Food transformation process description using PO^2 and FoodOn

Patrice Buche[a,b], Julien Cufi[b], Stéphane Dervaux[c], Juliette Dibie[c], Liliana Ibanescu[c], Alrick Oudot[b], Magalie Weber[d]

[a] *LIRMM, Univ Montpellier, CNRS, INRIA GraphIK, Montpellier, France*
[b] *IATE, Univ Montpellier, INRAE, CIRAD, Montpellier SupAgro, Montpellier, France*
[c] *UMR MIA-Paris, AgroParisTech, INRAE, Université Paris-Saclay, Paris, France*
[d] *INRAE, Nantes, France*

Abstract

The food production and processing sector are facing sustainability challenges of growing complexity. To tackle these challenges, data and knowledge from many different domains may be structured and stored using an ontology, a semantic model. In this paper we present a core ontology designed to model processes and observations from food domains. Three datasets structured according with this ontology are stored into a repository. Dedicated tools were designed to assist domain experts in integrating and querying data. Semantic integration of data from food transformation domains may enable new decision support tools for new products with good qualities and eco-friendly properties.

Keywords: food component modeling, ontology ecosystem, semantic integration, domain ontology, food transformation process, observation

---

1. Introduction

The food production and processing sector are facing sustainability challenges of growing complexity, such as global warming, the increase in over-

Email addresses: patrice.buche@inrae.fr (Patrice Buche), julien.cufi@inrae.fr (Julien Cufi), stephane.dervaux@inrae.fr (Stéphane Dervaux), juliette.dibie@agroparistech.fr (Juliette Dibie), liliana.ibanescu@agroparistech.fr (Liliana Ibanescu), alrick.oudot@inrae.fr (Alrick Oudot), magalie.weber@inrae.fr (Magalie Weber)

weight, obesity or population aging. To tackle these challenges, data and knowledge from many different domains in agriculture, food production, nutrition and health need to be modeled and structured in a way that allows storing very precise information about observations for the whole production, transformation and consuming processes, and proposing new methods for meta and multi-criteria analysis.

This paper presents our contribution to the knowledge and data representation task in food processing. In the framework of our recents projects, scientists in food process, oral physiology and sensory perception, eco-design and computer science, built a first ontology for the eco-design of transformation processes (9). This ontology has been reengineered to PO^2, a process and observation ontology in food science (12), to integrate data for three domains. The first domain concerns the formulation of dairy products, taking into account nutritional and sensory properties (11; 17). The second domain is about the manufacturing of meat products and involves an industrial partner, Solina Group, that designs and produces ingredient-based functional and culinary solutions for the food industry. The third domain is biorefinery.

Three datasets, structured according to PO^2 ontology, are stored into a repository.

Two tools were developed for integrating and querying those datasets.

The paper is organized as follows. Section 2 presents PO^2, a core ontology designed to model observations for food transformation process. Section 3 gives details about three available datasets structured according to PO^2 ontology. Section 4 presents two tools developed for integrating and querying those data.

2. PO^2 ontology

PO^2 (Process and Observation Ontology) allows one to represent a food transformation process described by a set of experimental observations available at different scales and changing over time through the different unit operations of the production process.

PO^2 has been developed using the Scenario 6 of the NeON methodology (18), i.e. reusing, merging and re-engineering ontological resources. PO^2 ontology reuses BFO (BFO), IAO (IAO), OWL-TIME (TIM), QUDT (QUD), and SSN/SOSA (13; SOS).

PO^2 ontology version 2.0, implemented in OWL 2 (OWL), is published

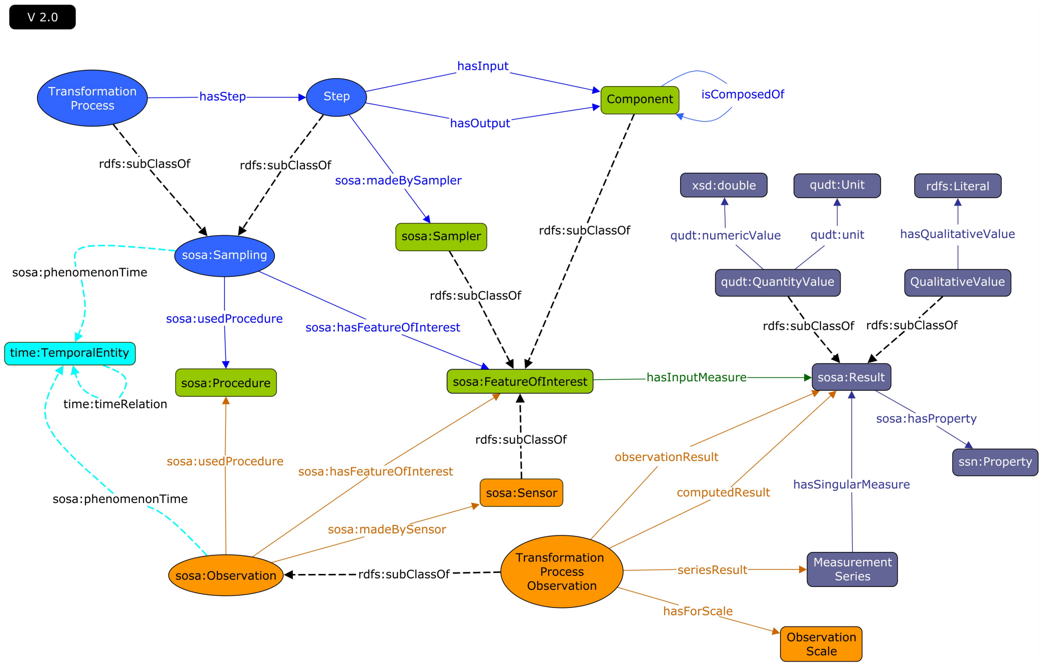

Figure 1: PO^2 ontology.

on the AgroPortal ontology library (PO2), and is Creative Commons Attribution International (CC BY 4.0) [(CCB)].

PO^2 contains 67 concepts and 79 relations. Figure 1 gives an excerpt of PO^2. There are 3 parts in PO^2:

1. a group of concepts concerning an `Observation`, i.e. the act of carrying out a `Procedure` in ordre to calculate a value of an observable property of a `FeatureOfInterest`;
2. a group of concepts concerning the `Result` of an `Observation` associated with units of measurements;
3. a group of concepts for the description of the production process: `Process` and `Step`. Each step is characterized by a temporal duration and has a collection of inputs and outputs.

The design of PO^2 allows to model a process as a sequence of unitary operations and also their inputs and outputs. An ongoing work about aligning concepts from PO^2 with the concept `Food Transformation Process` from FoodOn (10) (and its sub-concepts) should give more insights about the expressiveness of these two ontologies concerning the description of a food transformation process.

3. Domain Ontologies and Available Datasets

PO^2 (Process and Observation Ontology), briefly presented in Section 2, is the core ontology (the common part) of three domain ontologies. The first domain ontology concerns the formulation of dairy products, taking into account nutritional and sensory properties (11; 17). The second domain is about the manufacturing of meat products. The third domain is biorefinery.

Three datasets, each one structured according to the PO^2 ontology for each domain, are stored into a RDF-repository. 53 research projects are modeled: there are 997 itineraries involving 2900 steps including 369 observations records for a total of 5 557 631 RDF triples.

This RDF repository structured by PO^2 integrates into an homogeneous format different data sources which initially has heterogeneous formats.

4. Tools

`PO²Manager` is a tool designed to assist domain experts when extending the PO^2 domain ontology with concepts from other existing ontologies. It provides navigation in the concepts hierarchy, search for reusing existing semantic resources and assistance in adding a new concept. `PO²Manager` encodes the user guidelines, it is a real help for domain experts when exploring existing resources and reduces errors in data annotation. `PO²Manager` is a standalone application developed in Java.

`SPO²Q` is a Web application designed to assist users to query the PO^2 RDF repository. A set of SPARQL queries are pre-defined for users which are not familiar with SPARQL. In an advanced usage of `SPO²Q`, complex SPARQL queries may be defined.

5. Conclusion

This paper presents how we manage to model, store and query food processing for three domain using a core ontology designed to model processes and observations according to SSN/SOSA, the latest recommandation of W3C. Three datasets are stored into a RDF repository with 5.5 M triples.

Dedicated tools were designed to assist users in integrating and querying data.

As shown in (17), the semantic integration of data from dairy products using PO^2 allows i) to query in a uniform way all available data about cheese production; ii) to estimate missing data on cheese rheology, and iii) to help a Life Cycle Assessment practitioner to transfer knowledge from one domain to another by suggesting relevant parameters to be measured. In (16), PO^2 is used to learn probabilistic relational models and in (15) PO^2 is used to discover causal relations. One interesting use of semantic integration is to allows better results for decision support systems (14).

An ongoing work is to align concepts from PO^2 and FoodOn (10) when describing the manufacturing process for a meat product (i.e. sausage) and to compare the expressiveness of PO^2 and FoodOn for this task.

Future works should be done to develop decision support tools which will enable formulating new foods answering specific quality properties and produced with a controlled environmental impact. This would be a big step towards more sustainable food systems.

Acknowledgements

This work was supported by the NutriSensAl Project and FUI Metyl@b Project (BPI France) coordinated by Solina company.

References

[BFO] Basic Formal Ontology (BFO). `https://github.com/BFO-ontology/BFO`.

[CCB] Creative Commons Attribution International 4.0 International (CC BY 4.0). `https://creativecommons.org/licenses/by/4.0/`.

[SOS] Extensions to the Semantic Sensor Network Ontology. `(https://www.w3.org/TR/vocab-ssn-ext/)`.

[IAO] Information Artifact Ontology (IAO). `https://bioportal.bioontology.org/ontologies/IAO`.

[PO2] Process and Observation Ontology . `http://agroportal.lirmm.fr/ontologies/PO2`.

[QUD] Quantities, Units, Dimensions and Types (QUDT) Schema . (http://qudt.org/doc/2016/DOC_SCHEMA-QUDT-v2.0.html).

[TIM] Time Ontology in OWL. (https://www.w3.org/TR/owl-time/).

[OWL] Web Ontology Language (OWL). https://www.w3.org/2001/sw/wiki/OWL.

[9] Dibie, J., Dervaux, S., Doriot, E., Ibanescu, L., and Pénicaud, C. (2016). [MS]^{2}O - A multi-scale and multi-step ontology for transformation processes: Application to micro-organisms. In Haemmerlé, O., Stapleton, G., and Faron-Zucker, C., editors, *Graph-Based Representation and Reasoning - 22nd International Conference on Conceptual Structures, ICCS 2016, Annecy, France, July 5-7, 2016, Proceedings*, volume 9717 of *Lecture Notes in Computer Science*, pages 163–176. Springer.

[10] Dooley, D. M., Griffiths, E. J., Gosal, G., Buttigieg, P. L., Hoehndorf, R., Lange, M., Schriml, L. M., Brinkman, F. S. L., and Hsiao, W. W. L. (2018). Foodon: a harmonized food ontology to increase global food traceability, quality control and data integration. In *npj Science of Food*.

[11] Ibanescu, L., Allard, T., Dervaux, S., Dibie, J., Guichard, E., Pénicaud, C., and Raad, J. (2018). A use case of data integration in food production. In Cimiano, P. and Corby, O., editors, *Proceedings of the EKAW 2018 Posters and Demonstrations Session co-located with 21st International Conference on Knowledge Engineering and Knowledge Management (EKAW 2018), Nancy, France, November 12-16, 2018*, volume 2262 of *CEUR Workshop Proceedings*, pages 17–20. CEUR-WS.org.

[12] Ibanescu, L., Dibie, J., Dervaux, S., Guichard, E., and Raad, J. (2016). PO^2 - A Process and Observation Ontology in Food Science. Application to Dairy Gels. In Garoufallou, E., Coll, I. S., Stellato, A., and Greenberg, J., editors, *Metadata and Semantics Research - 10th International Conference, MTSR 2016, Göttingen, Germany, November 22-25, 2016, Proceedings*, volume 672 of *Communications in Computer and Information Science*, pages 155–165.

[13] Janowicz, K., Haller, A., Cox, S. J. D., Phuoc, D. L., and Lefrançois, M. (2018). SOSA: A lightweight ontology for sensors, observations, samples, and actuators. *CoRR*, abs/1805.09979.

[14] Lousteau-Cazalet, C., Barakat, A., Belaud, J. P., Buche, P., Busset, G., Charnomordic, B., Dervaux, S., Destercke, S., Dibie, J., Sablayrolles, C., and Vialle, C. (2016). A decision support system for eco-efficient biorefinery process comparison using a semantic approach. *Comput. Electron. Agric.*, 127:351–367.

[15] Munch, M., Dibie, J., Wuillemin, P., and Manfredotti, C. E. (2019). Towards interactive causal relation discovery driven by an ontology. In Barták, R. and Brawner, K. W., editors, *Proceedings of the Thirty-Second International Florida Artificial Intelligence Research Society Conference, Sarasota, Florida, USA, May 19-22 2019*, pages 504–508. AAAI Press.

[16] Munch, M., Wuillemin, P., Manfredotti, C. E., Dibie, J., and Dervaux, S. (2017). Learning probabilistic relational models using an ontology of transformation processes. In Panetto, H., Debruyne, C., Gaaloul, W., Papazoglou, M. P., Paschke, A., Ardagna, C. A., and Meersman, R., editors, *On the Move to Meaningful Internet Systems. OTM 2017 Conferences - Confederated International Conferences: CoopIS, C&TC, and ODBASE 2017, Rhodes, Greece, October 23-27, 2017, Proceedings, Part II*, volume 10574 of *Lecture Notes in Computer Science*, pages 198–215. Springer.

[17] Pénicaud, C., Ibanescu, L., Allard, T., Fonseca, F., Dervaux, S., Perret, B., Guillemin, H., Buchin, S., Salles, C., Dibie, J., and Guichard, E. (2019). Relating transformation process, eco-design, composition and sensory quality in cheeses using PO2 ontology. *International Dairy Journal*, 92:1 – 10.

[18] Suárez-Figueroa, M. C., Gómez-Pérez, A., and Fernández-López, M. (2012). The neon methodology for ontology engineering. In Suárez-Figueroa, M. C., Gómez-Pérez, A., Motta, E., and Gangemi, A., editors, *Ontology Engineering in a Networked World*, pages 9–34. Springer.