



**HAL**  
open science

## How the PO2/TransformON and FermentON domain ontologies can help us achieve interoperability and reuse of data on (bio)processes.

Magalie Weber, Ibanescu Liliana, Cufi Julien, Dervaux Stéphane

### ► To cite this version:

Magalie Weber, Ibanescu Liliana, Cufi Julien, Dervaux Stéphane. How the PO2/TransformON and FermentON domain ontologies can help us achieve interoperability and reuse of data on (bio)processes.. BioIndustry 4.0 - Summer Webinar 2024, Aug 2024, online, France. hal-04810088

**HAL Id: hal-04810088**

**<https://hal.inrae.fr/hal-04810088v1>**

Submitted on 28 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License



Summer Webinar 29-30 August 2024

## ➤ How the PO2/TransformON and FermentON domain ontologies can help us achieve interoperability and reuse of data on (bio)processes

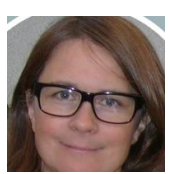
Magalie Weber, UR BIA, Nantes

Liliana Ibanescu, UMR MIA Paris-Saclay, Palaiseau

Julien Cufi, UMR IATE / Dpt TRANSFORM, Montpellier

& Stéphane Dervaux, UMR MIA Paris-Saclay, Palaiseau





Sylvie Clerjon  
IR INRAE  
QUAPA Clermond



Amidou Traoré  
IR INRAE  
QUAPA Clermond



Arnaud Charleroy  
IE INRAE  
MISTEA Montpellier



Nicolas Forquet  
AEDI-Reversaall Lyon-  
Villeurbanne



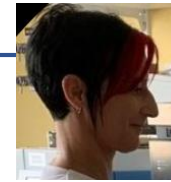
Liliana Ibanescu  
MC AgroParisTech  
MIA Paris-Saclay



Cédric Baudrit  
CR INRAE  
I2M Bordeaux



Christophe Fernandez  
IE INRAE  
I2M Bordeaux



Géraldine Lucchi  
IR INRAE  
CSGA Dijon



Michel Visali  
IR INRAE  
CSGA Dijon



Eric Latrille  
IR INRAE  
LBE Narbonne

Emilie Fernandez  
CDD AI  
LBE Narbonne



Claire Mayer  
DR INRAE  
IATE Montpellier



Hélène Falentin  
IR HDR INRAE  
STLO Rennes



Magalie Weber  
IR INRAE  
BIA Nantes



Stéphane Dervaux  
IE INRAE  
MIA Paris-Saclay

**Food, Bioproducts  
& Waste**  
  
**0138 Mathematics & Digital  
4259 Technologies**



Julien Cufi  
IR INRAE  
Dpt Transform



Mélanie Münch  
IR INRAE  
STLO Rennes



Adrien Réau  
IE INRAE  
IATE Montpellier



Virginie Rossard  
IE INRAE  
LBE Narbonne



Valérie Guillard  
PR UM  
IATE Montpellier



Valérie Gagnaire  
DR INRAE  
STLO Rennes



Anne Thierry  
CR INRAE  
STLO Rennes

 **Microbiology  
& the Food Chain**

 **Plant biology  
& Breeding**

 **Human Nutrition  
& Food Safety**

 **Agroecosystems**



Caroline Pénicaud  
CR HDR INRAE  
SAYFOOD Saclay



Emeline Goussé  
IE INRAE  
STLO Rennes



Leslie Lhomond  
IE INRAE  
IATE Montpellier



Hélène Coussy  
MC HDR UM  
IATE Montpellier



Reine Barbar  
MC Institut Agro  
IATE Montpellier



Florence Valence  
IR INRAE  
CIRM BIA



Alrick Oudot  
CDD IE INRAE  
IATE Montpellier



Patrice Buche  
IR HDR INRAE  
IATE Montpellier



Luc Menut  
AI INRAE  
IATE Montpellier



Hervé Guillemin  
IR INRAE  
UMR PAM  
Dijon/Poligny



Evelyne Aguera  
IE INRAE  
UE PechRouge



Florence Jacques  
IE INRAE  
UE PechRouge



Gaëlle Tanguy  
IE INRAE  
STLO Rennes



Geneviève Gésan-Guiziou  
DR INRAE  
STLO Rennes



Camilo Corrales-Munoz  
IR INRAE-TBI Toulouse

**Numerous collaborators...**



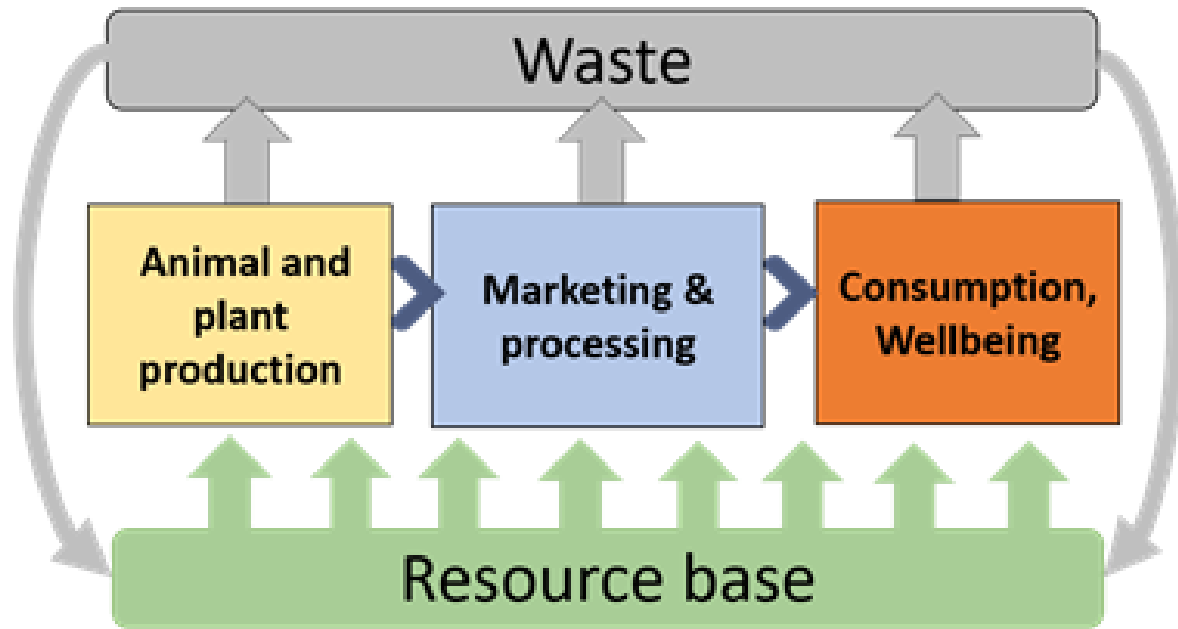
Marc Perez  
IE INRAE  
SPO Montpellier



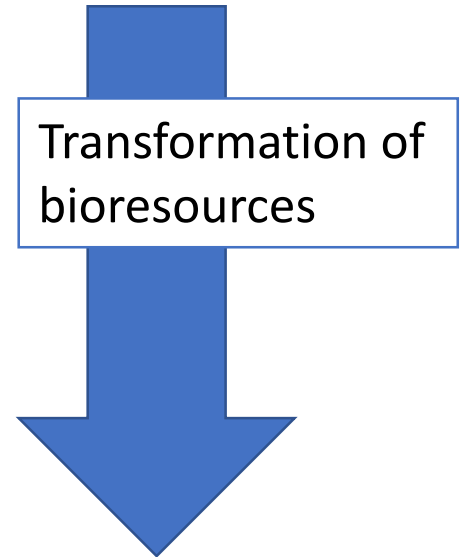
## ➤ Challenges & solution

A tour on PO<sub>2</sub>/TransformON and FermentON

# ➤ Food, Bioproducts and Waste engineering in the context of bioeconomy and global agri-food systems



Socio-economic drivers  
Environmental drivers  
Nutrition & Health



**Towards a circular & sustainable agri-food system**

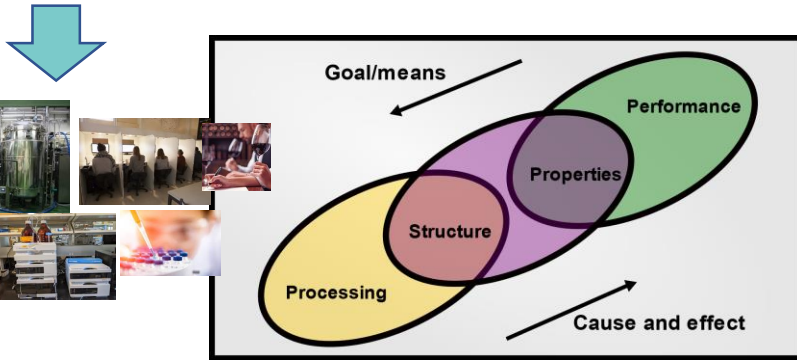
## Research challenges /ontology purposes:

- Designing food quality, from agricultural production to the impact of diets on health and the environment
- Designing biomass construction/deconstruction processes for the production of bioproducts and the use of waste (or residues) generated by human activities.



# ➤ Knowledge discovery in Food and Bioproducts engineering is challenging

Complex research questions on both process performances & product usages...



**Multidisciplinary fields & disciplines :** Chemistry, Physics, Biology, Microbiology, Process engineering, Analytical chemistry, Sensometrics, ....

**Heterogeneous experimental data,** multi-source, multi-format, multi-scale, sometimes incomplete, imprecise...

**How to reason under uncertainty ??**

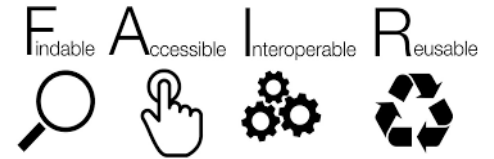
Need for  
STRUCTURING data

STANDARDISING metadata

QUERYING &  
LINKING data (W3C)

SHARE -----> Re-USE

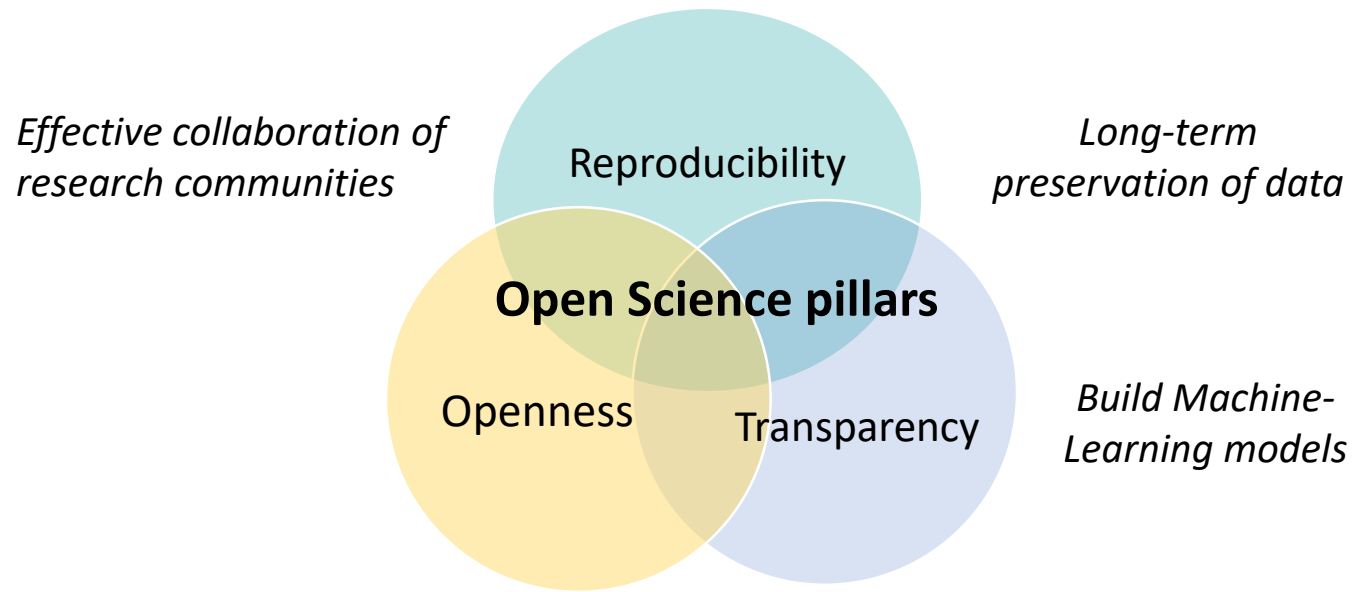
STATISTICAL MODELLING &  
MACHINE LEARNING



Feedback with rules & constraints

**KNOWLEDGE DISCOVERY**

## ➤ How to collect, manage and provide FAIR data (by-design)

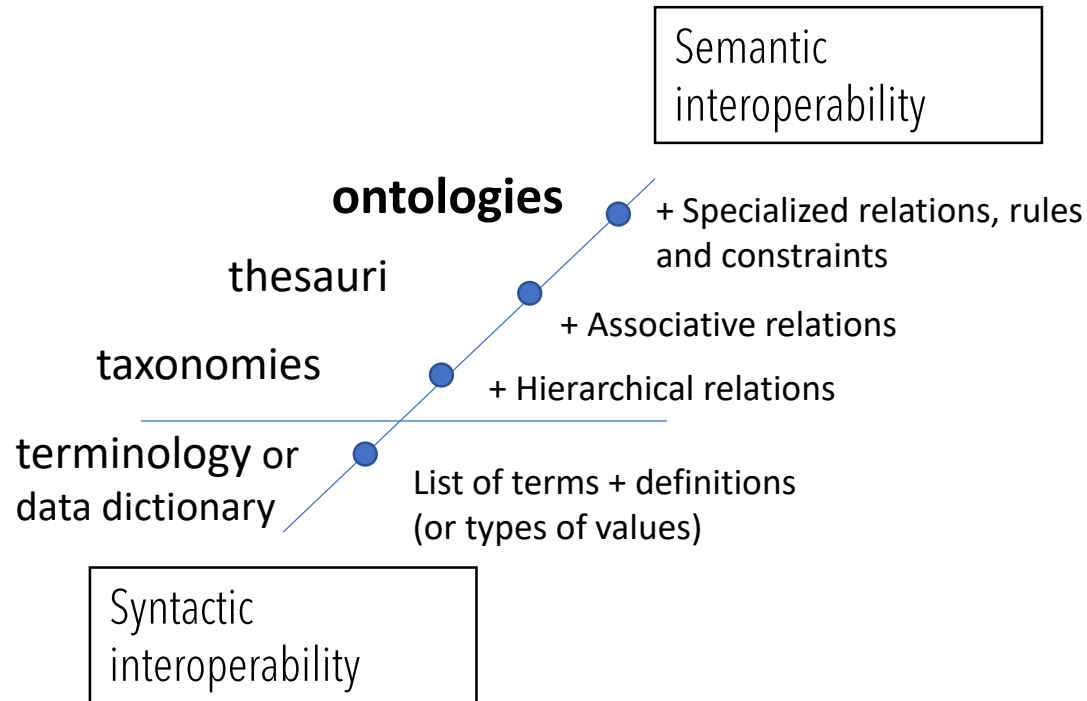


In the context of « Data-driven research » and Open Science, we need to **formalize knowledge** to make it "explicit" and "shareable": **not just data** but also **information about data (metadata)** to share the context and provenance of data



Ontologies are a key solution for interoperability & reusability (I & R)

# ➤ Why are ontologies so useful?



Syntactic interoperability provides a **consistent data format** that can be interpreted by different systems.



Semantic interoperability is **unambiguous**, leaving few room for error or misinterpretation.

Semantically interoperable systems share the **same conceptual understanding of what the underlying data conveys**, thanks to ontologies.



**There is a broad spectrum of ontologies :**

- **Upper-level : foundational ontologies**
  - *BFO, DOLCE, ...*
- **Mid-level : core ontologies, task ontologies**
  - *SOSA/SSN, PROV-O, Time ontology...*
- **Application level : domain ontologies , termino-ontologies**
  - *Vocabulary specific to the field (labels+ definitions)*

➤ *The domain ontologies, PO2/TransformON and FermentON are based on the PO2 core model*

**Ontologies allow us to reach *Semantic interoperability***



# ➤ The Process and Observation Ontology (PO2)

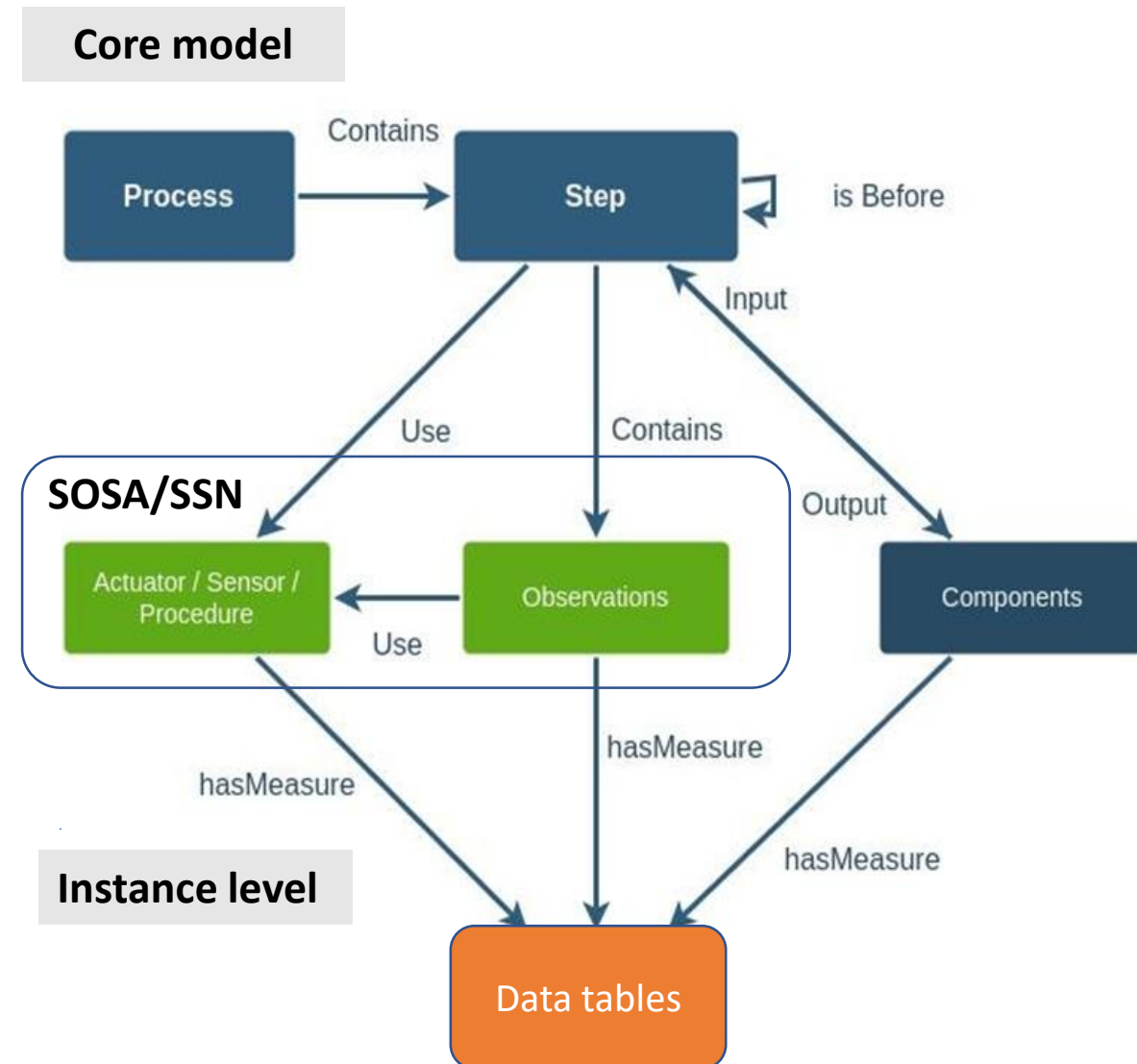
A generic model in OWL,

**reusing SOSA/SSN concepts**

together with other ontologies (BFO, IAO, Time ontology, QUDT), metadata schemas (DCAT, Schema.org), and the Unified Code for Units of Measure (UCUM)

**designed to describe :**

- ❖ Processes as a sequence of steps
  - ❖ Input and output components
  - ❖ Observations with materials and methods
  - ❖ Results described in data tables at the instance level
- **Quantitative and qualitative values in the data tables conform to the I-ADOPT Variable Design Patterns** (<https://catalogue.fair-impact.eu/resources/i-adopt>)



# ➤ A unified vocabulary (TransformON) specializing the PO2 core concepts

## Process Part

### PO2 / Component

#### living organism

- ▶ algae (as living organism)
- ▶ animal (as living organism)
- ▶ bacteria (as living organism)
- ▶ fungi (as living organism)
- ▶ lichen (as living organism)
- ▶ plant (as living organism)

#### substance

- ▶ biochemical constituent
- ▶ feed
- ▶ food
- ▶ non-food substance
- ▶ water (generic)

### PO2 / Process

#### physiological process

- ▶ human physiological process
- ▶ microbial physiological process
- ▶ plant physiological process

#### planned process

- ▶ characterization process
- ▶ transformation process

### PO2 / Step

#### characterization step

#### physiological process step

#### transformation step

- ▶ cleaning
- ▶ handling
- ▶ harvesting
- ▶ packaging
- ▶ pre-processing
- ▶ processing
- ▶ slaughtering
- ▶ storage
- ▶ transport

## Result Part

### PO2 / Attribute

#### calculation outcome

- ▶ experimental data attribute
- ▶ LCIA
- ▶ nutritional score

#### inherent quality

- ▶ identification attribute
- ▶ label or labelling claim
- ▶ physical state
- ▶ status of food name

#### measurement attribute

- ▶ biological attribute
- ▶ mensuration
- ▶ physico-chemical attribute
- ▶ quantity
- ▶ temporality

## Observation Part

### PO2 / Material

- ▶ measuring instrument
- ▶ processing equipment

### PO2 / Method

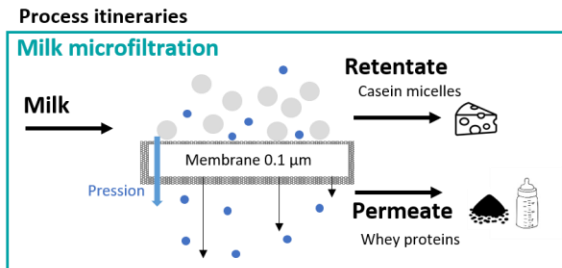
- ▶ analytical method
- ▶ assessment method
- ▶ computation and modelling
- ▶ semi-empiric instrumental method
- ▶ standard operating procedure

### PO2 / Scale

- ▶ measurement scale
- ▶ process scale

# ➤ A variety of use cases and research objectives have been addressed

**Objective: Predicting the performance of the milk microfiltration process**

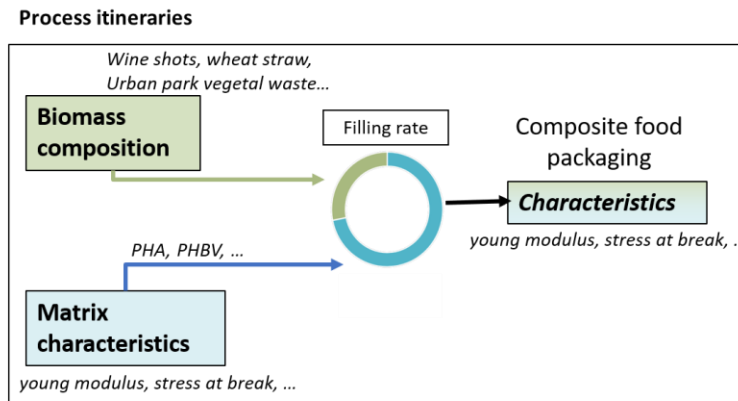


Baudrit et al. 2022

**Problem to be solved:**

- Very heterogeneous and incomplete data
- Existing models limited to a specific range of operating conditions

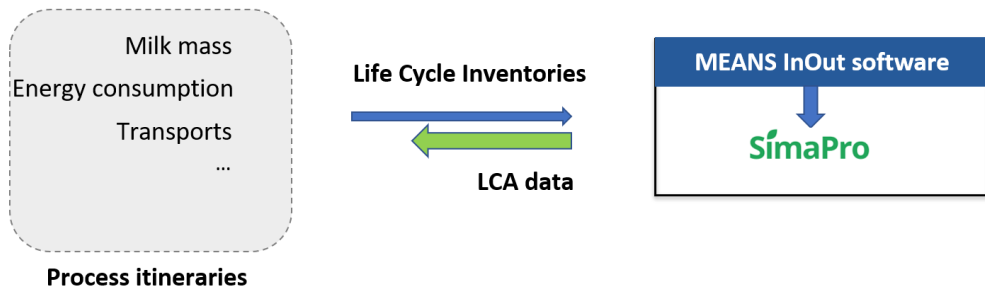
**Objective: optimization of composite food packaging formulation**



Münch et al. 2022

**Problem to be solved : find what biomass and what loading rate to achieve the desired mechanical characteristics**

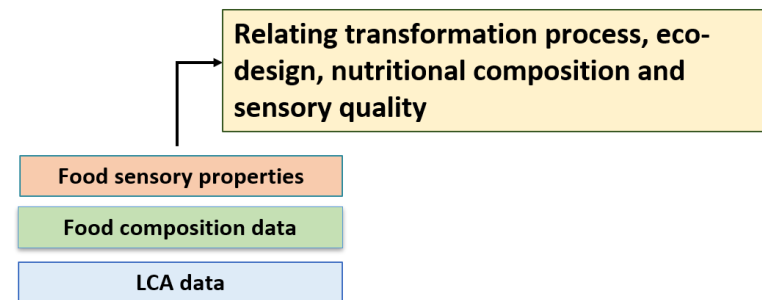
**Objective: to study the environmental impacts of cheeses**



Auberger et al. 2022

**Problem to be solved : providing the process flowchart and inventory data needed for Life Cycle Impact Assessment**

**Objective: to design healthy, sustainable and acceptable food**



Pénicaud et al. 2019

**Problem to be solved: integration of heterogeneous data to provide a global quality score**

# ➤ What's next ?

Ongoing project FermentON

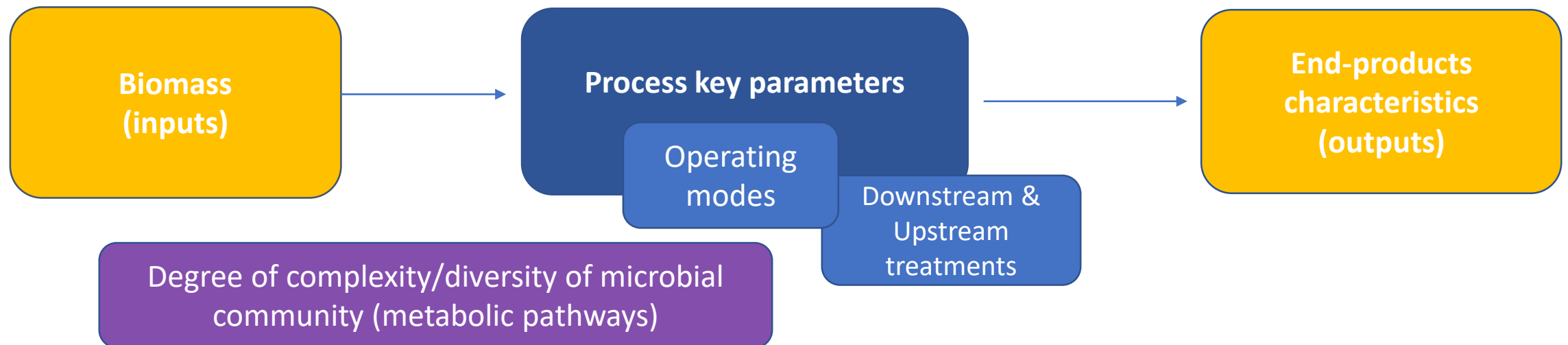


**INRAE**

PO2 presentation on Summer webinar BioIndustry 4.0  
30/08/2024 M. Weber, L. Ibanescu, J. Cufi, S. Dervaux

## ➤ FermentON : extension of PO2/TransformON to bioprocesses

**Scope:** key concepts to describe **fermentation processes** (= using microbial agents) and, more broadly, **biological processes** (= using microbial agents and/or enzymes).



### Examples of questions to be answered :

What combination of substrate/agent/target product and available equipment is needed to choose the cultivation method?

What products result from metabolic activity (intermediates and/or metabolites)? With which enzymes?

# ➤ Use cases in FermentON

## Food fermentation

(i) cheese,  
(ii) mixed dairy and plant-based yogurts,  
and (iii) fermented vegetables

## Wine making

A continuous line from vine to wine & a multiscale study of alcoholic fermentation

## Composite making

Production of PHA/PHBV for composite packaging materials

## Environmental biorefinery & biotech

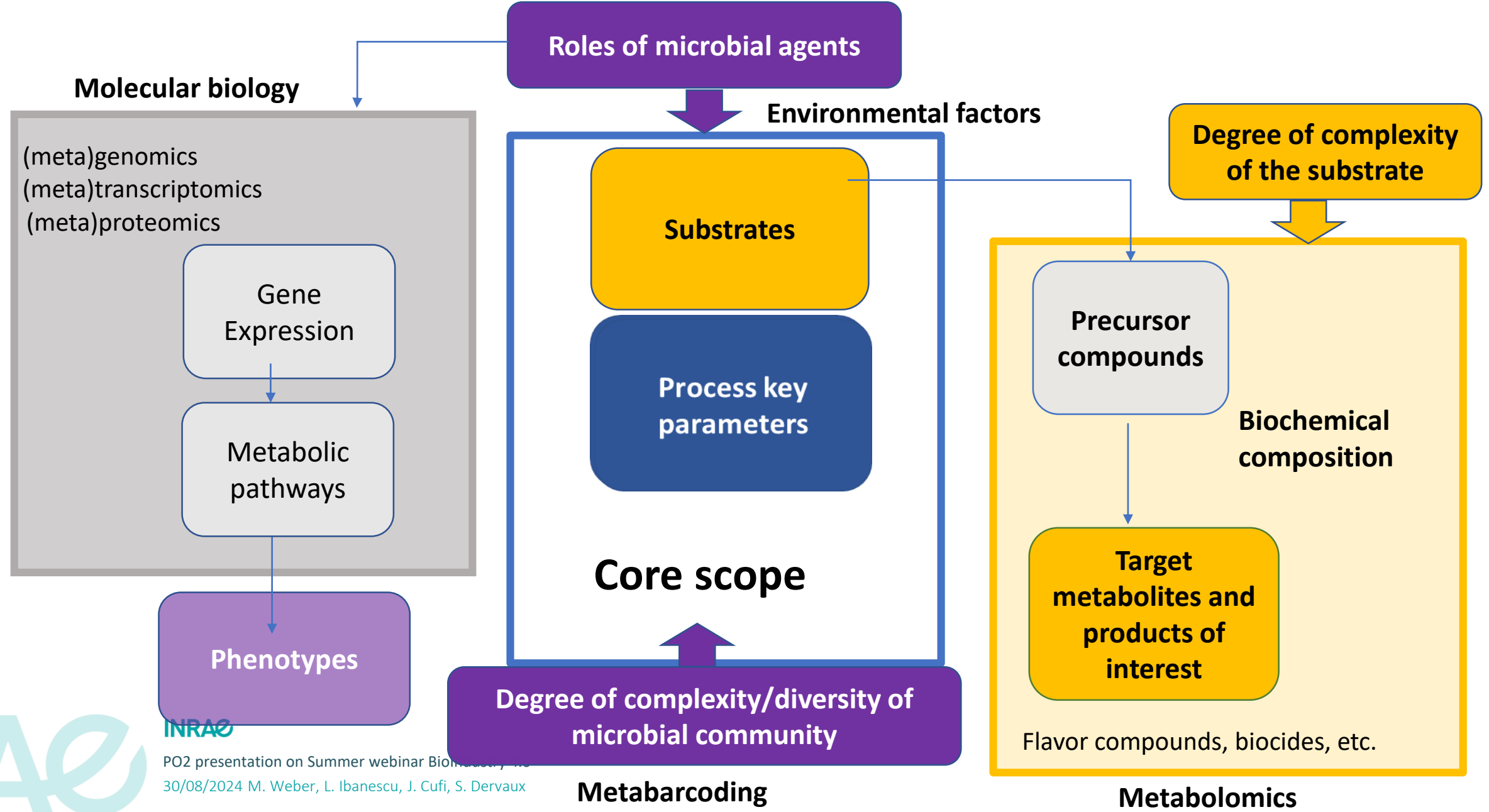
(i) Electrofermentation  
(ii) Design of soft sensors

Fermented food  
& Beverage

Bioproducts,  
matter & energy



# > Scope to be covered





➤ **How to produce FAIR by-design and AI-ready data  
in practice ?**

Deep dive into the ecosystem !



# ➤ How to produce FAIR by-design and AI-ready data in practice ?

*Producing FAIR data requires Ontology driven Information System*

These Information System address main needs like

- Data Entry
- Data Querying

... and relies on a Software Architecture who ensure

- Security and Access Management
- Interoperability



## ➤ Data entry

### *Desktop Application PO<sup>2</sup> Manager*

- Main features

- Data entry with visualization directly through the application
- Import/Export of Excel template files
- Data quality checks
  - Terminological consistency checks
  - Data checks on allowed value range (including dimensions for SI units)
- Semantization process *i.e* remote publication to save and share data

- Benefits

- Ontologies are stored on a centralized database
  - **One shared vision applied to ≠ context**
- Data understanding through visualisation of the studied process
- Standardized vocabulary used but no « Semantics prerequisites » required
- Linked with international W3C standards
  - (e.g., Semantic Sensor Network Ontology, Time Ontology, ...)



# ➤ Data entry

## PO<sup>2</sup> Manager User Interface

The screenshot displays the PO<sup>2</sup> Manager software interface. On the left, a tree view lists various itineraries under the 'ECOBIOCAP project'. The main window shows 'Itinerary 3 -- name: PHBV-CM(10%)' with a 'Product of interest: biobased composite'. A central graph visualizes the process flow. Nodes are color-coded: red circles for components (WS1, CM2, CM, PHBV0, PHBV, CM(dried), PHBV(dried), PHBV-CM(10%), PHBV-CM(10%)<sub>2</sub>, PHBV-CM(10%)<sub>2</sub>, PHBV-CM(10%)<sub>film</sub>), black circles for steps (cutting/milling, drying, extrusion, drying, thermopressing), and black/green circles for observations (cutting/milling, extrusion, thermopressing). A legend at the bottom left explains the node types: Black nodes: steps without observation; Black/Green nodes: steps with observation; Red nodes: compositions; Graph tips: Use shift key + mouse to add a new link; Use del key on selected link or step to remove it.

**Process itinerary is shown as a graph**

Red Circle :  
Component

Black Circle :  
Step

Black + Green:  
Observations





# ➤ Data entry

## PO<sup>2</sup> Manager User Interface : link with underlying ontologies

The screenshot displays the PO<sup>2</sup> Manager software interface. On the left, a tree view shows the project structure under 'Planet-Milling itineraries for a collection of crop byproducts', with 'ultra fine milling (Ball milling step 1.1)' selected. The main window is titled 'Step' and contains fields for 'Step type' (ultra fine milling), 'Step name' (Ball milling step 1.1), 'Date' (2017-06-28), 'Description' (initial water content 2.7%), 'Time', and 'Time duration'. Below these is a 'Materials & Methods' section with a table for parameters:

#	attribute	object	value	unit	comment	
0	Treatment duration		23	h		⊕ ⊖
1	mass	ball of milling machine	48	kg		⊕ ⊖

Below the parameters table are two 'Composition' windows. The first is for 'Flax fibre (1)' with an 'Input' type, containing a table:

#	attribute	component	value	unit	comment	
0	Mass	flax fibre	1000.0	g		⊕ ⊖

The second is for 'Flax fibre (2)' with an 'Output' type, containing a table:

#	attribute	component	value	unit	comment	
0	Mass	Flax fibre	970.0	g		⊕ ⊖

At the bottom left, there are three icons: 'online' (Wi-Fi symbol), 'TransformON' (molecular structure), and 'Planet-Milling ...' (milling machine).

SOSA System  
& SOSA Procedure

Input  
Component

Output  
Component

# ➤ Data entry

## PO<sup>2</sup> Manager User Interface : link with underlying ontologies

The screenshot displays the PO<sup>2</sup> Manager User Interface. On the left, a tree view shows the project structure under 'Planet-Milling itineraries for a collection of crop byproducts', with 'ultra fine milling (Ball milling step 1.1)' selected. The main area is divided into two sections: 'Observation' and 'Materials & Methods'. The 'Observation' section contains form fields for 'Observation type', 'Date', 'Time', 'Time duration', 'Scale', 'Repetition', and 'Objects observed'. The 'Materials & Methods' section shows 'Mastersizer 2000 (Malvern 2000)' and 'granulometry (standard operating condition - SOP4)'. Below these is a table for 'Observation 1' with columns for '#', 'attribute', 'object', 'value', 'unit', and 'comment'. The table contains data for Volume D50, D10, D90, Span, and Specific surface. A second table for 'Observation 2' shows 'Sampling time (min)' and corresponding values for the same attributes. Annotations include 'SOSA ObservationCollection' pointing to the form and 'SOSA Observation' pointing to the table. A purple box highlights the entire data entry area, and a red box highlights the table for Observation 1.

**PO2 Observation**

#	attribute	object	value	unit	comment
	Volume D50		8.101	µm	
	Volume D10		1.504	µm	
	Volume D90		54.353	µm	
	Span		6.657	1	
	Specific surface		1.36	m2/g	

#	Sampling time (min)	Volume D50 (µm)	Volume D10 (µm)	Volume D90 (µm)	Span (1)	Specific surface (m...
0	0	328.124	81.068	1020.171		
1	90	135.194	14.095	725.032	5.259	0.168
2	180	91.252	10.398	486.668	5.219	0.217
3	270	75.803	9.248	364.788	4.69	0.245
4	360	50.76	7.263	217.859	4.149	0.348
5	420	39.139	6.32	169.538	4.17	0.412

## ➤ Data query

*Web application SPO<sup>2</sup>Q*

### • Main Features

- Search engine with two modes :
  - **Simplified** for standard data users via classic web forms
    - No « technical » knowledge required
  - **Advanced** for « power users »
    - Assisted complex query writing
    - Fine-tuning of the SPARQL query
- Automatic conversion : querying data with ≠ units of measure
- Save most frequent queries for re-execution
- Query result export (CSV, JSON, ...)

### • Benefits

- Controlled vocabularies used to query data
- Query public and private project's data
- Cross querying between opened data sources



# ➤ Data query

SPO<sup>2</sup>Q User interface : Simple mode

The screenshot displays the 'Simple PO<sup>2</sup> Query' interface. At the top left is the INRAE logo. The main area is divided into several sections:

- Selected dataset:** A box containing the text 'Planet-Milling\_itineraries\_for\_a\_collection\_of\_crop\_byproducts'.
- Selected process:** A box containing the text 'by process name : doNotFilter'.
- Step type: milling:** A section with three dropdown menus: 'Step Material', 'Input component', and 'Output component'. A green callout box with an arrow points to the 'Step Material' dropdown, containing the text 'Filter on a step in a selected dataset/process'.
- Observation type: Observation:** A section containing a sub-form for 'volume D50'. This sub-form has a 'Show objects' checkbox, a table for filter configuration, and an 'Observation Material' dropdown. The table is as follows:

Filter type	Value	Unit
≤	10	um

A green callout box with an arrow points to the 'volume D50' sub-form, containing the text 'Want to restrict to the set of observations associated with Volume D50 ≤ 10 μm'. A white callout box with an arrow points to the table, containing the text 'Add an observation property result to filter on'.

At the bottom of the interface, there are two buttons: 'Add an observation' (dark blue) and 'Execute form' (dark red).

# ➤ Data query

## SPO<sup>2</sup>Q User interface : Advanced mode

**INRAE** Simple PO<sup>2</sup> Query Form SPARQL Results

### SPARQL Editor

Planet-Milling\_itineraries\_for\_a\_collection\_of\_crop\_byproducts 1000 lines

```
1
2 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
3 PREFIX PO2: <http://opendata.inrae.fr/PO2/core/>
4 PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
5 PREFIX time: <http://www.w3.org/2006/time#>
6 PREFIX sesame: <http://www.openrdf.org/schema/sesame#>
7 PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
8 PREFIX sosa: <http://www.w3.org/ns/sosa/>
9 PREFIX ssn: <http://www.w3.org/ns/ssn/>
10 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
11 PREFIX core: <http://opendata.inrae.fr/PO2/core/>
12 PREFIX qudt: <http://qudt.org/schema/qudt/>
13 PREFIX schema: <http://schema.org/>
14
15
16 SELECT ?project_name_1 ?process_1 ?sampleNameLabel_1 ?process_sample_code_1 ?itineraryLabel ?step_1 ?observation_11 ?property_111
   ?valueOrigin_prop_111 ?unitOrigin_prop_111 WHERE {
17
18     #####Step 1 filtered by type: milling
19     ?itinerary rdf:type <http://opendata.inrae.fr/PO2/core/Transformation_Process>.
20     ?itinerary PO2:hasForStep ?stepURI_1.
21     ?itinerary sesame:directType ?direct_process_1.
22     ?itinerary skos:prefLabel ?itineraryLabel.
23     ?direct_process_1 skos:prefLabel ?process_1.
24     FILTER (langMatches( lang(?process_1), "en" ) || langMatches( lang(?process_1), "" )).
25     optional {
26         ?direct_process_1 PO2:sampleCode ?process_sample_code_1 .
27     }
28     optional {
29         ?direct_process_1 PO2:sampleName ?process_sample_name_1 .
30         ?process_sample_name_1 skos:prefLabel ?sampleNameLabel_1.
31         FILTER (langMatches( lang(?sampleNameLabel_1), "en" ) || langMatches( lang(?sampleNameLabel_1), "" )).
32     }
33     optional {
34         ?direct_process_1 PO2:projectName ?project_name_1 .
```

### SPARQL Templates

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX PO2: <http://opendata.inrae.fr/PO2/core/>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX time: <http://www.w3.org/2006/time#>
PREFIX sesame: <http://www.openrdf.org/schema/sesame#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX sosa: <http://www.w3.org/ns/sosa/>
PREFIX ssn: <http://www.w3.org/ns/ssn/>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX core: <http://opendata.inrae.fr/PO2/core/>
PREFIX qudt: <http://qudt.org/schema/qudt/>
PREFIX schema: <http://schema.org/>

SELECT ?project_name_1 ?process_1 ?sampleNameLabel_1 ?process_sample_code_1 ?itineraryLabel ?st-
[...]
}
```

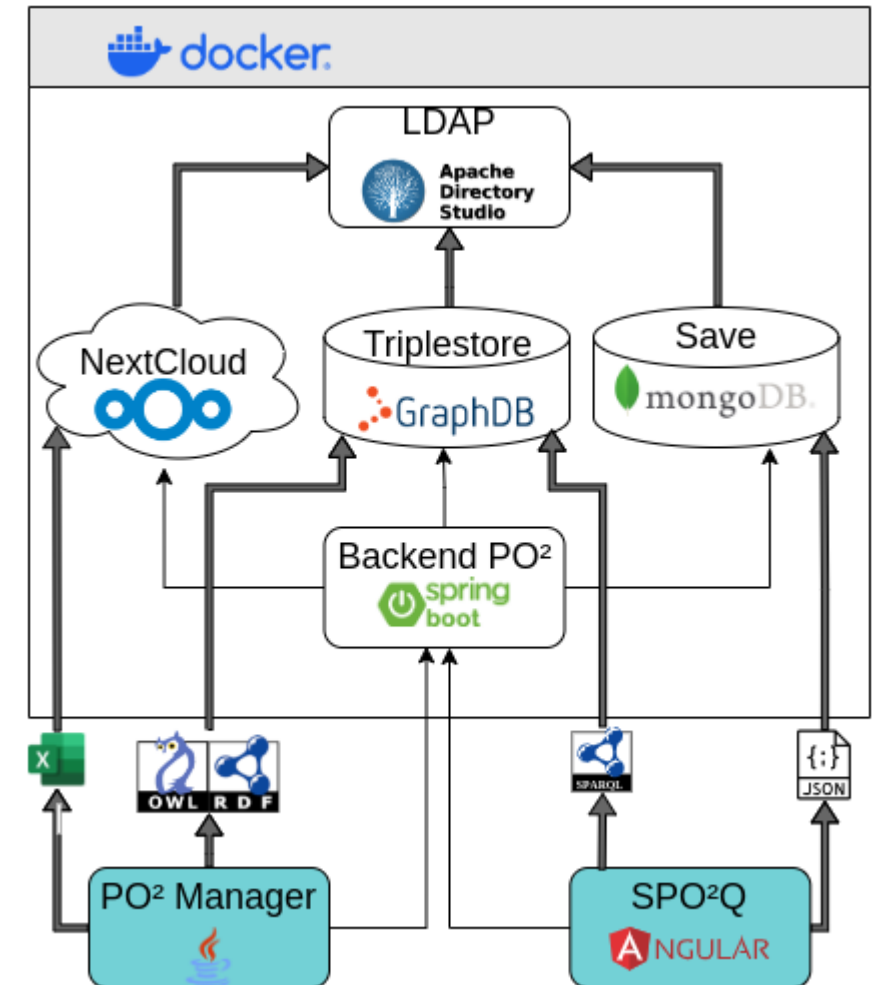
#### Step 1 filtered by type: milling

```
?itinerary rdf:type <http://opendata.inrae.fr/PO2/core/Transformation_Process>.
?itinerary PO2:hasForStep ?stepURI_1.
?itinerary sesame:directType ?direct_process_1.
?itinerary skos:prefLabel ?itineraryLabel.
?direct_process_1 skos:prefLabel ?process_1.
FILTER (langMatches( lang(?process_1), "en" ) || langMatches( lang(?process_1), "" )).
optional {
    ?direct_process_1 PO2:sampleCode ?process_sample_code_1 .
}
optional {
    ?direct_process_1 PO2:sampleName ?process_sample_name_1 .
    ?process_sample_name_1 skos:prefLabel ?sampleNameLabel_1.
    FILTER (langMatches( lang(?sampleNameLabel_1), "en" ) || langMatches( lang(?sampleNameLabel_1), "" )).
}
optional {
    ?direct_process_1 PO2:projectName ?project_name_1 .
```



## ➤ Security and Access Management

- **Data is FAIR but not necessarily immediately open**
  - ➔ Possibility of private project creation
- Account management internal LDAP shared with PO2 Manager and SPO<sup>2</sup>Q
  - One single account per user
  - Flexibility in account creation (not necessarily institutional)
- Management of read/write access by users
- Querying private and public data





## ➤ What's next ?

- Collaborative ontology construction
  - Share and improve domain ontology by integrating scientific point of views
  - Workflow construction based on Git/Gitlab
- Develop and enforce AI applications
  - Improve quality checks and user's feedback
- Large dataset management
  - Be able to perform semantic queries on large dataset (Ex : Long times series)
    - Proof of Concept « Best of two worlds » in progress
      - BigData technology stack for performance
      - Semantic technology stack for knowledge representation
    - SANSA <https://sansa-stack.net/>
- Interface improvement (UX design)





**In a nutshell**

# Summary



Data of interest

STRUCTURING data

STANDARDISING metadata

QUERYING & LINKING data (W3C)

AI-READY Data

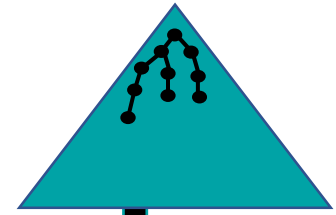
Re-USE

SHARE (PUBLISH)

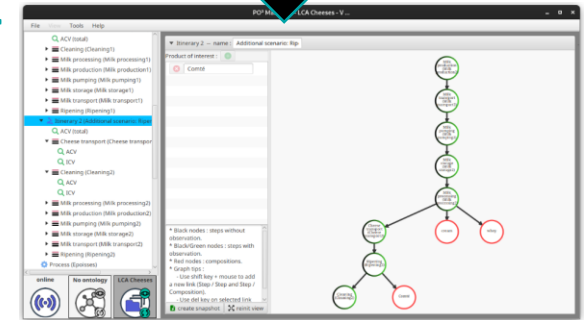
STATISTICAL MODELLING & MACHINE LEARNING

Open data publication option

PO2 core and domain ontologies



PO2 Manager



SPO2Q



Graph DB



FAIR by-design Data



**INRAE KNOWLEDGE DISCOVERY**

PO2 presentation on Summer webinar BioIndustry 4.0  
30/08/2024 M. Weber, L. Ibanescu, J. Cufi, S. Dervaux

# ➤ PO2/TransformON : An ontology for modelling biomass transformation process and characterization of food, feed, bioproducts & biowaste

npj | science of food

Explore content ▾ About the journal ▾ Publish with us ▾

[nature](#) > [npj science of food](#) > [articles](#) > article

Article | [Open Access](#) | [Published: 04 September 2023](#)

## PO2/TransformON, an ontology for data integration on food, feed, bioproducts and biowaste engineering

[Magalie Weber](#) , [Patrice Buche](#), [Liliana Ibanescu](#), [Stéphane Dervaux](#), [Hervé Guillemin](#), [Julien Cufi](#), [Michel Visalli](#), [Elisabeth Guichard](#) & [Caroline Pénicaud](#)



INRAE

PO2 presentation on Summer webinar BioIndustry 4.0  
30/08/2024 M. Weber, L. Ibanescu, J. Cufi, S. Dervaux

  
MINISTÈRE  
DE L'ENSEIGNEMENT  
SUPÉRIEUR  
ET DE LA RECHERCHE  
*Liberté  
Égalité  
Fraternité*

 Ouvrir  
la science !



*2023 Open Science Research Data Award from the French Minister for Higher Education, Research and Innovation*



Documentation and support:  
<https://quantum.mia-ps.inrae.fr/PO2/>



Magalie Weber  
IR INRAE  
BIA Nantes



Julien Cufi  
IR INRAE  
Dpt Transform



Liliana Ibanescu  
MC AgroParisTech  
MIA Paris-Saclay



Stéphane Dervaux  
IE INRAE  
MIA Paris-Saclay



Mélanie Münch  
IR INRAE  
STLO Rennes

STLO



## Core team for IT development & knowledge engineering



Alrick Oudot  
IE INRAE  
IATE Montpellier



Patrice Buche  
IR HDR NRAE  
IATE Montpellier



Luc Menut  
AI INRAE  
IATE Montpellier



Hervé Guillemain  
IR INRAE  
PAM Dijon/Poligny



& PLASTIC Platform



## ➤ Softwares

- Dervaux et al. 2023: PO2 Manager, an annotation tool to edit biomass transformation and characterization itineraries using the Process and Observation (PO2) Ontology. [⟨swh:1:dir:b9983f579ca087d946a9c3e4818e4507de395937⟩](https://doi.org/10.26434/chemrxiv-2023-b9983). [⟨hal-04313202⟩](https://hal.archives-ouvertes.fr/hal-04313202)
- Dervaux et al. 2024: PO2 Engine, an application programming interface associated with the Process and Observation (PO2) software ecosystem. [⟨swh:1:dir:076aa8ecc70160ac3c9c333570c8d495d7e2ff19⟩](https://doi.org/10.26434/chemrxiv-2024-076aa). [⟨hal-04387669⟩](https://hal.archives-ouvertes.fr/hal-04387669)
- Dervaux et al. 2024: Simple PO2 Query (SPO2Q), a querying tool to retrieve biomass transformation and characterization itineraries using the Process and Observation (PO2) Ontology. [⟨swh:1:dir:360ce0db73073210c9e95ee9ee44cb50fc31d9bd⟩](https://doi.org/10.26434/chemrxiv-2024-360ce). [⟨hal-04501660⟩](https://hal.archives-ouvertes.fr/hal-04501660)



# Reused standards and ontologies

- SOSA/SSN <https://www.w3.org/TR/vocab-ssn/>
- BFO <https://obofoundry.org/ontology/bfo.html>
- IAO <https://obofoundry.org/ontology/iao.html>
- Time Ontology <https://www.w3.org/TR/vocab-owl-time-rel/>
- QUDT <https://qudt.org/>
- UCUM <https://ucum.org/>
- Schema.org <https://schema.org/>
- DCAT <https://www.w3.org/TR/vocab-dcat/>
- I-ADOPT <https://i-adopt.github.io/ontology/>





# Bibliographic references

- Ibanescu et al. 2016: *PO2 - A Process and Observation Ontology in Food Science. Application to Dairy Gels*. In: Metadata and Semantics Research. *Communications in Computer and Information Science*, vol 672. Springer [Communication MTSR2016]. [10.1007/978-3-319-49157-8\\_13](https://doi.org/10.1007/978-3-319-49157-8_13). [hal-01418097](https://hal.archives-ouvertes.fr/hal-01418097)
- Pénicaud et al. 2019: Relating transformation process, eco-design, composition and sensory quality in cheeses using PO2 ontology. *International Dairy Journal*, 92, pp.1-10. [10.1016/j.idairyj.2019.01.003](https://doi.org/10.1016/j.idairyj.2019.01.003). [hal-02043246](https://hal.archives-ouvertes.fr/hal-02043246)
- Baudrit et al. 2022: Decision support tool for the agri-food sector using data annotated by ontology and Bayesian network: a proof of concept applied to milk microfiltration. *International Journal of Agricultural and Environmental Information Systems*, 2022, 13 (1), [10.4018/IJAEIS.309136](https://doi.org/10.4018/IJAEIS.309136). [hal-03738973](https://hal.archives-ouvertes.fr/hal-03738973)
- Auberger et al. 2022 Adapting MEANS-InOut LCA software to food engineering, in relation to the PO<sup>2</sup> food ontology and PO<sup>2</sup>-BaGaTel food engineering database. *13th International Conference on Life Cycle Assessment of Food 2022 (LCA Foods 2022)*, Oct 2022, Lima, Peru. [hal-03837121](https://hal.archives-ouvertes.fr/hal-03837121)
- Münch et al. 2022a: Combining ontology and probabilistic models for the design of bio-based product transformation processes. *Expert Systems with Applications*, Elsevier, 203, pp.117406. [10.1016/j.eswa.2022.117406](https://doi.org/10.1016/j.eswa.2022.117406). [hal-03662183](https://hal.archives-ouvertes.fr/hal-03662183)
- Münch et al. 2022b - Formalising contextual expert knowledge for causal discovery in linked knowledge graphs about transformation processes: application to processing of bio-composites for food packaging. *International Journal of Metadata, Semantics and Ontologies*, 16 (1), pp.1-15. [10.1504/IJMSO.2022.131129](https://doi.org/10.1504/IJMSO.2022.131129). [hal-04115029](https://hal.archives-ouvertes.fr/hal-04115029)
- Münch et al. 2022c: Biocomposites from poly(3-hydroxybutyrate-co-3-hydroxyvalerate) and lignocellulosic fillers: Processes stored in data warehouse structured by an ontology. *Data in Brief*, 42, pp.108191. [10.1016/j.dib.2022.108191](https://doi.org/10.1016/j.dib.2022.108191). [hal-03650668v2](https://hal.archives-ouvertes.fr/hal-03650668v2)
- Weber et al. 2023: PO2/TransformON, an ontology for data integration on food, feed, bioproducts and biowaste engineering. *npj Science of Food*, 2023, 7, pp.47. [10.1038/s41538-023-00221-2](https://doi.org/10.1038/s41538-023-00221-2). [hal-04197618](https://hal.archives-ouvertes.fr/hal-04197618)





**INRAE**

PO2 presentation on Summer webinar BioIndustry 4.0  
30/08/2024 M. Weber, L. Ibanescu, J. Cufi, S. Dervaux