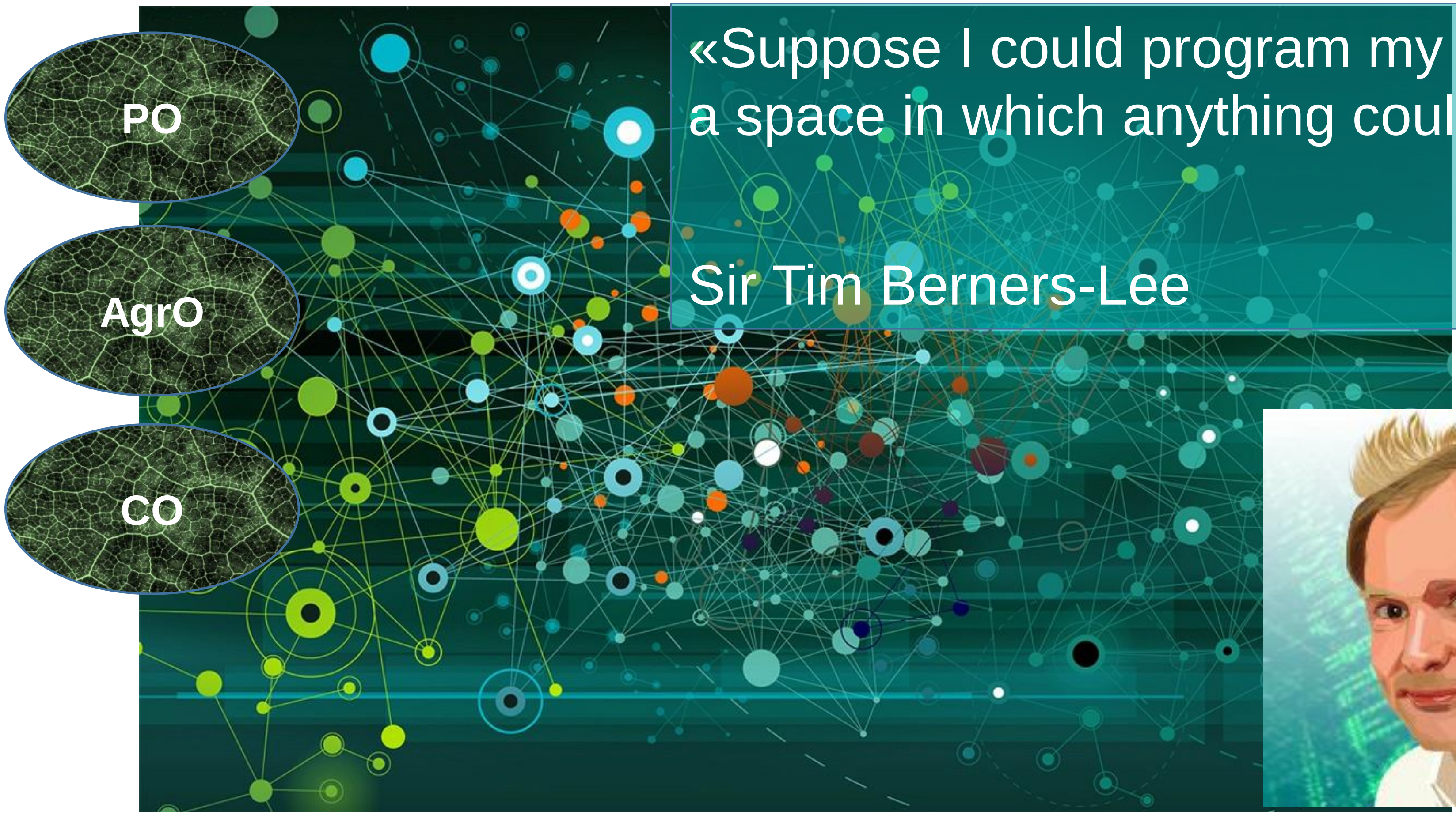
CONCLUSIONS

Multi-scale, multi-source, heterogeneous



(*Marshall-Colon et al. 2017*)

Plant phenomics datasets are in nature



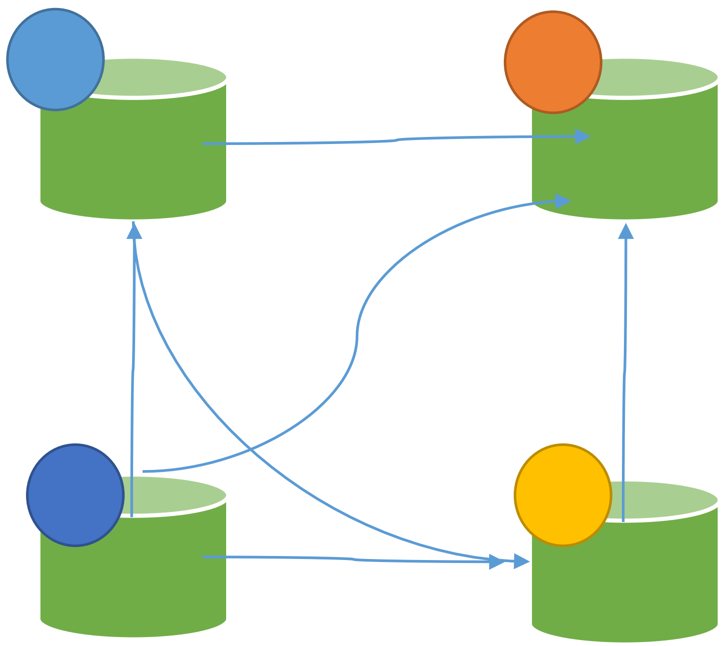
«Suppose I could program my computer to create a space in which anything could be linked to anything »

Sir Tim Berners-Lee

- PO: http://obofoundry.org/ontology/po
- AgrO: http://obofoundry.org/ontology/agro
- CO: https://www.cropontology.org

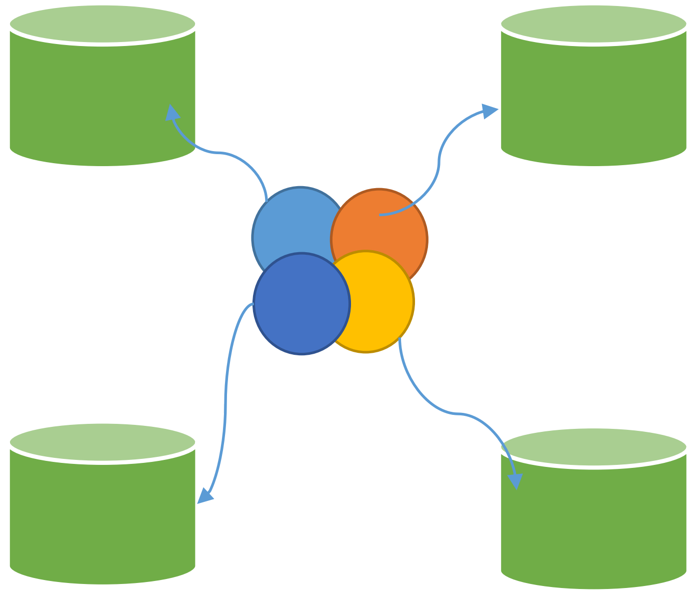
Plant phenomics community has adopted Semantic Web

Traditional systems



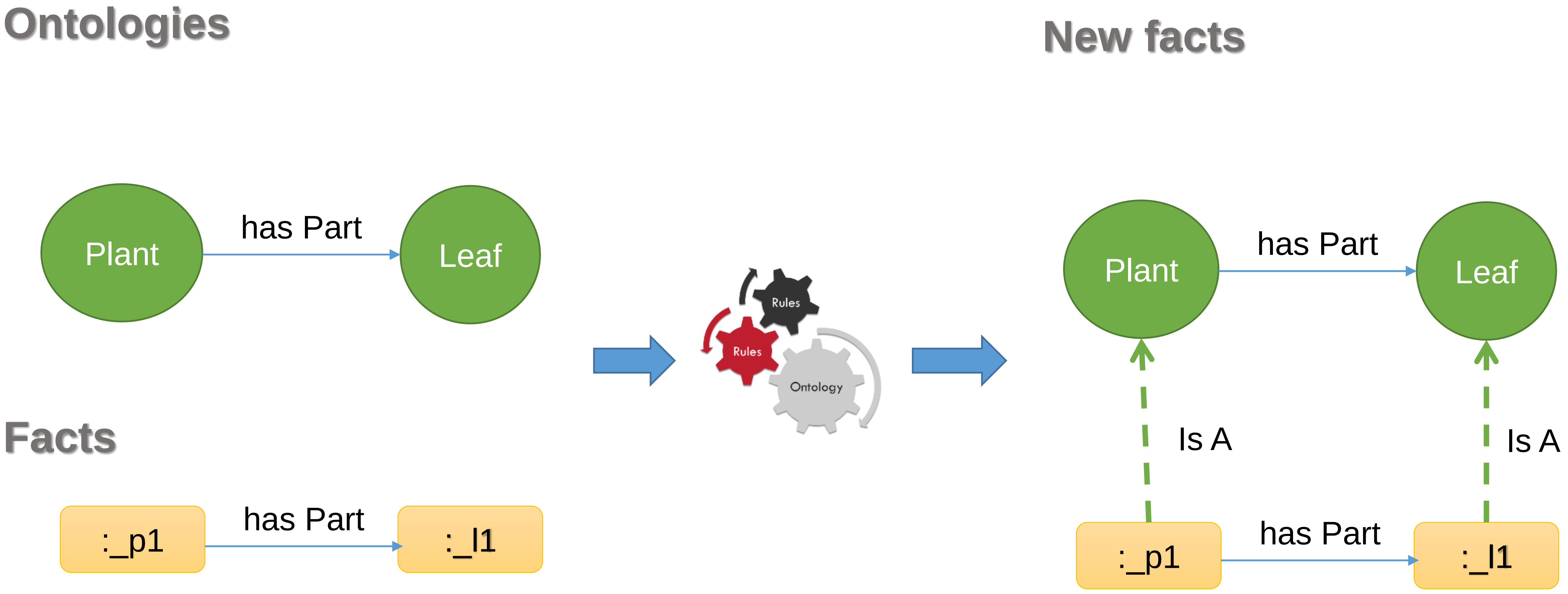
- Several Apps
- Several integrations

(PHIS, www.phis.inra.fr, Neveu et al., 2019)

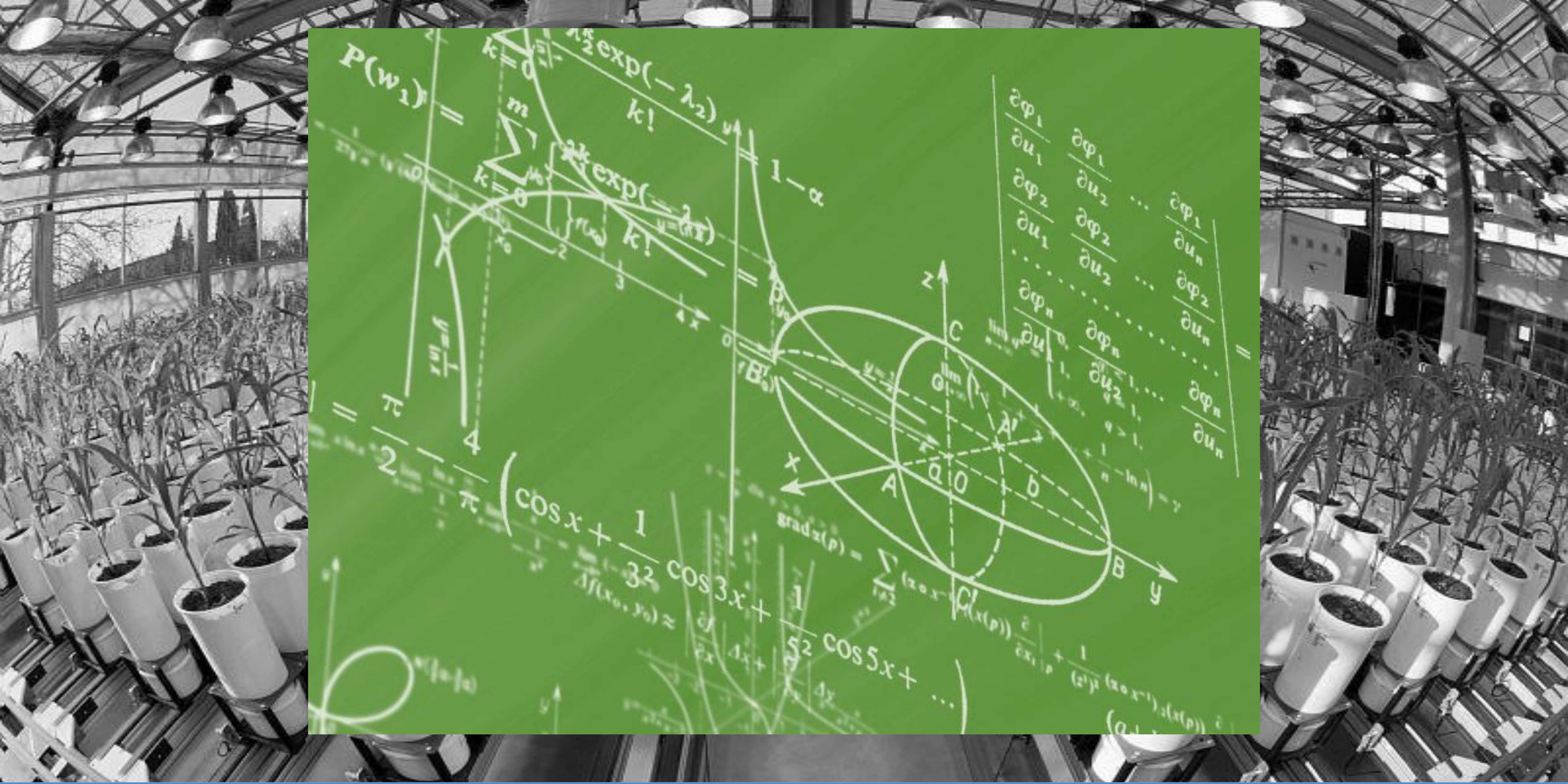


- Several Apps
- 0 integrations

PHIS, an ontology-based information system

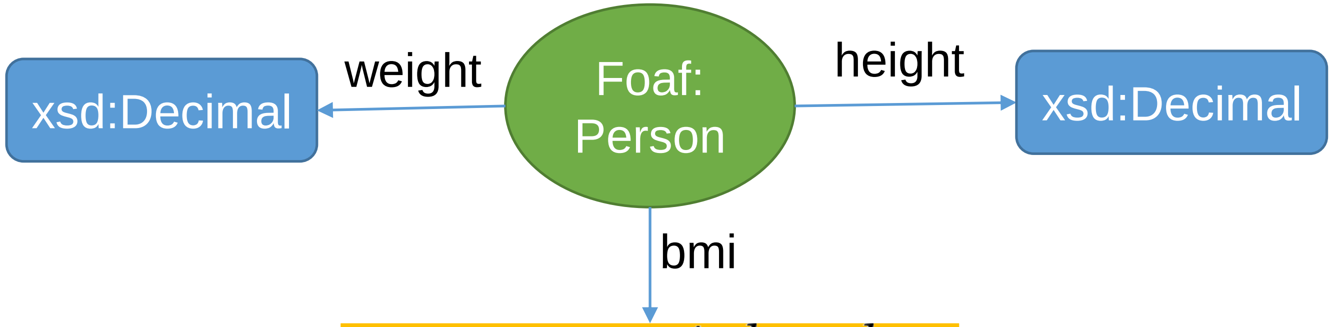


Reasoning services for harmonising data



Numerous equations relate plant phenomic traits

Ontologies



$$BMI = \frac{weight}{height^2}\left(\frac{kg}{m^2}\right)$$

New facts

URI	BMI
:john	24,691358
:peter	31,1111111
:camilo	25,390625
:alexa	21,4532872
:maria	25,4693878

Facts

URI	height (m)	weight (kg)
:john	1,80	80
:peter	1,50	70
:camilo	1,60	65
:alexa	1,70	62
:maria	1,75	78

Following the reasoning logic …..

Different units

height	units
1,80	m
1,70	m
68,89	inches

OM

QUDT

Increasing the complexity

Research questions

- How can mathematical relationships in the form of simple equations be represented and annotated using semantic web approaches?
- How to automate the harmonisation of quantity values in a given context exploiting ontologies and metadata? (e.g. exploit unit ontologies)
- How to use inference services to automate the calculation and assignment of quantity values derived from such mathematical relationships?

Lack of studies exploiting inference services

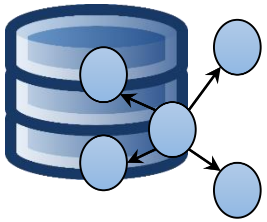
MOTIVATION & PROBLEM

RELATED WORKS

METHODOLOGY

EVALUATION PLAN

CONCLUSIONS



RDF/OWL data

Reasoning

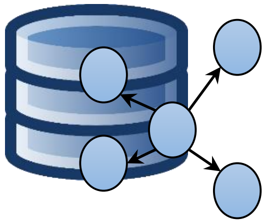
Query

Representing and computing mathematical expressions

Ontology-based Information representation

- Function ontology *(Meester et al. 2016)*
- Unit ontologies *(OM, UO, QUDT)*
- OpenMath + RDF *(Wenzel & Reinhardt, 2012)*

- No info about how to perform the computation



RDF/OWL data

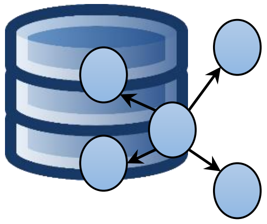
Reasoning

Query

Ontological reasoning

- Extend a query rewriting algorithm *(Bischof et al, 2013),*
- New datatype *(Parsia et al., 2008)*
- Semantic Web Rules (SHACL, SWRL)

- No use of unit ontologies
- Equations structured as text



RDF/OWL data

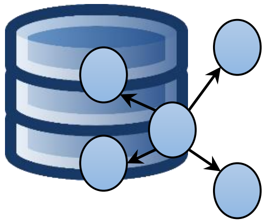
Reasoning

Query

SPARQL extensions

- Query structure different from mathematical objects
- Lack of modularity
- A query is not FAIR

- Extend SPARQL functions
  (Hogan et al., 2020)



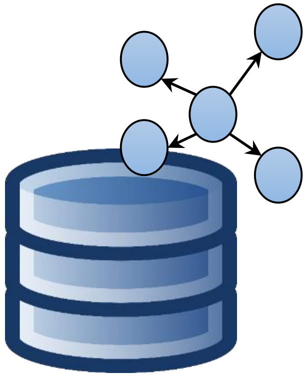
RDF/OWL data

Reasoning

Query

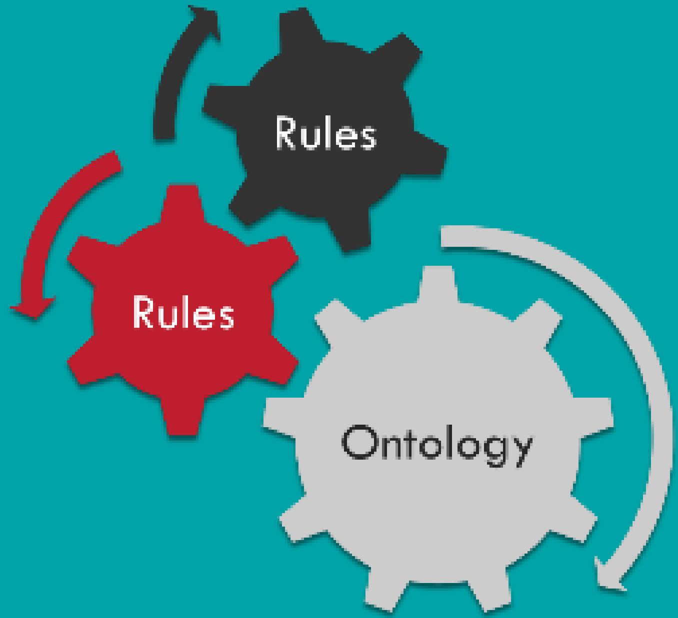
Representation features

1. FAIR equations
2. Based on ontology terms
3. Based on W3c recommendation (e.g. RDF)



Reasoning & computation features

1. Perform equations
2. Derive new facts
3. Exploit domain ontologies (e.g. units)



We propose an ontological framework

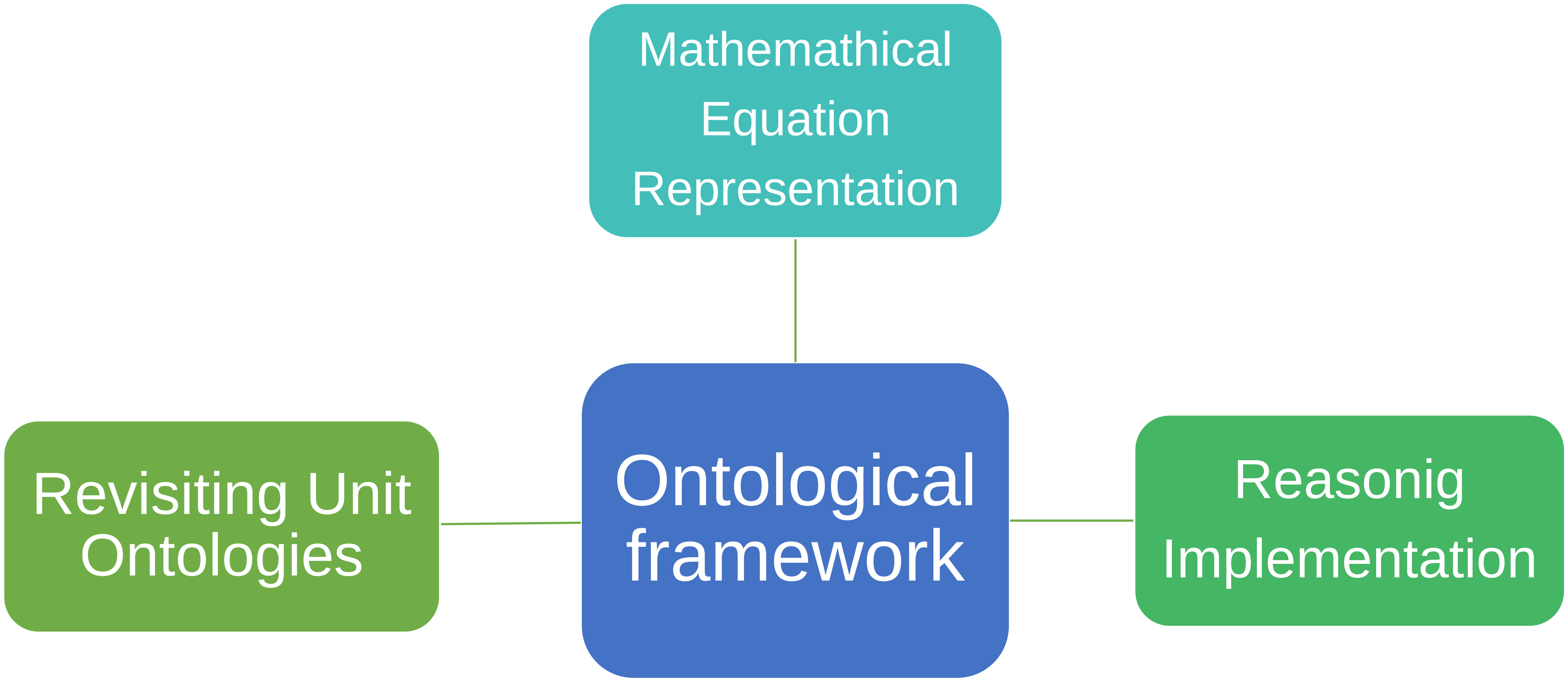
MOTIVATION & PROBLEM

RELATED WORKS

METHODOLOGY

EVALUATION PLAN

CONCLUSIONS



Research methodology and approach

Use metadata from unit ontologies to infer and unify heterogenoeus measurement

Perform Unit Conversion

$$cm \rightarrow \mathrm{m}$$

Normalise the temperature observation depending on contextual data from crops

Thermal time

$$\boldsymbol{Temperature \rightarrow Thermal\,Time}$$

Two case studies

Dimension units

$$1\,m^2 = 10000\,cm^2$$
$$1\,cm^2 = 1\,m^2 \times 10^{-4}$$

Light units

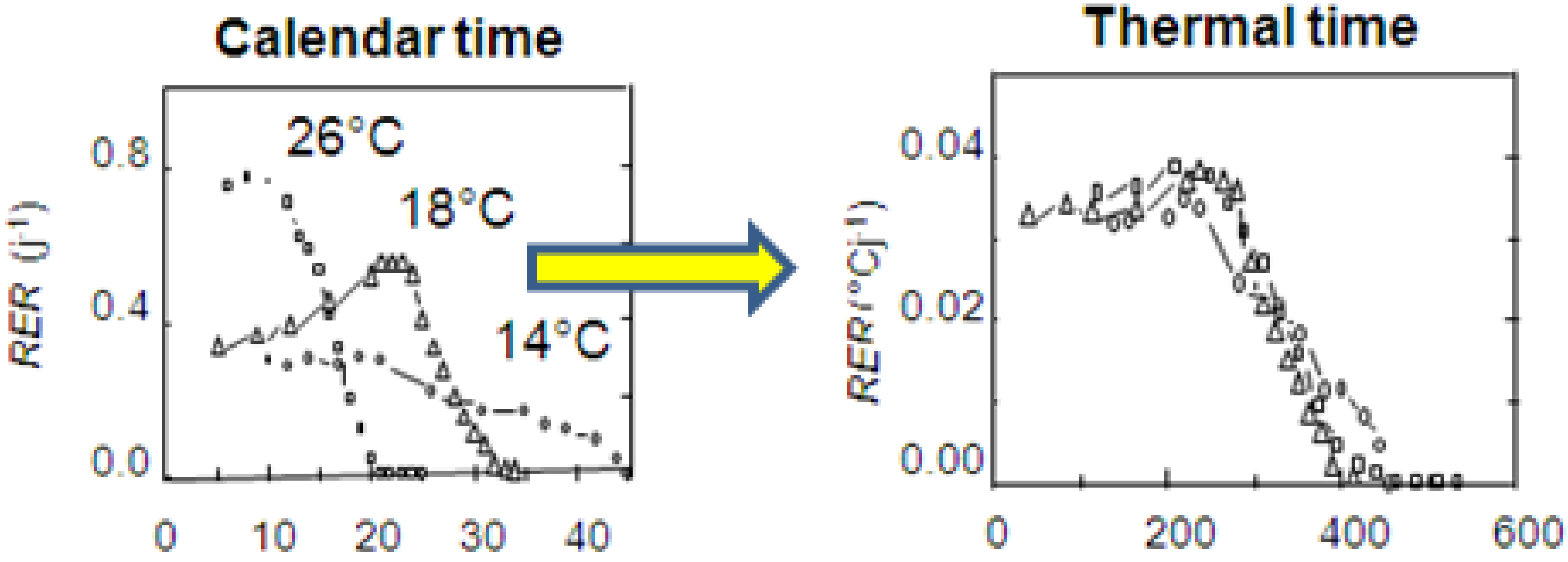
Global Solar Radiation (**Rs**)
Photosynthetically active radiation (**PAR**)

$$R_s(J \cdot cm^{-2}) \rightarrow PAR(\mu mol \cdot m^{-2} s^{-1})$$

Perform unit conversion

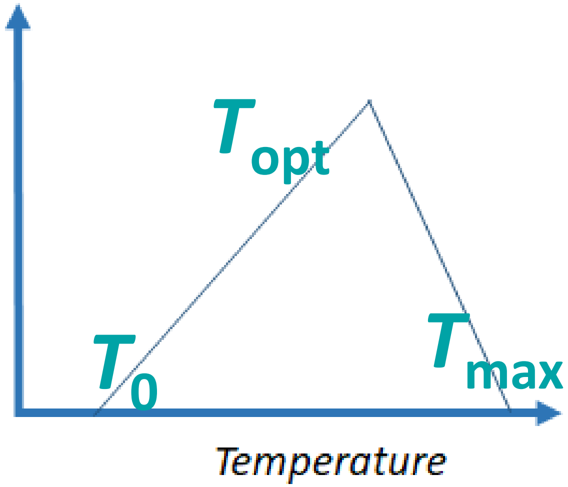
Thermal time definition

(i.e. growing degree units) a process handled by biologists and agronomists used to normalise several temperature-dependent processes such as leaf-progression.



(Granier et Tardieu)

Bilinear Model

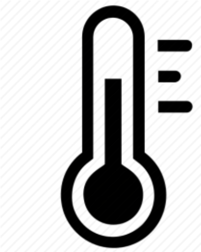


if $T > T_0 \leq T_{opt}$ **then** $T - T_0$

if $T > T_{opt} \leq T_{max}$ **then** $T - T_{max}$

T_{opt}	30.9 °C
T_{max}	37.8 °C
T_0	7 °C

T_{opt}	31.6 °C
T_{max}	38.9 °C
T_0	12 °C



Thermal time equation

MOTIVATION & PROBLEM

RELATED WORKS

METHODOLOGY

EVALUATION PLAN

CONCLUSIONS

Resources for the evaluation plan

Datasets

Unit ontologies

UO

OM

QUDT

Domain ontologies

PO

AgrO

CO

PHIS

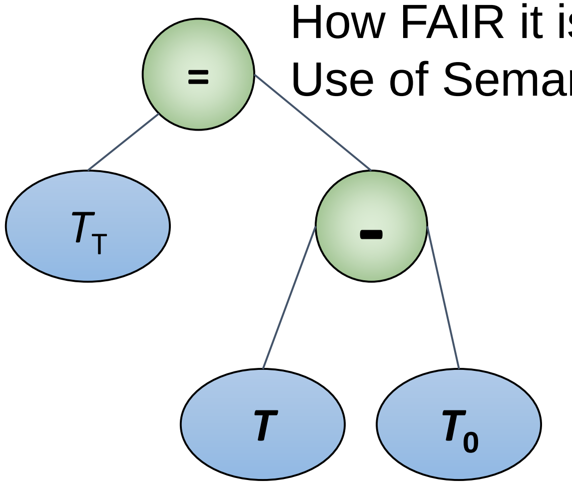
GraphDB

16 GB RAM
1,8 GHz, 8-cores

Assesing the framework

Assessing the equation representation

How FAIR it is?
Use of Semantic Web technologies



Assessing the unit conversion modulo

Which one is better for unit conversion?

UO **OM** **QUDT**

Assessing the nested equation.

Thermal time

$$°F \rightarrow °C$$

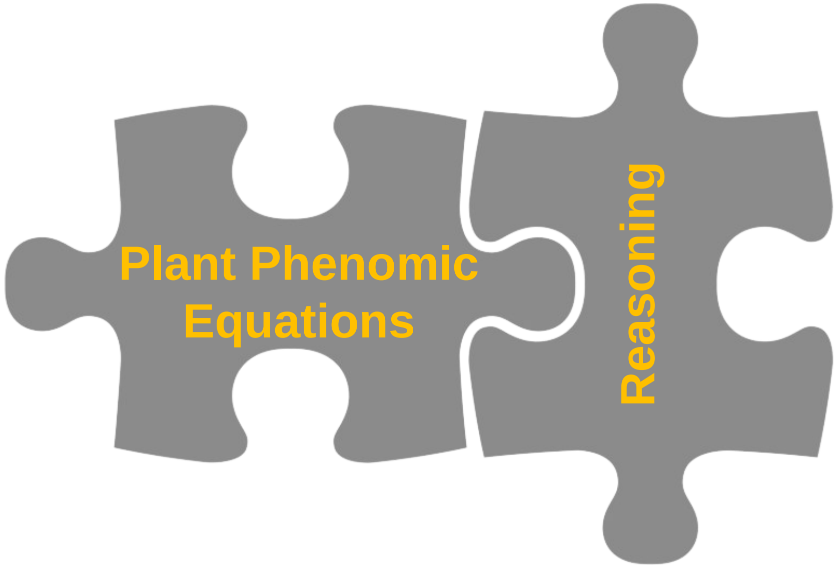
MOTIVATION & PROBLEM

RELATED WORKS

METHODOLOGY

EVALUATION PLAN

CONCLUSIONS



+ FAIR equations
+ Effective use of domain ontologies (units)
+ Enrich existing data with derived values

Therefore, the neglected numerical relationships will be easier to express and accionable

The framework can be used by others domains dealing with numerical attributes and mathematical equations.

Conclusions

THANK YOU

luis-felipe.vargas-rojas@inrae.fr
@LuisVargasRo

Supervisors:

- *Danai Symeonidou*
- *Llorenç Cabrera-Bosquet*
- *François Tardieu*

See also: Vargas-Rojas, F. (2021, June). Ontological Formalisation of Mathematical Equations for Phenomic Data Exploitation. In *European Semantic Web Conference* (pp. 176-185). Springer, Cham.