To ontology, and beyond! Modernizing a widely adopted reference model for agriculture

Joep Tummers[1,†], Daoud Urdu[1] and Roel Veraart[2]

[1]Wageningen Economic Research, Droevendaalsesteeg 4, 6708PB Wageningen, The Netherlands
[2]Wageningen University & Research, Philosophy Group, Hollandseweg 1, 6706KN Wageningen, The Netherlands

Abstract

Introduction: Once upon a time, there was a reference model for the agricultural domain named rmAgro. This model was widely used by multiple agricultural information system providers in the Netherlands and contained extensive domain knowledge. Although the model remains popular in the field of agri-information modeling, it is presented in a conventional approach: UML class diagrams. In a sector like agriculture, which is in dire need of interoperability, there is only one way forward: an Ontological publication of the model!

Methods: We are currently modernizing rmAgro into an ontology as part of the NXTgen HighTech programme: "Dutch Data Line". We are converting numerous classes and attributes from the UML model. The main question that we need to address is: How do you report into a future-proof model? This work is performed in close collaboration with domain experts from Wageningen University, software developers, and connections with other ontologies such as FoodOn, AIM, and SAREF4AGRI are being established.

Results: A first version of the ontology is being implemented in a use case focusing on grassland management with a digital calendar. This presentation will explore the design patterns, challenges encountered, and the strategies employed to navigate the complexities of ontological development within the agricultural context.

Conclusion: The ontological rendition of rmAgro holds significant promise for replacing its UML-based predecessor, marking a pivotal step towards modernizing agricultural data systems. While this transformation is still a work in progress, it represents a critical move towards adopting ontologies for meaningful data sharing, integration, and interoperability in agriculture.

Keywords

agri-food systems, data-sharing, semantics, data spaces

JOWO Workshop – IFOW 2024, July 15-19, 2024, Enschede, NL
joep.tummers@wur.nl (J. Tummers); daoud.urdu@wur.nl (D. Urdu); roel.veraart@wur.nl (R. Veraart)
https://www.wur.nl/en/persons/joep-tummers.htm (J. Tummers);
https://www.wur.nl/en/persons/daoud-urdu.htm (D. Urdu); https://www.wur.nl/en/persons/roel-veraart.htm (R. Veraart)
0000-0001-8144-0822 (J. Tummers); 0000-0003-2864-6658 (D. Urdu); 0000-0001-9137-2704 (R. Veraart)

1. Background

Agriculture has a long history in industrialization, mechanization, and nowadays, with digitization, making data interoperable is an ongoing effort. In the Dutch context, the Reference Model Agriculture (rmAgro) has been utilized for decades by various applications. This model acts as a benchmark for defining data exchanged within the agricultural industry. Its primary aim is to offer a detailed description and definition of object classes recognized in agriculture and their characteristics. Initiated and maintained by Wageningen University, rmAgro adheres to the latest design principles in information technology, evolving continuously to meet new and changing functional requirements. The model is presented through UML class diagrams and is accessible via a web interface, providing a visual representation of the model.

In an increasingly digitized agricultural sector, the collection of data is expanding, yet farmers are gradually losing control over their own data. Furthermore, the data is scattered across different systems resulting in a lack of interoperability. The NXTgenHighTech "Dutch Data Line," was established in response to this concern. The project aims to develop a farm data space, which is a blueprint for the "Information House for Agriculture" (Dutch: "Informatiehuis Landbouw"). This instantiation of the blueprint demonstrates how farm data spaces can facilitate data interoperability and utilization for software to both supply chains and government bodies. This simplifies data reporting and usage for insights and presentations for farmers, improving their position in the data economy.

The farm data space is conceptualized not as a data lake or physical location but as a federative system of individual spaces. These spaces are interconnected to enable mutual data reuse and are linked to stakeholders in the supply chain, government agencies, knowledge institutions, and service providers.

Adopting a farmer-centric approach, the Dutch Data Line project utilizes the SOLID protocol, ensuring that data remains with the farmer in a so-called "Pod". This approach necessitates a common interoperability stack for the farm data space and a uniform data annotation within the pod, enabling multi-application use. This effort focusses on the considerations for the development of an ontology that supports the use case of grassland management and usage with a digital calendar. For this use case, close interactions are established with two existing, but different, application developers.

Our challenge lies in the existing agricultural reference model, rmAgro, which is a platform independent model and thus necessitates to be directly implemented in the farm data space despite the technological differences. Consequently, we aim to re-envision rmAgro within a new framework that is more robust and sustainable.

Research Question:
How can we develop an agricultural reference ontology that is compatible and interoperable with the semantics of a farm data space?

2. Methods

Within the Dutch Data Line project, we embarked on developing a use case centered around two existing applications utilized by farmers for grassland management. The initial step involved identifying the data objects essential for these applications. Subsequently, these data objects were mapped onto existing standardized agricultural ontologies such as AIM and SAREF4AGRI, with rmAgro UML model as a starting point. Despite the availability of some relevant objects (e.g., farm, field), we found these ontologies insufficiently comprehensive in defining specific concepts and properties to encapsulate the full scope of the grassland management application. This realization prompted us to integrate it with other agricultural ontologies wherever feasible and necessary while revisiting the rmAgro model.

The transformation from a UML-based model to an ontology posed substantial challenges, necessitating extensive domain knowledge. To navigate this complex process, we enlisted the expertise of a diverse group of professionals from the agricultural domain, including scientists and domain experts from Wageningen University, as well as software developers. In the first iteration, we crafted an initial version of the ontology, focusing primarily on farms and their fields, with a particular emphasis on the geographic characteristics of the fields. This task was accomplished using the TopBraid Composer tool. The development process required a series of critical decisions to establish a logical hierarchy that would seamlessly integrate with existing applications. These decisions were rigorously validated through a collaborative and iterative process involving multiple domain experts, ensuring the ontology's alignment with real-world agricultural practices and semantics.

3. Results

The initial iteration of the ontology was seamlessly integrated into the premiere version of the farm data space, where it played a crucial role in annotating the data stored within a pod. This integration enabled semantic interoperability, allowing two distinct grassland management applications to leverage the same data effectively. Highlighting its accessibility and intent for widespread adoption, the ontology will be made publicly available on rmAgro.org. This current version encompasses a broad array of concepts, setting a robust foundation for further development and application in the agricultural domain.

4. Conclusion

Although this is merely the initial and emerging version, its potential is unmistakable. It has already facilitated interoperability between two applications dedicated to grassland management. Future ontological developments will focus on expanding the model to cover all facets of grassland usage applications. Currently, it encompasses only farming organizations and fields, but forthcoming updates will include additional aspects related to crop registration. Moreover, efforts to enhance integration with SAREF4AGRI, AIM, and FoodOn are underway, promising a more deliberate, comprehensive and interconnected framework.