

Crop Planning and Production Process Ontology (C3PO), a new model to assist diversified crop production

Baptiste Darnala^{1,2}, Florence Amardeilh², Catherine Roussey³ and Clément Jonquet^{1,4}

¹LIRMM, University of Montpellier, CNRS, Montpellier, France

²Elzeard, Cité du Numérique, Bègles, France

³TSCF, Université Clermont Auvergne, INRAE, Aubière, France

⁴MISTEA, University of Montpellier, INRAE, Institut Agro, Montpellier, France

Abstract

Crop planning and production process ontology (C3PO) is a representation of agricultural knowledge for diversified crop production. It is composed of eight modules to take into account all the elements involved in crop production as well as the complexity for representing farming practices and constraints. Especially, the crop management module is sliced in three different layers to capture the generic information about crop itineraries, the planned production process and the tasks eventually really achieved by various agents in the plots. To implement this ontology, we designed a new pipeline, presented here, integrating various existing tools to ease the ontology documentation, population and maintenance tasks.

Keywords

ontology, agriculture, semantic resource, knowledge representation, crop itinerary

1. Introduction

Agricultural production practices are changing and crop production are facing multiple vulnerabilities: economic, regulatory, environmental and social (isolation, skills, overload, etc.). Moreover, the pressure created by distribution chains and the consumers' demand for better and healthier vegetables, produced in local area, force farmers to improve crop management. Several scientific studies [1, 2] have shown diversity --both in time and space-- of cultivated plants:

- improves risks management with respect to weather and economic changing context;
- delivers a natural defence system against diseases and pest attacks;
- increases stability and resilience of agro-ecosystems.

However, with new agroecological techniques, farmers' workload, management complexity and mental load also increase [3, 4]. Between 2019 and 2020, the SME Elzeard conducted over


IFOW'21: Integrated Food Ontology Workshop, September 15-18, Bolzano Italy


✉ baptiste.darnala@elzeard.co (B. Darnala); florence.amardeilh@elzeard.co (F. Amardeilh);

catherine.roussey@inrae.fr (C. Roussey); jonquet@lirmm.fr (C. Jonquet)

🌐 <https://www.elzeard.co/> (B. Darnala)

🆔 0000-0002-6306-4437 (F. Amardeilh); 0000-0002-3076-5499 (C. Roussey); 0000-0002-2404-1582 (C. Jonquet)

 © 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)

150 interviews with farmers, agricultural advisors, teachers and researchers. The company identified several technological barriers, among them the need for knowledge sharing and for new operational tools to assist farmers in their daily crop management: optimizing their crop rotations, their yields and finding alternatives to chemicals.

To address the previous identified need, Elzeard develops a web and mobile application to assist farmers in their planning and production activities. In order to encode the knowledge in the system and to provide right decision-support and adapted recommendations, Elzeard have designed an ontological model for representing plot management and crop itineraries, called *Crop Planning and Production Process Ontology* (C3PO). This work is done in collaboration with the ANR D2KAB project consortium ¹ and the MesclunDurab consortium ². The objective of this ontology is to structure a knowledge graph about agroecological crop practices, that will become a common good for the benefits of the various actors of the vegetable production sector. The knowledge graph will be accessible via a Web portal, called *La Serre des Savoirs* being currently developed. *La Serre des Savoirs* will leverage the information contained in the knowledge base into a more usable way for the different actors of the food industry. The C3PO ontology will also be published and shared on the AgroPortal ontology repository [5] and adhere to the FAIR principles (Findable, Accessible, Interoperable, Reusable) [6].

In this paper, we present preliminary results about the design requirements and implementation workflow of the C3PO ontology. Section 2 covers related research regarding crop production models; Section 3 explains the specificity of the diversified vegetable crop production representation; Section 4 outlines the first version of the C3PO ontology; Section 5 describes the ontology creation pipeline used to develop, document and populate the ontology with reference data.

2. Related work

Semantic Web technologies were chosen to represent the interlinked knowledge needed to manage crop planning and production process in its digital application because of their two key advantages. Not only do these technologies allow knowledge to be represented through formal and logical semantics enabling native reasoning (the ontology), but they also allow this knowledge to be interconnected with open linked data thanks to shared or aligned ontologies, thus promoting data interoperability (graph database).

For designing the C3PO ontology, we firstly examined existing ontologies [7] that could be reused to represent plant characteristics and crop itineraries, that is to say general information on a sequence of farming practices to ensure optimal growth of all crops in a plot. Some agronomic resources about plants or crops are already available in a semantic format like the agroecological knowledge base GEKO [8] or in various thesaurus and agronomic reference ontologies like Plant Ontology [9], Crop Ontology [10], AGROVOC [11], TAXREF-LD [12] or FrenchCropUsage [13]. Nevertheless, these semantic resources did not describe farming practices and their interactions from an agricultural production perspective.

Searching AgroPortal's content for "crop management plan" or "crop itinerary", returned

¹<http://www.d2kab.org/>

²<https://maraichage.wixsite.com/mesclun>

several results. The first one is the Durum Wheat ontology but it only contains "itinerary" as a <skos:Concept>, i.e. as a terminological definition. We also found a "TechnicalItinerary" as a <owl:Class> into MS2O but this ontology is now deprecated. Another ontology identified was AgroRDF [14]. It represents and describes the farm practices as a set of technical processes. Unfortunately some of their classes are only defined in German and there is no sequence of processes or plan.

Lately we also discovered Agronomy Ontology (AGRO) [15]. The farming practices are defined as a set of technical processes, named planned processes, in the context of agronomic experiments. This ontology is based on Basic Formal Ontology (BFO) [16]. The main scope of application of BFO seems to focus on the concrete representations of the real world and not really on the abstract representations of the human mind that may help in decision making. That is why AGRO is oriented towards the recording of precise observations concerning specific agronomic experiments rather than towards the planning of agricultural production from reference data. We are already planning future work to interconnect the definitions of common Farming practices with those of the Agronomic Ontology.

We also looked for process representation from other domains, such as the W3C Recommendation Prov Ontology [17]. Prov-O is a process-flow model in the sense that it traces the provenance and evolution of Activity concepts, interacting with Agents and other Entities. It is a very interesting model because it brings a representation of agents, activities but also entity collection, ordered or not, that can be very useful to represent a vegetable production. Indeed, with Prov-O, we could represent the farm, its actors represented by the producers for example, and the different activities realised on plants or in the plots. But after testing its implementation on our use cases, we chose not to use Prov-O to represent crop planning and production process as it does not allow a theoretical representation of activities with relative dates placed into the calendar. We rather decided to use Prov-O to keep trace of the updates on the various entities in the knowledge graph data.

We studied another process ontology dedicated to transformation processes in food science, called Process and Observation Ontology (PO2) [18]. This ontology is also derived from BFO principles and has the same drawbacks that AGRO. It was a good start for implementing operational processes composed of steps, products, observations, measures, and agents. But it was not adapted to distinguish the reference data about the crop itineraries from the observed data in a specific plot.

Recently, we discovered Valueflows³, an ontology that describes flows of economic value according to a three layers representation (Knowledge, Plan, Observation). As described in Section 4, we used the same three-layers division to design the ontology part about crop management process. To the best of our knowledge, this is the first conceptualisation of a sequence of cultivation processes associated with a production plan linked to some generic knowledge about those plans.

³<https://lab.allmende.io/valueflows>

3. Representing diversified vegetable production

The workflow of vegetable production and farm management requires a high level of knowledge with regards to production techniques, especially to adapt to the on-going agroecological transition. Moreover, with increasing regulatory standards and consumer pressure, it becomes very complicated for farmers to produce diversified, tasty and healthy vegetables without too many economic losses. It is therefore necessary to relieve the mental burden of farmers by identifying all the knowledge related to agricultural techniques that allow them to organise, in time and space, the schedules of each of the crops as well as their association and rotation over a given season.

To face this challenge, the digital services offered by Elzeard to vegetable farmers rely on the exploitation of the technical crop itineraries. A crop itinerary can be defined as *a generic process consisting of a succession of technical tasks and recommendations for agricultural production*. For a given crop, the crop itinerary has to be specialised according to the cultural context and the farmer's cultivation methods. Elzeard's decision support system assists farmers in the adaptation from a generic crop itinerary to a custom-made one, therefore optimizing crop production.

The scope of the Crop Planning and Production Process Ontology (C3PO) presented in this paper therefore targets information about the various possible crop itineraries and their dependence with the whole agro-ecosystem. It has to take into account the general information about the crops and cultivars but also the constraints about the farming system (organic or conventional), the location of the farm along with the soil and climate conditions, the client orders, the crop rotations and associations, the farming equipments, the potential diseases or aggressors and crop protection solutions, etc.

4. Ontology modelling

As the scope of this ontology is quite large and complex, we have chosen to split it into eight modules where each part defines a particular need of crop production. The modules interact with each other to represent the entire planning and production process on a farm, from seed selection to crop delivery. Moreover, in order to be interoperable according to FAIR principles, the modules have been linked to external ontologies for direct reuse or alignment wherever possible as described in Figure 1.

- The Plant module represents cultivated plants (e.g., "Carrot" vegetables), their botanical family (e.g., "Apiaceae"), their usage category (e.g., "Root vegetable") and the different cultivars (e.g., "Nantaise") with their characteristics like the earliness, size, colour, shape. This module reuses the French Crop Usage thesaurus [13] for vernacular names, crop usages and aligns with TAXREF-LD [12] for scientific names and taxonomic organization of families and species and plant development stage with BBCH (*Biologische Bundesanstalt, Bundessortenamt und CHemische Industrie*)⁴.

⁴<http://ontology.irstea.fr/bbch/core>

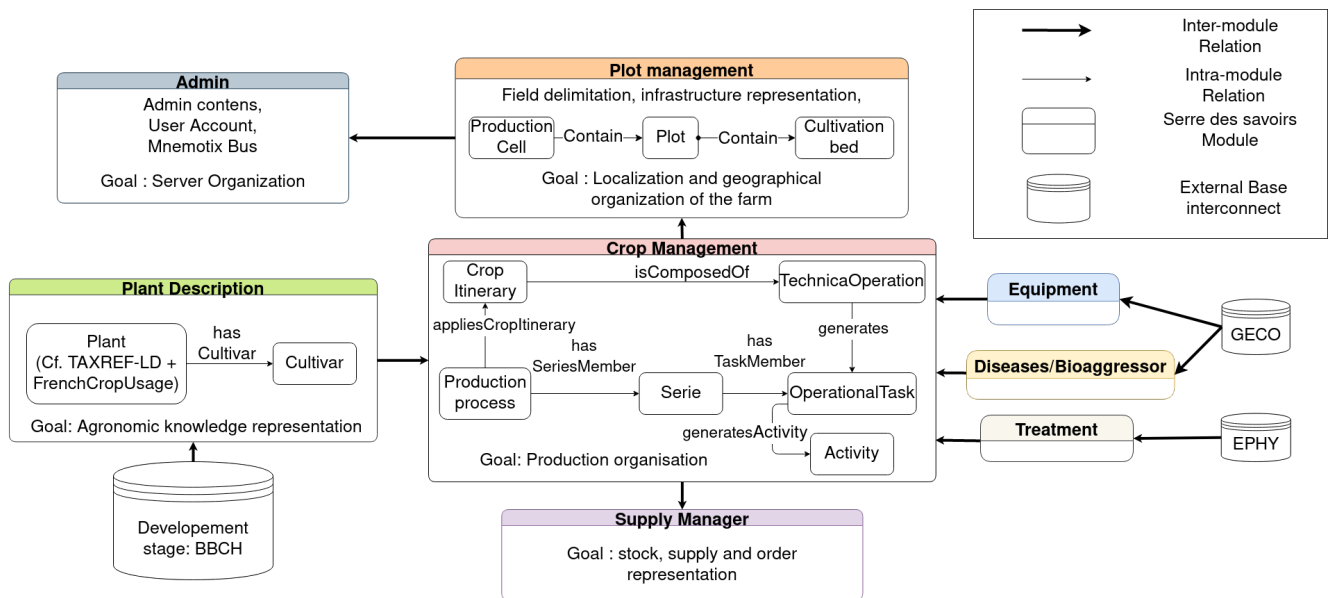


Figure 1: Overview of the Crop Planning and Production Process Ontology.

- The Aggressor module defines potential diseases and pests of each cultivated plants with their description and protection solutions. For this we will rely on the GECO (*GEstion des Connaissances*)⁵ ontology model and knowledge base. We are working with the GECO's team in D2KAB to document the model and align it with other resources such as Agrovoc, Wikidata or TaxRefld to complete with additional data about those living organisms.
- The Treatment module contains the description of every inputs that enhances plant growth or plant defence. For this we will rely on an OWL version of French ANSES's E-Phy inputs catalogue developed in D2KAB.
- The Plot module represents the spatial organisation of the farm, the division into plots and even beds for diversified vegetable production, the different infrastructures present on the farm like the irrigation equipment, etc. For this we have not yet identified an external semantic resource that goes into sufficient details regarding the plot description dedicated to diversified crop productions.
- The Admin module imports and links with existing ontology models such as FOAF, Time Ontology, the Event Ontology or Prov-O, to represents agents, organisation or administrative entities.
- The Supply Manager module organises the stocks and the product delivery. For this module, we are discussing with the DataFoodConsortium⁶ to reuse part of their supply chain ontology model, which they are currently also aligning on ValueFlows model.
- The Crop Management module represents all the crop production process. It is the core of the C3PO ontology, importing all other modules needed to organise the crop production.

⁵<https://geco.ecophytopic.fr/>

⁶<https://www.datafoodconsortium.org/en/>

- The Vocabulary module, formalized in SKOS, organises reference values used for the definitions of class instances, like units of measures (aligned with QUDT definitions), the climate choices, the farming practices, the earliness levels and so on. Each time it is possible, reference values are aligned to AGROVOC concepts.

The Crop Management module is the most complex. To better take into account the crop planning and production process, we need to represent crop itineraries from different perspective: the generic theoretical information as advised in agricultural literature –how tomatoes should be grown under various constraints; the planned process –how a farmer plan to produce tomatoes according to its own constraints; the observation data –how a farmer eventually produced tomatoes taking into account possible hazards.

Those three dimensions better capture changes between plans and results in order to optimise production from one year to another. Additionally, we want to encourage the sharing of observation data between farmers to update theoretical knowledge if needed to propose adaptations to climate change for example.

To describe the process according to those separate layers we get inspiration from the ValueFlows model, itself inspired by the REA ontology [19] about economic phenomena occurring in Enterprise Information System. REA introduced a distinction between Knowledge Infrastructure and Operational Infrastructure whereas ValueFlows model goes further distinguishing the Knowledge level (core concepts, guidelines, recipes, procedures), the Plan level (offers, requests, schedules and plans) and the Observation level (what really happened). This three-layer representation completely fits our requirements for modelling the crops' operational planning and production process. Figure 2 describes the production flow with an example. The Knowledge Level describes the generic concepts about plants, technical operations that are part of a itinerary procedure, implementing specific farming practices organised as a hierarchy, the various contextual conditions that may affect the expected yield, the relative periods of the different steps of the process (germination, plantation, maintenance, harvest, conservation), etc. Then the Plan Level represents the production process that is deduced from the itinerary procedure, the cultivars used as a seed material, the planned operational tasks, from soil preparation to harvest, according to market opportunities, delivery dates and volumes, etc. Moreover, in a diversified crop production, such as for the vegetables, it is important to take into account the fact that a production process can be sliced into various series of the same crop in order to have longer lasting harvests. A farmer repeats more or less the same tasks of a production process, at different dates to benefit from a continued flow of production. Lastly, the Observation Level declines the planned operational tasks into activities that are assigned to a specific location (plots and/or beds), persons or teams and at a certain date. It gives an historical view of the achieved task with complementary data about the equipment used, the inputs dose or volumes, the time spent to do the activity, etc. It also represents observation data (weather conditions, hazards, presence of an agressor, and so on).

5. Pipeline of ontology production

To start implementing our ontology, we used the Linked Open Terms (LOT) [20] methodology as it proposes an agile and iterative approach to ontology development. First, we worked with

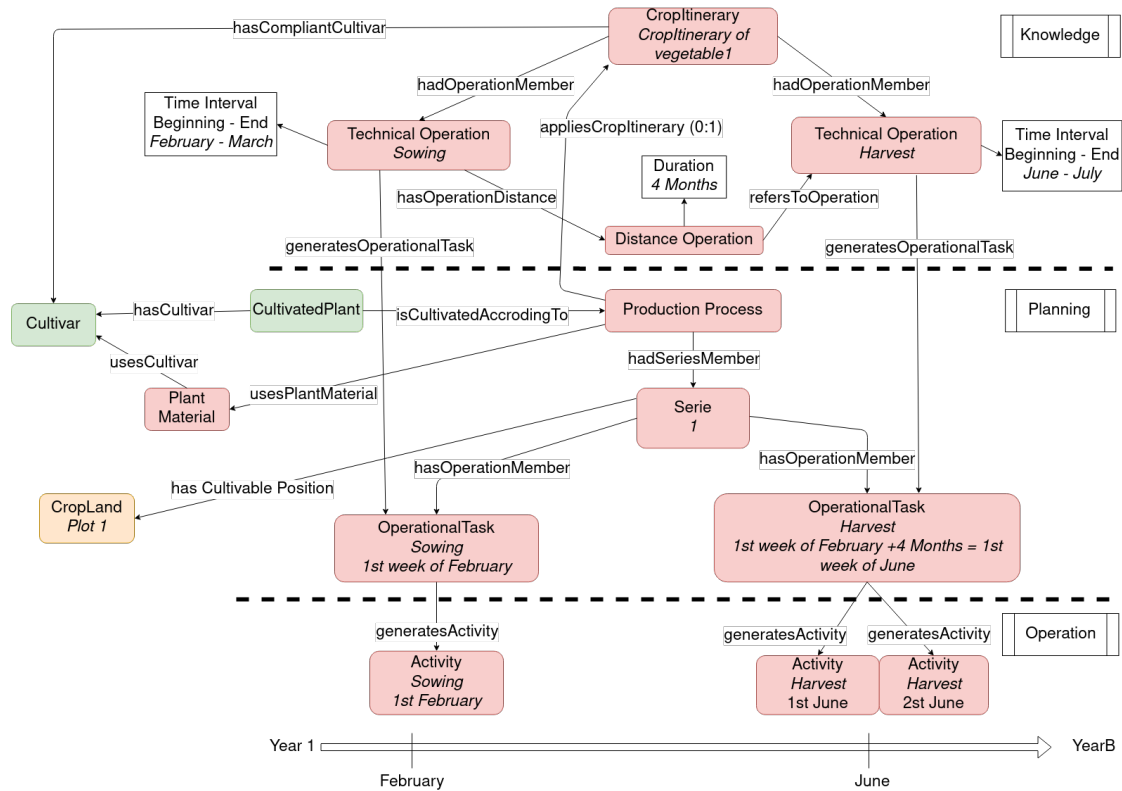


Figure 2: Example of a production process with concepts represented in the three layers of the C3PO module

several domain experts to identify a set of competency questions along with the functional specification of Elzeard’s application. For example, “what are the ordered technical tasks to cultivate organic tomatoes in Mediterranean climate?”

We gathered additional requirement specifications from various domain-oriented text books, guidelines procedures, crop datasets or dedicated knowledge bases such as GECCO. We also chose to reuse existing ontologies and thesauri in order to avoid reinventing the wheel and facilitate the interoperability of our model, either upper-level metadata vocabularies such as FOAF, EVENT Ontology or OWL-TIME or domain-specific ontologies such as FCU or TAXREF-LD.

At first, to model the ontology we used a graphical model by using the online graphic tool “Draw.io”⁷ to ease the discussions with the domain experts. Then, we had to manually transcript the concepts’ definitions of this graphical model in an ontology editor such as Protégé [21]. Not only this method is prone to human error when transcribing the ontology model but it may lead to serious problems when keeping the graphical model up to date, still useful to provide an overview of the ontology model. It appeared that some modifications made into Protégé were not always updated on the graphical model. To counter this problem, an ontology creation pipeline was built by combining different existing tools in the same workflow.

⁷<https://www.draw.io>

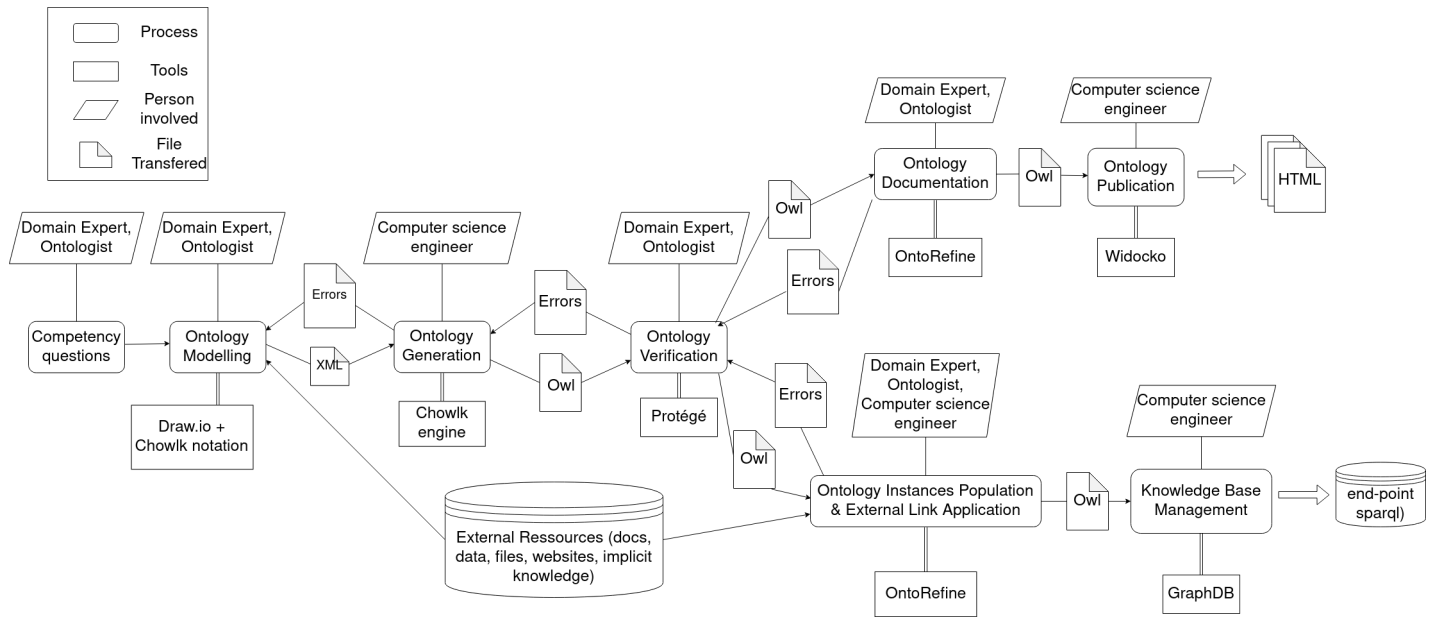


Figure 3: C3PO's pipeline for ontology production.

The first tool of that pipeline is still Draw.io but we added Chowlk [22], a graphical extension for an ontology representation. Thanks to a standardized graphical notation to represent ontological classes, properties and restrictions, Chowlk automatically transcribes Draw.io's XML output into an OWL representation with the Turtle syntax. Then, the generated TTL file can be imported into Protégé for verification and validation by the ontologist. Then the ontology model can be imported in a RDF triple store. GraphDB⁸ has been chosen for its performance and reasoning capabilities as the RDF storage of the future Knowledge Base. Moreover, to simplify the documentation process, GraphDB includes OntoRefine, an integrated upgraded version of the open-source tool OpenRefine⁹. It enables data transformation for converting tabular data (e.g. CSV) into RDF data. There are two uses for this tool. The first one is to import the concepts and properties definitions (all labels, comments and other annotation properties) that are maintained outside the Draw.io to make the graphical model easier to read. The second one is to populate the data graph with knowledge provided by domain experts in tabular files as it is often more convenient for them. Once the ontological model is complete, the last tool used is Widoco [23] to automatically generate all the ontology documentation.

The C3PO ontology documentation is currently being finalised and will be published at its namespace URL¹⁰. Meanwhile, the C3PO's owl files are available at *Serre des Savoirs* project's GitLab repository¹¹ to collect future contributions and revision requests.

⁸<https://www.ontotext.com/graphdb>

⁹<https://openrefine.org/>

¹⁰<http://www.elzeard.co/ontologies/c3po>

¹¹<https://gitlab.com/serre-des-savoirs/c3po>

6. Conclusion

To comply with operational needs for Elzeard's digital application, we modelled a new ontology for representing Crop Planning and Production Process dedicated to diversified crop production. Even if we tried to reuse existing thesaurus and agronomy ontologies, there was no existing ontology that captured the specific need of representing crop itineraries from the perspective of the farmers' production processes. The C3PO ontology is used to construct a knowledge base for the storage of the Elzeard web application data and into *la Serre des Savoirs*, a web portal that will give access to extensive knowledge about plants and their crop itineraries on a practical way for the agriculture actors. Various domain expert partners were included in our approach to question and identify the main concepts and properties: scientists (D2KAB, MESCLUNDURAB) and agricultural professionals (crop farmers, networks of agricultural advisors, teachers.). A pipeline was initiated and tested to ease the different steps of ontology creation, documentation and population, integrating various independent tools. In future work, we will align our concepts and properties to BFO models, such as Agronomy Ontology or PO2. We will also publish the ontology model according to the FAIR principles on AgroPortal and our GitLab for tracking issues and new contributions. For the first time, an ontology with information oriented on French farming practices is developed, both in French and in English. However, the ontology is intended to become a global open knowledge base about all farming practices and crop itineraries related to diversified crop production such as market gardening, arboriculture or agroforestry.

Acknowledgments

This work was partially achieved with support of the Data to Knowledge in Agronomy and Biodiversity (D2KAB – www.d2kab.org) project that received funding from the French National Research Agency (ANR-18-CE23-0017). We also acknowledge support from the National Office for Biodiversity with MesclunDurab grant and the Nouvelle-Aquitaine Region with "Social Innovation AMI" and "Digital Prototypes" grants. We also thank Dr. Kevin Morel for his feedback and all contributors from MesclunDurab and D2KAB projects.

References

- [1] F. Isbell, P. Adler, N. Eisenhauer, D. Fornara, K. Kimmel, C. Kremen, D. Letourneau, M. Liebman, H. Polley, S. Quijas, M. Scherer-Lorenzen, Benefits of increasing plant diversity in sustainable agroecosystems, *J. Ecol.* 105 (2017) 871–879. URL: <https://doi.org/10.1111/1365-2745.12789>. doi:10.1111/1365-2745.12789.
- [2] R. Paut, R. Sabatier, M. Tchamitchian, Reducing risk through crop diversification: An application of portfolio theory to diversified horticultural systems, *Agricultural systems* 168 (2019) 123–130.
- [3] K. Morel, F. Léger, Strategies to manage crop planning complexity in very diversified direct selling farming systems: the example of organic market gardeners, *Proceedings of the 5th International Symposium for Farming Systems Design* 93-94 (2015).

- [4] B. Dumont, Why working conditions are a key issue of sustainability in agriculture? a comparison between agroecological, organic and conventional vegetable systems, *Journal of Rural Studies* 56 (2021) 53–64.
- [5] C. Jonquet, A. Toulet, E. Arnaud, S. Aubin, E. Dzalé-Yeumo, V. Emonet, J. Graybeal, M.-A. Laporte, M. Musen, V. Pesce, P. Larmande, Agroportal: an ontology repository for agronomy, *Computers and Electronics in Agriculture* 144 (2018) 126–143. URL: <http://agroportal.lirmm.fr/>.
- [6] I. J. A. a. Mark D. Wilkinson, Michel Dumontier, The fair guiding principles for scientific data management and stewardship, *Scientific Data* 3 (2016). URL: <https://www.nature.com/articles/sdata201618.pdf>. doi:10.1038/sdata.2016.18.
- [7] B. Drury, R. Fernandes, M.-F. Moura, A. de Andrade Lopes, A survey of semantic web technology for agriculture, *Information Processing in Agriculture* 6 (2019) 487–501.
- [8] V. Soullignac, F. Pinet, E. Lambert, L. Guichard, L. Trouche, S. Aubin, Geco, the french web-based application for knowledge management in agroecology, *Computers and Electronics in Agriculture* 162 (2019) 1050–1056.
- [9] Avraham, al, The plant ontology database: a community resource for plant structure and development stages controlled vocabulary and annotations, *Nucleic Acid Research* 36 (2008) 449–454.
- [10] E. Arnaud, L. Cooper, R. Shrestha, N. Menda, R. Nelson, L. Matteis, M. Skofic, R. Bastow, P. Jaiswal, L. Mueller, G. McLaren, Towards a reference plant trait ontology for modeling knowledge of plant traits and phenotypes, proceedings of the 4th Conference on Knowledge Engineering and Ontology Development, 4-7 October 2012, Spain (2012).
- [11] C. Caracciolo, A. Stellato, A. Morshed, G. Johannsen, S. Rajbhandari, Y. Jaques, J. Keizer, The agrovoc linked dataset, *Semantic Web Journal* 4 (2013) 341–348.
- [12] F. Michel, O. Gargominy, S. Terceirie, C. Faron-Zucker, A model to represent nomenclatural and taxonomic information as linked data. application to the french taxonomic register, taxref, Proceedings of ISWC 2017 Workshop on Semantics for Biodiversity (S4Biodiv 2017), Oct 2017, Vienna, Austria (2017) 1–12.
- [13] C. Roussey, Frenchcropusage: Thésaurus sur les cultures françaises. le thésaurus décrivant les cultures françaises par leur utilisation au format skos, FAIRsharing.org (2018). URL: [FAIRsharing.9228fv](https://fairsharing.org/9228fv). doi:10.25504.
- [14] D. Martini, M. Schmitz, E. Mietzsch, agrordf as a semantic overlay to agroxml: a general model for enhancing interoperability in agrifood data standards, Proceedings of EFITA-WCCA-CIGR Conference “Sustainable Agriculture through ICT Innovation”, June 23-27 2013, Torino, Italy (2013).
- [15] M. Devare, C. Aubert, M.-A. Laporte, L. Valette, E. Arnaud, P. Buttigieg, Data-driven agricultural research for development: A need for data harmonization via semantics, Proceedings of the Joint International Conference on Biological Ontology and BioCreative, Corvallis, Oregon, United States, August 1-4, 2016. (2016). URL: http://ceur-ws.org/Vol-1747/IT205_ICBO2016.pdf.
- [16] R. Arp, B. Smith, A. Spear, Building Ontologies With Basic Formal Ontology, MIT Press, 2015.
- [17] T. Lebo, S. Sahoo, D. McGuinness, K. Belhajjame, J. Cheney, D. Corsar, D. Garijo, S. Soiland-Reyes, S. Zednik, J. Zhao, PROV-O: The PROV Ontology, W3C Recommendation, World

Wide Web Consortium, 2013.

- [18] S. D. E. G. J. R. Liliana Ibanescu, Juliette Dibie-Barthelemy, Po² - a process and observation ontology in food science. application to dairy gels, Metadata and Semantics Research - 10th International Conference, Nov 2016, Gottingen, Germany 105 (2016) 155–165. URL: <https://hal-agroparistech.archives-ouvertes.fr/hal-01418097>. doi:10.1007/978-3-319-49157-8_13.
- [19] G. L. Geerts, W. E. McCarthy, The ontological foundation of rea enterprise information systems, Presentation to the Annual Meeting of the American Accounting Association, Philadelphia, August 2000 (2000). URL: https://www.williamemccarthy.com/s/The-Ontological-Foundations-of-REA-Enterprise-Information-Systems_2000.pdf.
- [20] M. Poveda-Villalón, A reuse-based lightweight method for developing linked data ontologies and vocabularies, in: The Semantic Web: Research and Applications, 2012, pp. 833–837.
- [21] M. A. Musen, The protégé project: a look back and a look forward, AI Matters 1 (2015) 4–12. URL: <https://doi.org/10.1145/2757001.2757003>. doi:10.1145/2757001.2757003.
- [22] S. Chavez-Feria, R. Garcia-Castro, M. Poveda-Villalón, Converting uml-based ontology conceptualizations to owl with chowlk, ESWC2021 Poster and Demo Track at 18th International Conference, ESWC 2021, Virtual Event, June 6–10, 2021 12731 (2021).
- [23] D. Garijo, Widoco: A wizard for documenting ontologies, The Semantic Web – ISWC 2017 10588 (2017). URL: https://link.springer.com/chapter/10.1007%2F978-3-319-68204-4_9. doi:https://doi.org/10.1007/978-3-319-68204-4_9.