



HAL
open science

An ontology-based pipeline to support the design of technical itineraries: application to composite food packaging

Mélanie Munch, Patrice Buche, Cristina Manfredotti, Liliana Ibanescu

► To cite this version:

Mélanie Munch, Patrice Buche, Cristina Manfredotti, Liliana Ibanescu. An ontology-based pipeline to support the design of technical itineraries: application to composite food packaging. IFOW 2021 Integrated Food Ontology Workshop, Sep 2021, Bolzano, Italy. hal-04218646

HAL Id: hal-04218646

<https://hal.science/hal-04218646v1>

Submitted on 26 Sep 2023

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 4.0 International License

An ontology-based pipeline to support the design of technical itineraries: application to composite food packaging

Mélanie MUNCH¹, Patrice Buche^{1,2}, Cristina Manfredotti³ and Liliana Ibanescu³

¹LATE, Univ Montpellier, INRAE, CIRAD, Montpellier SupAgro, Montpellier, France

²LIRMM, Univ Montpellier, CNRS, INRIA GrapphIK, Montpellier, France

³UMR MIA-Paris, AgroParisTech, INRAE, Paris-Saclay University, 75005 Paris, France

Abstract

Due to the complexity of their data collection and their multi-scale approach, technical itineraries are complex and challenging to model. In this short abstract, we present a complete pipeline that aims to answer two kinds of expert competency questions: some by directly querying the data, and the others by learning a model able to reason with this complexity. By answering those, POND allows us to discover new applications and reverse engineer possible modifications.

Keywords

Ontology, Knowledge Base, Probabilistic model, Causality, Reverse engineering

1. POND: PO² ONTology Discovery

Composite food packaging is a challenging technical itinerary to design, due to the multiple parameters to be processed and optimized. Because of its potential variability, strong methods are needed to handle both its uncertainty and the need for structuration for its parameters.

In this short abstract, we present the POND (Process and Observation ONTology Discovery) pipeline, focused on the integration and the reasoning on knowledge from different sources. Based on the PO² Ontology [1] (dedicated to represent Transformation Processes), it allows the representation and the probabilistic reasoning on complex technical itineraries.

POND is composed of three different steps 1:

- **Step 1. Knowledge Collection.** Knowledge comes initially from experimental data and experts interviews. Most of it is integrated to the PO² knowledge graph, which allows to semantize and structure complex and heterogeneous data from diverse sources. This structuration is exported as an RDF Graph that can then be queried through the SPARQL language. The expert has then the possibility of expressing Competency Questions (CQ) [2], which are natural-language questions that outline the scope of knowledge represented by an ontology. In our case, CQs can either be answered using a SPARQL

IFOW 2021 Integrated Food Ontology Workshop, September 15-18, Bolzano Italy

✉ melanie.munch@gmail.com (M. MUNCH); patrice.buche@inrae.fr (P. Buche);

cristina.manfredotti@agroparistech.fr (C. Manfredotti); liliana.ibanescu@agroparistech.fr (L. Ibanescu)



© 2021 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



CEUR Workshop Proceedings (CEUR-WS.org)

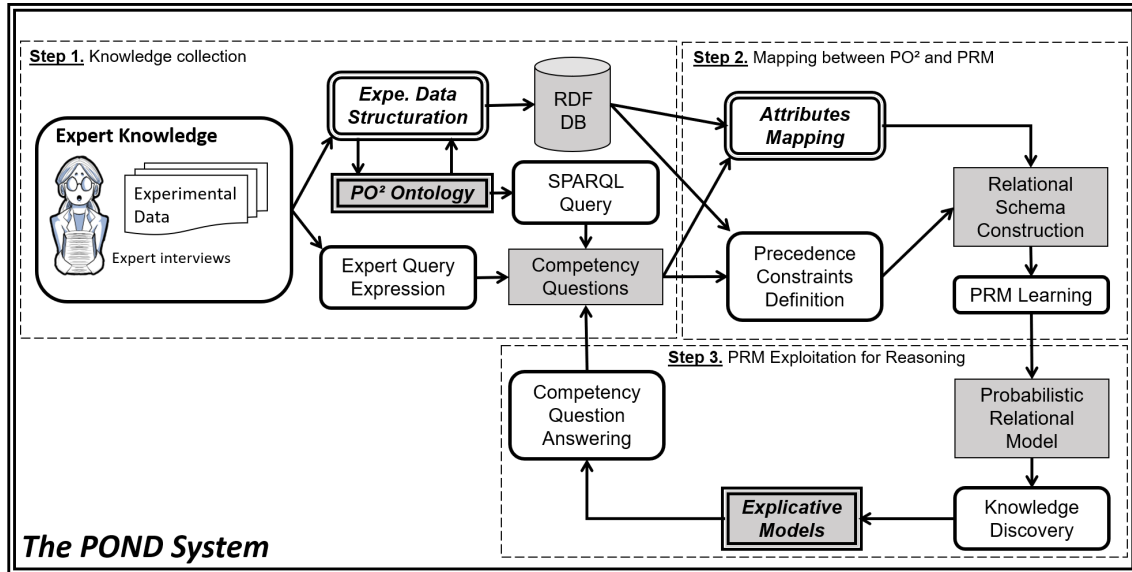


Figure 1: Presentation of the POND Pipeline, which combines ontologies for data representation and structuration, and Probabilistic Relational Models for reasoning under uncertainty.

query, or through a probabilistic reasoning with a specific model whose learning is detailed in the following steps.

- **Step 2. Mapping between PO^2 and PRM.** Some CQs can only be answered through a probabilistic reasoning. This requires the introduction of the Probabilistic Relational Model (PRM) [3], which easily allows knowledge integration while robustly modeling uncertainty. The key idea is to exploit the expert knowledge defined in **Step 1** to narrow the search space and guide the learning towards a model close to the reality in order to allow probabilistic and causal reasoning. To achieve the integration of this expert knowledge, two sub-steps have been defined: the *Attributes Mapping* where the expert defines the ontology's attributes relevant for the problem; and the *Precedence Constraints Definition* where the expert gives potential causal relations that will influence the learning [4]. All of the information is then encompassed in a Relation Schema (RS), which is used to guide the PRM learning [5].
- **Step 3. PRM Exploitation for Reasoning.** Once the PRM learned, it then becomes possible to analyse the learned model in order to answer the expert's CQ.

2. Application

Composite food packaging's goal is to provide new formulations of packages through the mix of bio-masses and polymeric matrices. In our application, a potential CQ is the possibility of reverse-engineer possible bio-masses that could be used to enhance the final package's features. In this case, the expert expresses constraints they would like to apply to the final products

(such as better mechanical properties, for instance); from the PRM, it is then possible to deduce, through inference, scenarios (i.e. a set of conditions on the inputs) that would enhance the odds of satisfying the expert's constraints. These scenarios usually require a particular composition of the biomass, which can be used to query through different databases in order to discover and recommend new material that haven't been tested in the experiment yet.

3. Conclusion and Perspectives

In conclusion, we have in this short abstract presented a complete ontology-based pipeline that can integrate expert's knowledge in order to answer different kind of competency questions through two different means: 1) by SPARQL querying and 2) by learning a probabilistic model. This last solution offers a brand new look on the data, and even suggests other designs for the studied technical itinerary. On future works, it could be interesting to see how these discoveries can help formulate explanatory rules for the studied domain.

References

- [1] L. Ibanescu, J. Dibie, S. Dervaux, E. Guichard, J. Raad, Po2- a process and observation ontology in food science. application to dairy gels, *Metadata and Semantics Research* (2016) 155–165.
- [2] M. Uschold, M. Gruninger, *Ontologies: principles, methods and applications*, *The Knowledge Engineering Review* 11 (1996) 93–136. doi:10.1017/S0269888900007797.
- [3] D. Koller, N. Friedman, *Probabilistic Graphical Models: Principles and Techniques - Adaptive Computation and Machine Learning*, The MIT Press, 2009.
- [4] M. Munch, J. Dibie, P. Wuillemin, C. E. Manfredotti, Towards interactive causal relation discovery driven by an ontology, *Florida Artificial Intelligence Research Society Conference* (2019) 504–508.
- [5] B. T. L. Getoor (Ed.), *Introduction to Statistical Relational Learning (Adaptive Computation and Machine Learning)*, 2007.