Semantic modelling of Plant Protection Products data

Proof of Concept: Variable Rate Application of soil herbicide

Daoud Urdu, Daan Goense, Derek Scuffell, Hugo Mills, Patricia Perez, Conny Graumans

Semantic modelling of Plant Protection Products data

Proof of Concept: Variable Rate Application of soil herbicide

Daoud Urdu[1], Daan Goense[1], Derek Scuffell[2], Hugo Mills[2], Patricia Perez[3], Conny Graumans[4]

1 Wageningen Economic Research
2 Agrimetrics
3 LexAgri
4 AgGateway

Use case contributions: Jits Riepma (WUR), Corne Kempenaar (WUR)

Review: Sjaak Wolfert (WUR), Koos van der Meij (WUR)

General contributions: Harald Sundmaeker (ATB), Jules Bloem (WUR), Annette Breemer (WUR)

Wageningen Economic Research
Wageningen, May 2022

White Paper
2022-051
ISBN 978-94-6447-298-1

Urdu, D., D. Goense, D. Scuffel, H. Mills, P. Perez, C. Graumans, 2022. *Semantic modelling of Plant Protection Products data; Proof of Concept: Variable Rate Application of soil herbicide.* Wageningen, Wageningen Economic Research, White Paper 2022-051. 40 pp.; 13 fig.; 5 tab.; 10 ref.

To assess novel technologies and their sustainability potential in the food and agricultural domain, we need to understand the interrelation and complementarity of data, data models and business processes. Smart Farming techniques provide a proper ground to translate this field of research to applications that contribute to sustainability goals.

This work analyses the sustainability challenge that farmers face in their decision making process to select Plant Protection Products (PPP) supported by tools such as Farm Management Information Systems (FMIS). This work specifically analyses data models that describe the plant protection domain while it addresses the urgent need to integrate data from semantically disconnected domains and achieve interoperability. Therefore the PPP domain require an evaluation of current data infrastructure and its associated design patterns. Two design patterns for reference data models are evaluated: entity relationship diagrams and knowledge graph development. The Within Field Management Zoning use case of the IoF2020 project was selected to apply these design patterns with semantically the same data. The results are designed and proposed as follows:

1. An illustrative wireframe that puts data in context;
2. Competency questions that identifies the information need;
3. Harmonised reference data model that represents the PPP domain in UML;
4. An illustrative knowledge graph that represents the PPP domain in RDF.

The results show that design patterns that are needed to harmonise reference data comprise the following heuristics:

- Indicate the need (business, societal and environmental) with methods like wireframes and validate competency questions;
- Harmonise reference data on the semantic level;
- Scope the data model by the identification reusable candidate concepts and entities;
- Link concepts and entities with the use of knowledge graphs and semantic web specifications based on W3C standards, such as RDF.

By formally relating the specifications, the amount of time is reduced that is needed to redefine existing entities according to a new standard. Therefore we argue that the use of knowledge graphs could enable pathways for data to be interoperated. The use of knowledge graphs should address the issue of when semantically the same data are defined according to different standards by specifying how the standards are related. However, the activity to map these relations results in yet another set of conditions that need to be defined.

Key words: Sustainable Agriculture, Smart Farming, Data Harmonisation, Semantic Interoperability

This report can be downloaded for free at https://doi.org/10.18174/572900 or at www.wur.eu/economic-research (under Wageningen Economic Research publications).

P.O. Box 29703, 2502 LS The Hague, The Netherlands, T +31 (0)70 335 83 30,
E communications.ssg@wur.nl, http://www.wur.eu/economic-research. Wageningen Economic Research is part of Wageningen University & Research.

Wageningen Economic Research White Paper 2022-051 | Project code 2282300206

Cover photo: Shutterstock

Contents

List of Figures

List of Tables

Preface

Sustainability is a major challenge for agri-food systems. Total production and productivity per unit of land must increase while natural resources must be used more efficiently, and waste must be reduced. Pollution and other negative effects of e.g. the use of pesticides must be minimized or reduced to zero.

Research findings suggest that digital technologies such as Precision Agriculture, Smart Farming, Internet of Things and Artificial Intelligence in agriculture are key technologies to develop sustainable agriculture. Data and data exchange between various systems and devices play a key role in this development.

Seamless exchange of data between information systems has always been an important challenge. The rapid development of digital technologies and abundant availability of data make it even a more urgent issue. New information technologies such as semantic web and knowledge graphs are therefore welcomed to address this issue.

Therefore, I want to thank the authors of this report to deliver a proof of concept in the domain of plant protection products. It shows how these novel information technologies can enhance the development of standards and improve the process of data exchange.

I hope it will inspire the whole agri-food domain to build on this knowledge, contributing to a more sustainable food system.

Dr.ir. J. (Sjaak) Wolfert
Theme Ambassador Digital Innovation for Sustainable Food Systems
Wageningen Economic Research
Wageningen University & Research
The Netherlands

Acronyms

Abbreviation	Explanation
API	Application Programming Interface
CTGB	College voor de toelating van gewasbeschermingsmiddelen en biociden
EPPO	European and Mediterranean Plant Protection Organization
GLN	Global Location Number
GPC	Global Product Classification
GTIN	Global Trade Item Number
JSON-LD	JavaScript Object Notation for Linked Data
MDA	Model Driven Approach
PIM	Platform Independent Model
PoC	Proof of Concept
PPP	Plant Protection Product
RDF	Resource Definition Framework
UML	Unified Modelling Language
URI	Unique Resource Identifier
URL	Unique Resource Locator
VRA	Variable Rate Application
XML	Extensible Markup Language

Change history

Version	Date	To	Changes
0.4	10/2/21	AgGateway WG06 FarmInputs IoF2020 UC 1.1	First draft by WUR
0.5	29/3/21	AgGateway WG06 FarmInputs IoF2020 UC 1.1	Input from Agrimetrics
0.6	21/7/21	AgGateway WG06 FarmInputs	Input to knowledge graph, conclusion and discussion.
0.7	10/8/21	AgGateway WG06 FarmInputs	Last WG review
0.8	20/10/21	Use Case contributors & WG External review	Last comments & WUR review
0.9	8/11/21	AgGateway's Portfolio Management Center team	Clean draft, ready to publish
1.0	19/01/22	IoF2020 / Smart AgriHubs communication team	Management summary and lay-out changes

Management summary

A large team of organisations joined efforts in the Internet of Food and Farm 2020 (IoF2020) project to make use of data, facilitate decisions and optimise processes with the help of IoT technologies. IoF2020 was organised by so-called use cases that worked on such potentials. The *Within Field Management Zoning* use case defined specific field management zones that supports stakeholders to by develop and link devices that sense and actuate   with external data, mainly in potato to e.g. decrease the use of herbicides, fertiliser, or water.

To assess novel technologies and their sustainability potential in the food and agricultural domain, we need to understand the interrelation and complementarity of data, data models and business processes. Smart Farming techniques provide a proper ground to translate this field of research to applications that contribute to sustainability goals.

This work analyses the sustainability challenge that farmers face in their decision making process to select Plant Protection Products (PPP) supported by tools such as Farm Management Information Systems (FMIS). This work specifically analyses data models that describe the plant protection domain while it addresses the urgent need to integrate data from semantically disconnected domains and achieve interoperability. Therefore the PPP domain require an evaluation of current data infrastructure and its associated design patterns. Two design patterns for reference data models are evaluated: entity relationship diagrams and knowledge graph development. The Within Field Management Zoning use case of the IoF2020 project was selected to apply these design patterns with semantically the same data. The results are designed and proposed as follows:

1. An illustrative wireframe that puts data in context;
2. Competency questions that identifies the information need;
3. Harmonised reference data model that represents the PPP domain in UML;
4. An illustrative knowledge graph that represents the PPP domain in RDF.

The results show that design patterns that are needed to harmonise reference data comprise the following heuristics:

- Indicate the need (business, societal and environmental) with methods like wireframes and validate competency questions;
- Harmonise reference data on the semantic level;
- Scope the data model by the identification reusable candidate concepts and entities;
- Link concepts and entities with the use of knowledge graphs and semantic web specifications based on W3C standards, such as RDF.

By formally relating the specifications, the amount of time is reduced that is needed to redefine existing entities according to a new standard. Therefore we argue that the use of knowledge graphs could enable pathways for data to be interoperated. The use of knowledge graphs should address the issue of when semantically the same data are defined according to different standards by specifying how the standards are related. However, the activity to map these relations results in yet another set of conditions that need to be defined.

A limitation of study could be the lack of sufficient attention on the information security aspect with regard to the way URIs are organised. Another limitation could be an analysis of validation and verification methods for data models in general within the agri-food domain.

Suggestions for next steps include investigation of the prototype of the proposed design, while taking the value model into consideration. A multi-actor agile standardisation approach is suggested, with at least the following actors: farmer or farmer representative, machine manufacturer (i.e. AEF), governmental authority, plant protection products provider and a DSS provider.

1 Harmonisation on farm inputs data

Farming is the oldest industry on earth and arguably the most important. It feeds the world, employs millions and sustains life. Yet it is also often criticised because of its detrimental environmental impacts and resource consumption. For example, it emits 9% of UK greenhouse gas emissions and consumes 70% of our water.[1] The global challenge that agriculture faces is to grow more with less inputs or transform into a more circular way of production.

Farmers rely, besides on intuition, on diverse data and years of accumulated experience to produce our food. But farming food requires attention to detail and hard work despite often working with slim operating margins. To understand how we bring data and diagnostics to traditional broadacre farming so that growers can grow more produce from less, we have to compare how we make decisions on the farm today with how we will exploit data in the future to make decisions.

This report details the results of a short project to investigate how the process of informing farm decision making with data can be improved through improvements to how datasets are represented in data models. To do this, we focus on the increasingly challenging scenario of supporting farmers' decision making surrounding the application of plant protection products. As part of decision making, the plant protection domain is complex and evolving. Decision making on the usage and authorisation of plant protection products needs efficiency and flexibility of data exchange between stakeholders. Moreover, there is an increasing need to take into account data from traditionally disconnected domains.[2] For these reasons it is important to evaluate current processes and its supportive data infrastructures. Future data infrastructures should provide the efficiency and flexibility to meet the growing variety of different criteria towards achieving the business goals of different stakeholders. To evaluate the suitability of different approaches to improving data infrastructures we particularly focus on (1) the status of relevant existing data models, (2) identifying future competency questions that are likely to benefit from improving the data infrastructure, and (3) supportive technologies to meet these questions.

The work described in this report resulted from the FarmInputs initiative, which combines AgGateway's FarmInputs project (WG06) with Agrimetrics's investigation into building semantic models for Within Field Managing Zoning. The FarmInputs initiative was carried out in the period of January 2020 – April 2021 by Wageningen University & Research (WUR), Lexagri, Agrimetrics and AgGateway.

1.1 Scope and objective

FarmInputs is an AgGateway[3] initiative, with the objective to standardise and harmonise the data exchange concerning farm inputs for growing crops. This ultimately should contribute in a positive way to sustainable agriculture, track & trace of farm inputs, interoperability in data-sharing and should boost innovation. Farm inputs in this context are the main inputs for growing crops, such as: seed, fertiliser, water, energy and plant protection products.

Standardise / harmonise in this context means:
- On a semantic level: deliver reference data models (ontology, class model) for the specific data sets. Deliver recommendations in the use of identifiers such as Global Trade Item Number (GTIN), Global Location Number (GLN), Global Product Classification (GPC) and preferred classifications and code list such as the European and Mediterranean Plant Protection Organization (EPPO) standards.
- On the syntax level: specify, based on the data model, standard syntaxes, such as JSON to exchange relevant data sets.

---

[1] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/835762/agriclimate-9edition-02oct19.pdf
[2] https://www.proeftuinprecisielandbouw.nl/door-wildgroei-aan-software-komt-precisielandbouw-niet-van-de-grond/
[3] http://www.aggateway.org/

- On a business process level: promote the use of the recommended standards (e.g. the Cristal[4] recommendation for track and trace of plant protection products).

The AgGateway FarmInputs initiative is carried out by two working groups: one on reference data (WG06) and one on standardising Work-Orders, Work-Recommendations, Work-Records (WG07).

As far as the AgGateway initiative is concerned, this document relates to the first track about exchanging reference data, with a special focus on plant protection products data.

As part of the FarmInputs initiative, a Proof of Concept (PoC) was carried out. The goal of this PoC is to simulate an existing challenge, in this case the Variable Rate Application (VRA) of soil-herbicide and demonstrate the use of a harmonised reference data model with novel data integration technologies. For the farmer, this represents a change in the unit of decision making for application of herbicide changes from the traditional field to a sub-field area. In doing so the farmer can apply the appropriate levels of herbicide for the best crop outcomes, significantly reduce inputs and improve economics outcomes for their farm. Beside existing expected results, such as task maps and suggested use of products, future scenarios should provide new insights such as awareness of side-effects in our environmental footprint.

Moreover, this study considers alternative approaches of integrating different heterogeneous sources with linked data concepts and semantic web technologies. One of the global challenges faced by the agri-food tech industry is the harmonisation and validation of machine readable data. Many attempts are made to avoid formats in documents such as PDF. As a starting point for a discussion, a synthesis is given of the current approach versus alternative approaches and suggestions for next steps. This study investigates to what extent a RDF-based network (Resource Description Framework) is useful to link the different data sources and data consumer.

The Within Field Management Zoning[5] use case is a collaboration with partners who are already looking at an effective way to variably apply pre-emergence herbicides for potato growing. This is one of the IoF2020 use cases. Advances in the variable-rate application have the potential to reduce the costs of plant protection for the grower and also to reduce the plant protection footprint for potato production (Kempenaar, et al., 2017).

There is an opportunity to use the AgGateway WG06 plant protection product modelling work to provide a linked dataset that can be made available online in a convenient way (e.g. from the Agrimetrics Data Marketplace over API) and used by the use case to provide new information about plant protection products that can be linked to their own data.

At this stage, the project is focusing on how the ratio of organic matter to clay in the soil affects the efficacy of the herbicide. On the basis of dose-response standards for different soil ratios the rate of application is varied across a growing area.

In a future where precision agriculture is the norm, decisions will be made by machines on progressively smaller areas of the growing area, requiring real-time access to more higher definition data. These data will be sourced from multiple data providers in the agricultural data ecosystem. This raises a challenge for the sector on how data can be made findable, accessible, interoperable, and reusable so that it can be exploited with advances in precision farming.

The Global Open Data for Agriculture and Nutrition network (GODAN) recognised this when preparing a report for the 2016 G8 Summit ((GODAN), Global Open Data for Agriculture & Nutrition, 2016). The GODAN report sees the wider adoption of a common resource description framework as a critical step in achieving this. Achieving interoperability across multiple data sources was one of the motivators for the CGIAR to adopt the W3C Resource Description Framework (RDF) as the foundation of its CGIAR Big Data Platform (CGIAR, 2017). Early research in the SemaGrow (Wageningen Research, 2015) project has tested and shown the viability using RDF for interoperability on data-intensive agricultural systems, proposing a querying system that uses metadata about the data sources to optimise query execution.

4 https://croplife.org/wp-content/uploads/2017/02/CRISTAL-ON-BOARDING-HANDBOOK-160117.pdf
5 https://www.iof2020.eu/use-case-catalogue/arable/within-field-management-zoning

2 Material and Methods

The overall project approach consist of the following generic phases:

1. Implications of the use case Within Field Management Zoning and its effects for plant protection data for the environmental impact.
2. Modelling reference data for the domain of plant protection and list of relevant data publishers.
3. Application of the RDF approach.
4. Scoping the domain model.

Each phase is elaborated in the subsections below.

2.1 Within Field Management Zoning

The Within Field Management Zoning use case in variable-rate spraying has shown that the proposed operation technique is a viable approach in precision agriculture, with improved sustainability outcomes for the grower and the environment (Kempenaar, et al., 2017).

The efficacy of soil herbicides is influenced by soil parameters such as organic matter and clay content. Herbicides are absorbed by these soil particles and not available for herbicidal activity. Soils with higher organic matter or clay content will need a higher dosage of herbicides to effectively kill germinating weeds. However, a high dosage of herbicides on soils with lower organic matter or clay content can lead to crop damage. Making use of variable rate application, the soil herbicides can be applied by taking into account spatial variation of the soil. This can be done with the use of a soil map, describing this spatial variation in clay or organic matter content. These maps can be drawn by hand with the use of software and experiences. But there are also multiple companies providing soil scans. Most soil scans are performed physically on the field, but other providers create a soil map based on satellite imagery, historical soil maps and farmer experience.

With the use of decision support rules, a soil map can be translated into a variable dosage map. These rules can be used with, for example, the Akkerweb Herbicide application. This application combines the soil map and the type of soil herbicide to create a variable dosage map with the optimal dosage for each organic matter or clay content on the field. The farmer combines this dosage map with the routing, spray volume and spray boom width to create an operation map: a map describing what dosage to apply where. The operation task map can be downloaded in the right format for the specific sprayer in Shape or ISO-XML format. The board computer of the sprayer can read this operation task map and execute the variable rate application of the soil herbicide. This will lead to more efficient use and reduction of soil herbicides, a higher yield through less damage and lower environmental impact.

2.1.1 Environmental Impact

With new rules, based on the same data the environmental impact could be incorporated in farmers' decision making processes to protect the crops from diseases and pest. Currently the use case makes use of The Pesticide Yardstick.[6]This tool provides insights on the environmental impact points of each pesticide which is permitted in the Dutch market. The impact is based on the active ingredients in plant protection products and based on the dosage and the percentage drift. The tool makes an estimation on how much of these ingredients come into contact with the environment. Additionally, location (field, greenhouse, ditch-side) of usage is also relevant for indicating the type of impact on e.g. birds, butterflies and bees. The tool expresses the result in environmental impact points divided into the effect on aquatic organisms, soil organisms and groundwater. More impact points implies more damage to the environment.

6 https://www.pesticideyardstick.eu/

More recently, a project has been working on a follow-up to the Pesticide yardstick. The project Environmental Indicator Crop protection[7] aims to create an open calculation tool, which can be implemented in certification schemes. The new values, transparency and objectivity should go in hand with the acceptance and support of the market, sector, and society in an international environment. The tool should be based on international standards to be used for certification authorities, in collaboration with individual users and government entities.

2.2 Reference data modelling

One of the objectives of the FarmInputs initiative was to improve the plant protection part of the overall data reference model Agro (drmAgro) and deliver a standard reference data model that can be used as a basis for designing APIs. The following approach was used:

- Take the Homologa data model, College voor de toelating van gewasbeschermingsmiddelen en biociden (CTGB) data model + drmAgro data model to draft a reference class model for PPP reference data (master data), using Enterprise Architect as a tool.
- Transfer the class model from EA to TopBraid and work out the RDF details in TopBRAID.
- Evaluate the options of publishing an ontology on the internet, using the TopBraid RDF model as a source.
- Evaluate the options of specifying different types of interfacing (REST-API, GraphQL, etc.) using the RDF model as the source.

2.2.1 Data publishers

Data were sourced from two production databases, which have been created from product labels. Homologa, a dataset published by LexAgri, and the UK Pesticide Guide (UKPG), published by the British Crop Production Council (BCPC) and National Institute of Agricultural Botany (NIAB). For the Netherlands these are: CTGB and Nefyto (Dutch plant protection association). Additionally, there are case-specific datasets derived from systems which are in use by the farmer, such as field data, soil maps and crop-specific data. Derived from the decision support system, the dosing algorithms are also a dataset in scope. Characteristics of these datasets that aligned with concepts identified in the competency questions were the starting point for the domain modelling.

Table 1 *Overview of datasets considered as relevant.*

Dataset	**Publisher**	**Country**
Homologa	LexAgri	UK
UKPG	BCPC & NIAB	UK
Field data	Farmer	NL, Abbenes
Soil maps	Farmer	NL, Abbenes
Crop specific	Farmer	NL, Abbenes
Dosing algorithms	FarmMaps	Global
Dutch authorisations	CTGB	NL
Dutch authorisations	Nefyto	NL

Domain modelling was started with a sketching tool in co-operation with domain experts. After an initial design a first-pass RDF model was created using an RDF modelling tool. Various RDF modelling tools exist: Protégé is a commonly used tool and open source; TopBraid Composer from TopQuadrant is a commercial modelling tool that can be used to build RDF models.

[7] https://www.wur.nl/upload_mm/8/b/0/2c86316a-2441-4581-91ca-286db1df6e97_Gewasbescherming en milieu-impact.pdf

2.3 Resource Description Framework (RDF)

There is an increasing desire for data to be well described so that it can be read by humans and machines (Medicine, 2019). The greatest impedance to data interoperability is badly described data and data descriptions which require a human in the loop (HIL). The W3C Resource Description Framework gives an approach so that data that is schema-less (so universally readable), where the data and the meta-data are virtually co-located providing an apparent single knowledge graph.

The RDF structure is made of triples, an atomic unit of expression within an RDF graph. Each triple is comprised of a subject, predicate (or relationship) and object. A facile example for the spreadsheet cell is presented in Figure 1. The product 'VAYANTIS' has a seed coat colour of 'Blue'.

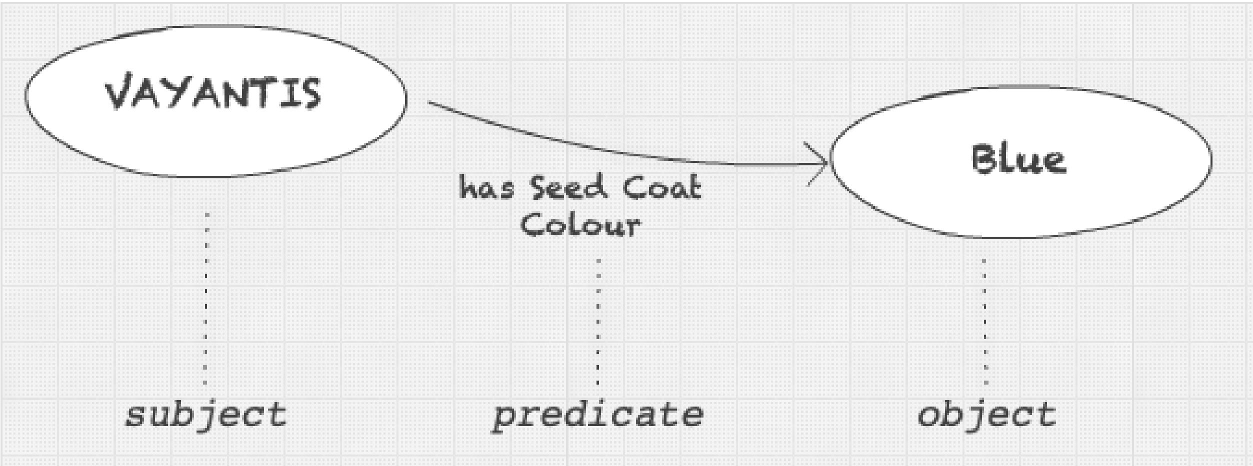

Figure 1 *Example Vayantis in spreadsheet cell.*

Product	Seed Coat Colour	Crop
VAYANTIS	Blue	Soya
REDIGRO PRO	Red	Wheat

Figure 2 *Example Vayantis as a triple.*

Using the RDF framework this would be expressed as a triple as shown below in Figure 2.

The item VAYANTIS (represented by a unique identifier) becomes the subject of the triple; the property, in this case Seed Coat Colour, becomes the predicate of the triple, and the data value becomes the object of the triple. Each of the triple's components are referred to as resources. Resources can continue to be linked together and the whole dataset can be expressed as graph. For example, the remainder of this data would form the knowledge graph below in Figure 3.

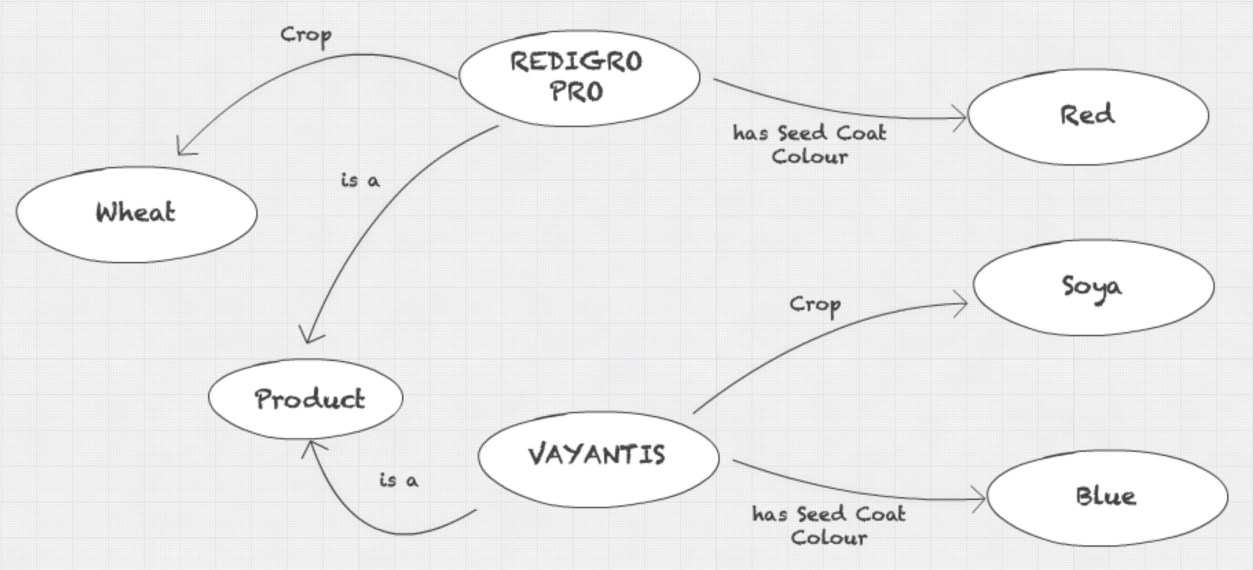

Figure 3 *Example Vayantis in a knowledge graph.*

A knowledge graph has no schema to consult before reading data. The format is homogenous throughout the dataset, and the meta-data (the data describing the data) is accessible within the graph, in this case the triples stating that REDIGRO and VAYANTIS are products.

2.3.1 Location Irrelevance

One of the fundamental features of an RDF knowledge graph is that all resources can be identified with a Uniform Resource Identifier. This is similar to a unique identifier or a primary key that is used in relational database systems, except that a URI is globally unique and extends the format defined by the W3C for Uniform Resource Locators (URLs). Describing entities in data and data models with the use of URIs means that data is better findable on the web with persistent identifiers. Such data could be published on a web server so that is it is dereferenceable – retrieving the data at a URI from the web yields information about that item, including links to other items identified by URIs which can themselves be dereferenced. This allows data to be distributed across multiple providers and authorities and yet still remain accessible and interconnected.

2.3.2 Objectives of the RDF Modelling

The objectives of the RDF modelling are:

1. Create a domain model for Plant Protection Products (PPPs), which can be serialised as an RDF graph to provide meta-data for plant protection product datasets.
2. Test the PPP domain model, as serialised RDF, using sample data.
3. Exemplify a suitable query of RDF data, described using the PPP model, that returns data fit-for-purpose for a precision farming operation such as prem-herbicide variable rate application.

Outputs from the RDF modelling:

- An RDF model that can be used to describe data held in different reference databases.
- Sample competency questions that represent what an expert or machine-agent would ask of the data.

2.4 Scoping of the model and knowledge graph development

With traditional data systems the data architect must define a universe of discourse – the boundaries in which data will be modelled and created and this is then used to define the conceptual model of the domain (the schema) in which ever proprietary language is chosen. With an RDF knowledge graph all concepts can be linked. It is therefore particularly important that any project defines a domain scope, or risks being overly broadly scoped, defining terms and relationships that provide no value or that take excessive time to create.

For scoping the model, the following questions were used:
- Which concepts in the PPP data are needed?
- Which concepts in the VRA data are needed?
- Which of the candidate concept definitions can we reuse?

To define the universe of discourse within the PPP domain, scoping the model was an essential step for making the boundaries clear in which the data will be modelled. This is an essential part of the design of an information system, where the conceptual model of the domain – a formal definition of what things are and how they are related – is developed. One of the tools available for doing this, is the RDF modelling paradigm. An RDF model specifies the concepts or collections of things, and relationships ('classes', and 'predicates' respectively, in RDF terminology) which are used in a field of knowledge. Further information can also be represented within the model, such as subclass/superclass relationships, and the relationships which are used with each class. In many respects, the RDF approach resembles the UML approach described in the previous section, to the extent that a mechanical transformation between the two approaches is feasible, although yielding poor-quality results. However, there are significant technical differences between the two models, particularly around the 'closed-ness' of the model, where RDF is more open to extension, but it is consequently easier to introduce incompatibilities into the data. In fact, a core principle of RDF modelling is that it allows for overlapping, interconnected and reusable models of different knowledge domains.

3 Results

Main results are elaborated in this section within the following order: (1) a harmonised reference data model that is the result of a participatory design session with key-actors for Plant Protection Products (PPP) data, (2) competency questions that are defined to identify specific information need, (3) a wireframe that is proposed based on the use case description, and (4) an initial knowledge graph that could provide increased interoperability on the domain layer.

3.1 Wireframe as a means to put data in context

As stated in Chapter 1, the objective of this PoC was to standardise and harmonise farm inputs on three levels: (1) semantic level, (2) syntax level and (3) a business process level. Figure 4 presents the results obtained from the preliminary analysis of scoping the model and developing the competency questions as reference data. These reference data are summarised in a wireframe, a hypothetical user interface.

The most interesting aspect of the wireframe is that it puts data in context and includes relevant PPP data and as well as the VRA calibration data. The VRA calibration data is the dosing algorithm for doing the calculation. This wireframe shows the user journey in formulating the competency questions presented earlier. The location is determined by the chosen cropping scheme from Akkerweb (Akkerweb, sd).

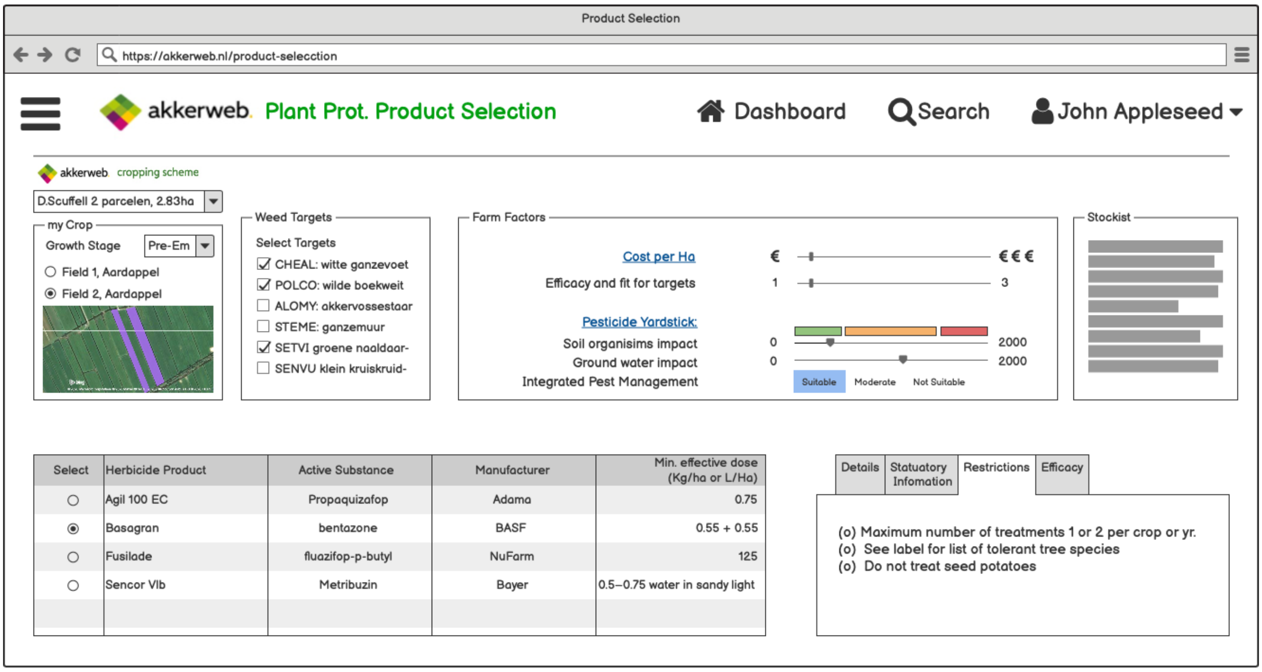

Figure 4 *Wireframe VRA Plant protection product selection.*

3.2 Competency questions for identifying the information need

Competency questions help to define the scope of projects to keep modelling activities relevant and appropriate to the recognised business value that the knowledge graph will deliver. A competency question is any question that an expert (human or machine-agent) would ask of the data. Competency questions fulfil three major needs on a project:

1. They provide a means to scope the project data.
2. As they have been generated by domain experts, they identify business value.
3. A competency question provides a means to test the linked knowledge graph through queries based on the competency question.

The two competency questions for this project are illustrated in Table 2. Analysis of the competency questions is used in scoping the project by determining the main concepts within scope, along with properties (predicates) that are needed for the knowledge base.

Table 2 *Competency Questions.*

Competency Question	**Candidate datasets**	**Domain Concepts**
Which <plant protection product> can I use at <Growth Stage> for my <crop type> crop that target <this weed/pest>	Homologa, UKPG, EUPD, BBCH	PP_product, growth_stage, crop, crop_type, target, weed,pest
Get me all products that contain only <chemical substance> (as their active substance) and get me their maximum application rates that are permitted for use in <my crop type>, targeting <my weed>	ChEMBL, ChemiSpider, UKPG, Homologa, EPPO	Biological_species, biological_group, product_application, application_rate, target_weed, chemical_substance

3.3 Harmonised reference data model: a multi-actor collaboration

The harmonised reference data model is specified for the PPP domain within WG06 of AgGateway (Plant Protection Products domain, part of drmAgro (Domain Reference Model Agro), 2020). Mainly three sources are considered to be relevant for the development of the harmonized reference data model: (1) the Homologa data model, (2) a composed sub-model of the reference model Agro (Wageningen Econmic Research, sd), and (3) EU policy documents concerning regulation of plant protection products.

The following classes below are examples that should provide understanding on how the harmonised reference data model is developed.

Use conditions

The class *PlantProtectionProductUseCondition : Public Class* is based on the “Data Dictionary for Plant Protection Products Data Migration” and contains important information on candidate attributes and associations. However, from a domain modelling perspective, it is questionable whether this use conditions should be split up into the following categories,for example, *PlantProtectionProductTargets, PlantprotectionProductBefinificaries,* or *PlantprotectionProductRestrictions.*

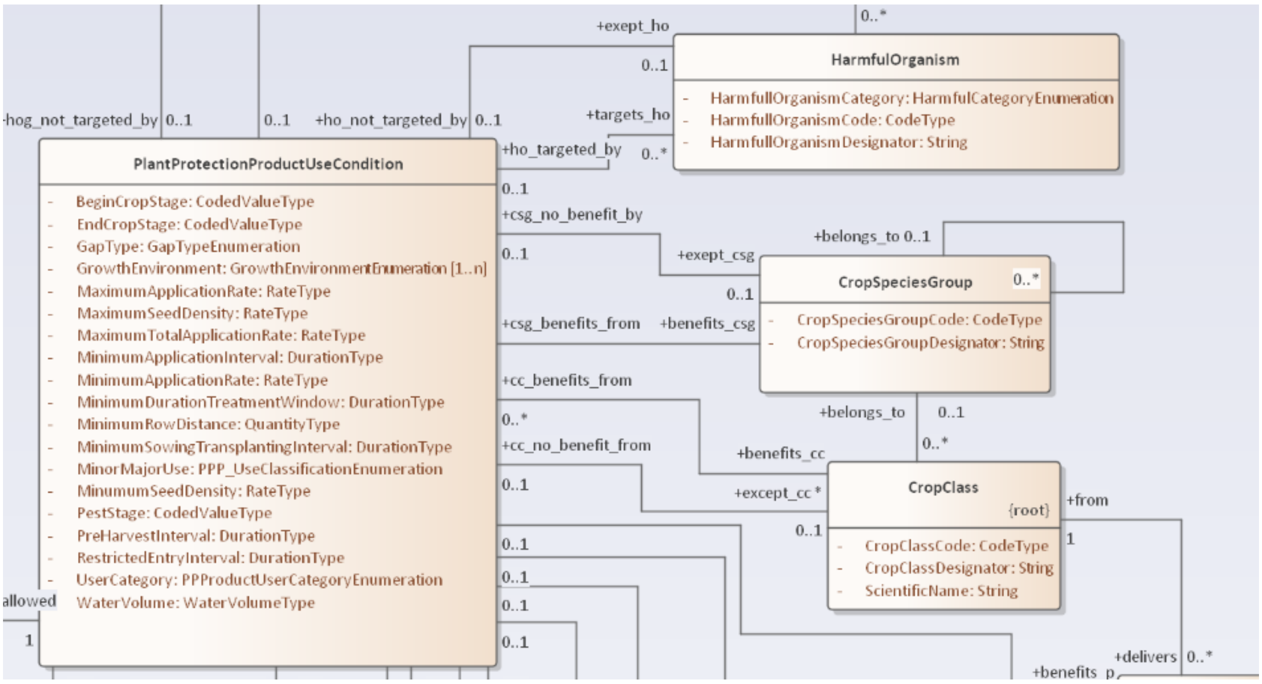

Figure 5 *Part of the reference data model.*

In Figure 5 a subset of the model is presented.
The model is aimed to be interoperable with other parts of the reference model drmAgro by reusing existing entities and attributes. Surprisingly, an essential part of the reference data model is an overview of the enumerations which might become more common in the future since list of crop species and harmful organisms are essential to determine use conditions based on regulations.
Part of the proposed class diagram for enumerations is shown in Figure 6. In Appendix 1 the complete reference data model and enumerations can be found.

Seed treatment
In *PlantprotectionProductUseConditions* several attributes address the use of plant protection products as seed treatment. These results, therefore, need to be interpreted with caution since effort is needed to translate the rate (kg/kg) of PPP in a batch of seed to the rate (kg/ha) of that product on the field. As expected, this use case contains an algorithm that requires some analyses and might result in renewed specified PPP and its use conditions.

Plant species as weed.
A *PlantProtectionProduct* , in general, targets a *HarmfulOrganism*, which is specified by a number of coding lists. One of the *HarmfulOrganism* categories is WEED, that consists of two possibilities of specification.One possibility is weed as a harmful organism and simply use the plant species code as code for the organism, while the other possibility is to make associations to *PlantSpecies* and give them the predicate “target”. The latter one imply weed not as a harmful organism.

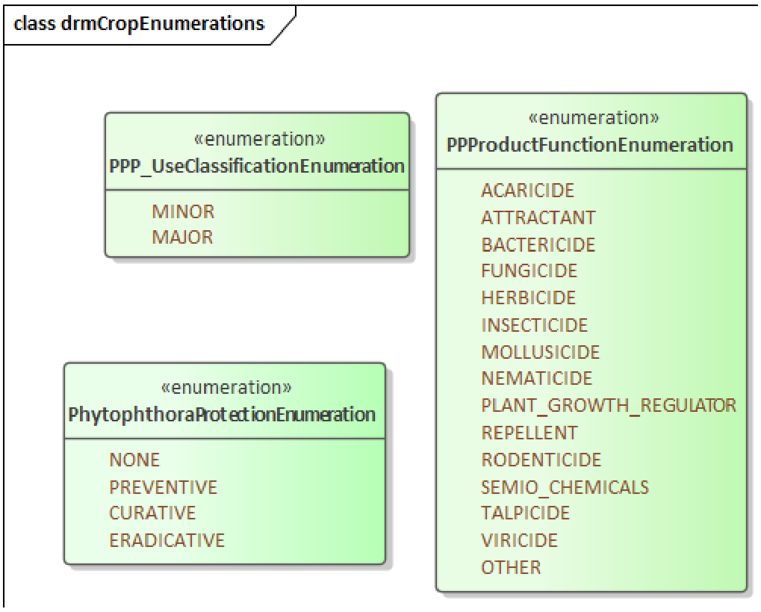

Figure 6 *Enumerations for plant protection reference data.*

Plant protection products require a number of code lists and identifier lists. It is not always clear which worldwide operating authority is, or should be, responsible for an international list. Where this is clear, like for example ISO for country codes, such a list is not always published following the FAIR principles. In this example there are multiple organisations publishing an equivalent. Some lists will stay on a national or regional basis, like for example growing periods. For the example of OperationTechnique a table is presented in Appendix 2 'Example reference data'.
What stands out in the reference data model are classes that represent reference data. One of the objectives of this PoC was to deliver recommendations on the use of identifiers, classifications and code lists. Further analysis showed that classes, as presented in Table 3, represent reference data. From this data we can see that still for most of the classes it is not clear what the source is or what it should be.

Table 3 *Classes that represent reference data from the reference data model.*

Class Name	Type of reference data	Standard Development Organisation	Comment
ApplicationSeason	Codelist (National)		
AuthorisingOrganisation	IdentifierList		
BBrand	IdentifierList		Required as reference data
Country	CodeList	ISO	
CropProductionperiod	CodeList (National)		
CropProductionPurpose	CodeList		
CulturalPractise	CodeList		
GrowthEnvironement	CodeList		
GrowthMedium	CodeList		
HarmfulOrganism	CodeList	EPPO	
harmfulOrganismGroup	CodeList	EPPO	
InfrastructureCategory	CodeList		
MaximumResidueLevel	CodeList		Junction table between Substance and Produce (Authorising organisation)
Manufacturer	IdentifierList		Required as reference data
OperationTechnique	IdentifierList	AEF	
PlantGroup	Codelist		
PlantGrowthAspect	IdentifierList		
PlantProtectionProduct	IdentifierList		Trade name – based on the registration number
PlantSpecies	Codelist	FAO, EPPO, GS1, VBN	Required as reference data. VBN for Dutch floriculture
Produce	IdentifierList	GS1	
ProductGroup	CodeList	Floricode	For categorisation within the floriculture, matched with the HS (Harmonised System)
Substance	Identifierslist		Active Ingredient
TradeMark	IdentifierList		
Variety	Identfierslist		

Aligned with the initial objective for RDF modelling, in this study an attempt is made on the automatic generation of an RDF model based on the UML model. This attempt faced the challenge of the merits of Enterprise Architect for transforming schemas. For example, some parts can be transformed with the OGC standard, while some other parts are not suitable. An example triple of the class from the automatic transformed RDF is shown below:

```
rmAgro:PlantProtectionProductUseCondition
 rdf:type owl:Class ;
 rdfs:comment "The legally required conditions or terms of use under
which the Authorisation of a plant protection product is valid" ;
 rdfs:label "PlantProtectionProductUseCondition" ;
 .
```

Figure 7 *Example triple of a PPP class expressed in Turtle (Terse RDF Triple Language).*

3.4 Knowledge graph development by scoping the data model

The full RDF expression of the PPP model is shown in Appendix 3, but the core concepts of the model are:

- Product – the formulated mixture of substances being sold under a given name.
- Authorisation – (part of) the set of rules under which the product may be used in a particular regulatory environment.
- Beneficiary and Outcome – the crop type, group of species, amenity or environment which is improved by the application of this product, and the intended effect on that beneficiary.
- Target and Effect – the organism or group of organisms on which this product is intended to have an effect, and that effect itself.
- Packaging – the way in which the product may be packaged for sale.
- Active ingredient – the chemical substance (or substances) in the product which cause the primary effect on the target.

Each Authorisation contains a set of rules for how the product may be applied: for example, how frequently, how many times in the lifetime of a crop, in what concentration (per square meter), how soon after application harvesting is allowed, at what stage(s) in the lifecycle of the Beneficiary or Target. The Beneficiary, in this model, may be more than just a plant species, as a PPP may be applied in order to benefit something which is not a crop (for example, a topical herbicide to remove weeds from a golf course, or to keep a rail line clear of growth). The Target may be a distinct species, collection of species (e.g. *fusarium*), or may even be the same as the Beneficiary (for example, plant growth regulators have the same Target and Beneficiary). While the Beneficiary is generally a legislative constraint – you cannot use the product on a beneficiary it's not regulated for – the Target is not, as that would prevent, say, a herbicide targeting one kind of weed from being applied if there is also a second kind of weed present.

Note that the Packaging part of the model here requires further work to support details of packaging for shipping (cartons, boxes, palettes, shipping containers). The detailed Packaging/shipping model was considered out of scope for this investigation. Figure 8 shows the conceptual PPP model. In Appendix 3, Figure 13, the model is presented in the context of a specific product application scenario.

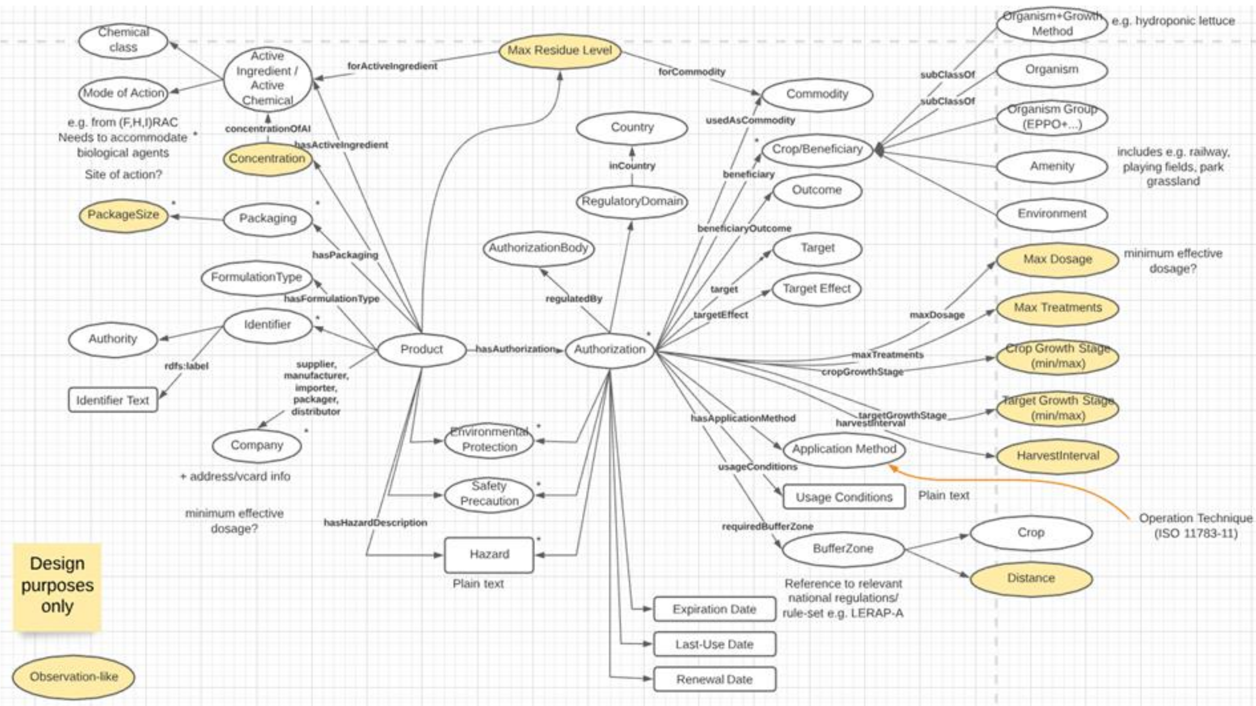

Figure 8 *PPP Knowledge graph - candidate concepts, classes (ovals), properties (boxes) and relationships (arrows).*

4 Discussion

This PoC was set out with the aim to assess the importance of reference data modelling as a way to support connecting data from multiple data domains. Taking this aim into consideration, semantics is assumed to be crucial for to identify the particular relevant domain concepts and data model entities.

The results of this PoC show that harmonisation of reference data comprises the following main activities: (1) indicate the need (business, societal and environmental) with methods such as wireframe and validate competency questions (2) harmonise reference data on the semantic level, (3) audit the data model to identify reusable candidate concepts and entities, and (4) link concepts and entities with the use of knowledge graphs based on W3C standards, such as RDF.. These activities are presented as building blocks in Figure 9.
Main results are elaborated in this section in the following order: (1) a harmonised reference data model that is the result of a participatory design session with key-actors for Plant Protection Products (PPP) data, (2) competency questions that are defined to identify specific information need, (3) a wireframe that is proposed based on the use case description, and (4) an initial knowledge graph that could provide increased interoperability on the domain layer.

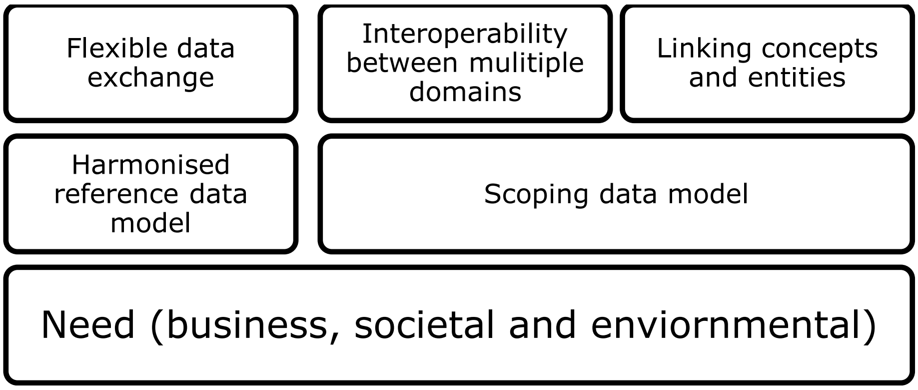

Figure 9 *Building blocks for harmonising reference data.*

Tool skills dependency
Surprisingly , the dependence of a modeller with regard to the skills of a tool such as Enterprise Architect results in a challenge to apply a certain design pattern. For example, we argue that design patterns Model Driven Approach (MDA) and Platform Independent Modelling (PIM) are important drivers for interoperability. Within the tool Enterprise Architect, there are tool-specific transition- and schema-generating templates, which require adaption capabilities of those templates or even writing new ones. This requires capabilities both from human side as well as tools functionality.

Use of existing standards
The dependency of skills to model certain domains or tool specific characteristics seem not to be the only challenges that emerged from findings of this study.
Another dispute that emerges from findings of this study is the the frequent use of existing standards that are platform specific, which prevents the adoption of preferred design patterns . We faced the problem that most of such standards are already platform specific. Examples are Geographic Markup Language (GML) for geometries and business entities such as those from UNCEFACT and Universal Business Language (UBL), all expressed in XML. This finding raises intriguing questions regarding the nature and the extent of MDA and PIM to which a model needs to be transformed to a platform specific version. A possible explanation could be that standards do not necessarily always follow the principle of using existing standards themselves. This results often in redefined basic elements, such as date and time, or geometries, such as point and line string,

while they are already defined as XML basic data type or as geometry by the Open Geospatial Consortium (OGC).

Requirement for complete language dependency
The use of existing standards implicates also the completeness of existing standards. In order to model a class defined as UML according the PIM principles, the UML specifications to be complete. Often role names and predicates are left out in associations, while they need to be specified to properly transform the model to RDF. When OWL needs to be used with defined complex relationships or precise constraints, the Object Constraint Language (OCL) need to be used for possibilities which are an addition to UML. The OCL constraints in the templates could be used to convert to the appropriate RDF/OWL constraints.

To harmonise data in a global level for crops brings several challenges. A challenge that stands out is the different type of crop use which asks for different growing conditions. For example, the crop strawberry has two uses as a commodity: as a fruit or as a salad. When strawberries are grown for salad use, residues could be different. This also applies to other crops, such as beetroot, baby leaf or forage. Other plant examples are sunflower within floriculture or oil production and edible flowers such as artichoke, pansies and roses. Therefore, a classification of usage is performed by the authors of this study. As seen in Table 4, the way information on strawberry is represented, for example in France, brings several implications: (1) column 'Crop original' provides attributes (metadata), (2) translated terms to other languages is done through a dictionary with unique identifiers and (3) this specific data point is connected to PPP reference data like MRLs (e.g. active ingredients) and logistics (e.g. trade names and registration numbers) through the same unique ID.

Table 4 *Example of different uses of strawberry as a commodity leads to harmonisation challenges.*

Crop group	**Crop (commodity)**	**Crop info**	**Crop original**
AROMATIC-AND-INFUSIONS	STRAWBERRIES: LEAVES (Fragaria ananassa)	DRY/OUTDOOR	Infusions séchées (feuilles et fleurs) de plein champ
		DRY	Infusions séchées (feuilles, fleurs, racines)
		OUTDOOR	Infusions de plein champ
		PROTECTED	Infusions sous abri
		-	Infusions sauf fines herbes et herbes aromatiques
FRUIT: BERRIES&SMALL-FRUITS	STRAWBERRIES (Fragaria ananassa)	-	Fraisier
		OUTDOOR	Fraisier de plein champ
		GLASS-HOUSE/UNDER-GLASS	Fraisier sous serre
		NURSERIES/SEEDLINGS	Traitements généraux
		PROTECTED	Fraisier sous abri
		NON-SOIL-BOUND/PROTECTED	Fraisier sous abri hors sol

In addition to previous implications, from a regulatory perspective PPP's needs to be curated cautiously to harmonise both products and chemicals as active ingredients. When it comes with the list provided by the EU Commission in which allowed PPPs are defined, rather chemicals are mentioned than products (EUR-Lex, 2021). These results therefore need to be interpreted with caution while PPPs are usually formulated with more than one active ingredient. Accordingly, if a PPP needs an approval, all active ingredients must be included in this list. Nonetheless, this process could become complex since some chemicals are particularly allowed for organic production under certain conditions of use and requirements (Homologa, 2021).

As stated previously, modelers that describe data in such a way that it is readable by humans and machines is becoming important for data interoperability. The RDF standard, published by the W3C, provides an approach for being able to do this by creating knowledge graphs. An initial objective of the project was to identify the differences, in terms of advantages and disadvantages of domain modelling as knowledge graphs and standards being described as class diagrams in UML.

In general, much time and effort is spent on specifying models for achieving higher interoperability. For example, to communicate and transfer data objects (instead of real objects) the objects need to be specified. Abstract models are used to specify the dataset, which is stored in memory or a database. After a serialisation process these datasets and models are transformed into specific formats, such as JSON, RDF/Turtle, XML, etc. Subsequently, the serialised data files are used to build applications and software. In Figure 10 an overview of these different layers is presented.

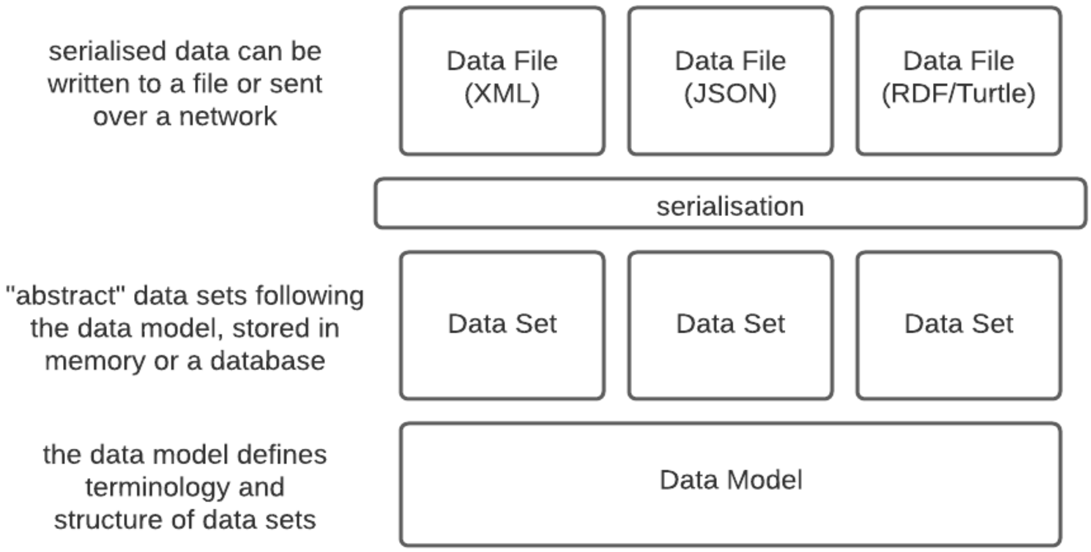

Figure 10 *Different layers of interoperability and verification.*

5 Conclusion

Summarised, this PoC presents the main characteristics of an RDF modelling paradigm in combination with reference data modelling with UML to represent the domain of plant protection products data. To test the domain model, as serialised RDF, sample data is being used from the UKPG dataset. However, one limitation of this PoC is the lack of example queries of the RDF data that return usable data for the precision farming operation of variable rate application for herbicide. Due to time constraints only the RDF model (and sample ingested data) is available for analysis and future work to apply and demonstrate the possibilities. The RDF model is intended to be able to describe the data held both in Homologa and the BCPC UKPG. A summary diagram of the complete model is included in Appendix 3, along with a sample showing the implementation of the model for a selected plant protection product.

Additionally, the sample competency questions are presented, which illustrates what an expert (either machine or human) would like to ask of the data. These questions supported the elicitation of concepts of the domain model.

To overcome our agricultural challenges, there is a need for an architectural tool which supports a data marketplace. Based on experiences within the IoF2020 project and the Working Group, main components of such a tool are: (1) Minimal Interoperability Mechanisms (MIMS), (2) a Linked Open Data Cloud to overcome physical boundaries and (3) alignment of different types of data models. For this use case these were: ADAPT as an established standard, the reference model expressed as UML and the knowledge graph that promises to make data machine and human interoperable. The main finding is the aspect of managing standardisation of installed bases for the last decades towards interoperability by design nowadays.

For the reference model, the main challenge was to cope with the complexity of different identifiers lists and coding lists. An example is the different operation techniques and cultural practices as published within AEF and AgroConnect. As shown in Appendix 2: Example reference data, there are specific types of cultural practices and operation techniques available as coded lists. The challenge is to integrate this dataset with the data ecosystem of the use case, which consists of different databases. A federated querying mechanism, which an RDF modelling paradigm provides, could provide a solution for this challenge. However, this PoC did not detect any evidence for the advantages in terms of costs and benefits. Questions like 'What are the business and value models for each stakeholder in the data-ecosystem of the PoC design?' remain unclear.

Concluding, we argue that the use of semantic web specifications could enable the interoperability of data. In this PoC, we were able to use semantic web specifications to indicate how data can be interoperated, by formally relating the specifications. This results in another set of requirements that need to be defined, i.e., how different things are related. However, it reduces the amount of time to redefine everything according to another new standard. It should deal with the problem of when data are defined according to different standards by defining how the different standard specifications are related.

A possible limitation of this PoC could be the lack of a deliberate information security approach which applies specifically for the way data is organised with URIs. Another limitation could be the validation and verification possibilities of the model within the agri-food chain.

A suggestion for next step could be to search for the prototype of such a design, while assessing the value model in mind. For this follow-up, it is recommended to adopt a multi-actor approach with the involvement of at least the following actors: a farmer's or farmers' representative, machine manufacturer (i.e. AEF), a policy/governmental institution, a plant protection products provider and a DSS provider. The possible activities could consist of mapping of different standards from SDOs such as EPPO, VBN and GS1.

References

(GODAN), Global Open Data for Agriculture & Nutrition. (2016, August). *A Global Data Ecosystem for Agriculture and Food.* Retrieved from Godan: https://www.godan.info/sites/default/files/documents/Godan_Global_Data_Ecosystem_Publication_lowres.pdf.

Akkerweb. (n.d.). *Akkerweb zet data voor je aan het werk*. Retrieved from Akkerweb: https://akkerweb.eu/en-gb/.

CGIAR. (2017). *About the platform: Using big data to solve agricultural development problems faster, better and at greater scale.* Retrieved from CGIAR: Platform for Big Data in Agriculture: https://bigdata.cgiar.org/about-the-platform/.

drmAgro (Domain Reference Model Agro). (2020, October 18). Retrieved from wecr.wur.nl: https://www.wecr.wur.nl/RM_AgroModels/SubModel PPP/index.htm.

EUR-Lex. (2021, 1 1). *Commission Regulation (EC) No 889/2008 of 5 September 2008 laying down detailed rules for the implementation of Council Regulation (EC) No 834/2007 on organic production and labelling of organic products with regard to organic production, labelling and co*. Retrieved from EUR-Lex Access to European Union Law: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32008R0889#d1e32-36-1.

Homologa. (2021, 6 23). *Which crop protection products are allowed in organic agriculture in the European Union (for both domestic production and import)?* Retrieved from Homologa.com: https://homologa.com/blog/uncategorized/which-crop-protection-products-are-allowed-in-organic-agriculture-in-the-european-union-for-both-domestic-production-and-import/.

Kempenaar, C., Been, T., Booij, J., Evert, F. v., Michielsen, J.-M., & Kocks, C. (2017). Advances in Variable Rate Technology Application in Potato in The Netherlands. *Potato Research*, 295-305.

Medicine, N. A. (2019, September 11). *Implementing Fair Data for People and Machines: Impacts and Implications*. Retrieved from Nationalacademies: https://www.nationalacademies.org/event/09-11-2019/implementing-fair-data-for-people-and-machines-impacts-and-implications.

Wageningen Econmic Research. (n.d.). *About Us*. Retrieved from rmagro.org: https://rmagro.org/.

Wageningen Research. (2015, October 31). *SemaGrow: Data intensive techniques to boost the real-time performance of global agricultural data infrastructures.* Retrieved from Research WUR: https://research.wur.nl/en/projects/semagrow-data-intensive-techniques-to-boost-the-real-time-perform.

Appendix 1 PPP Reference data model

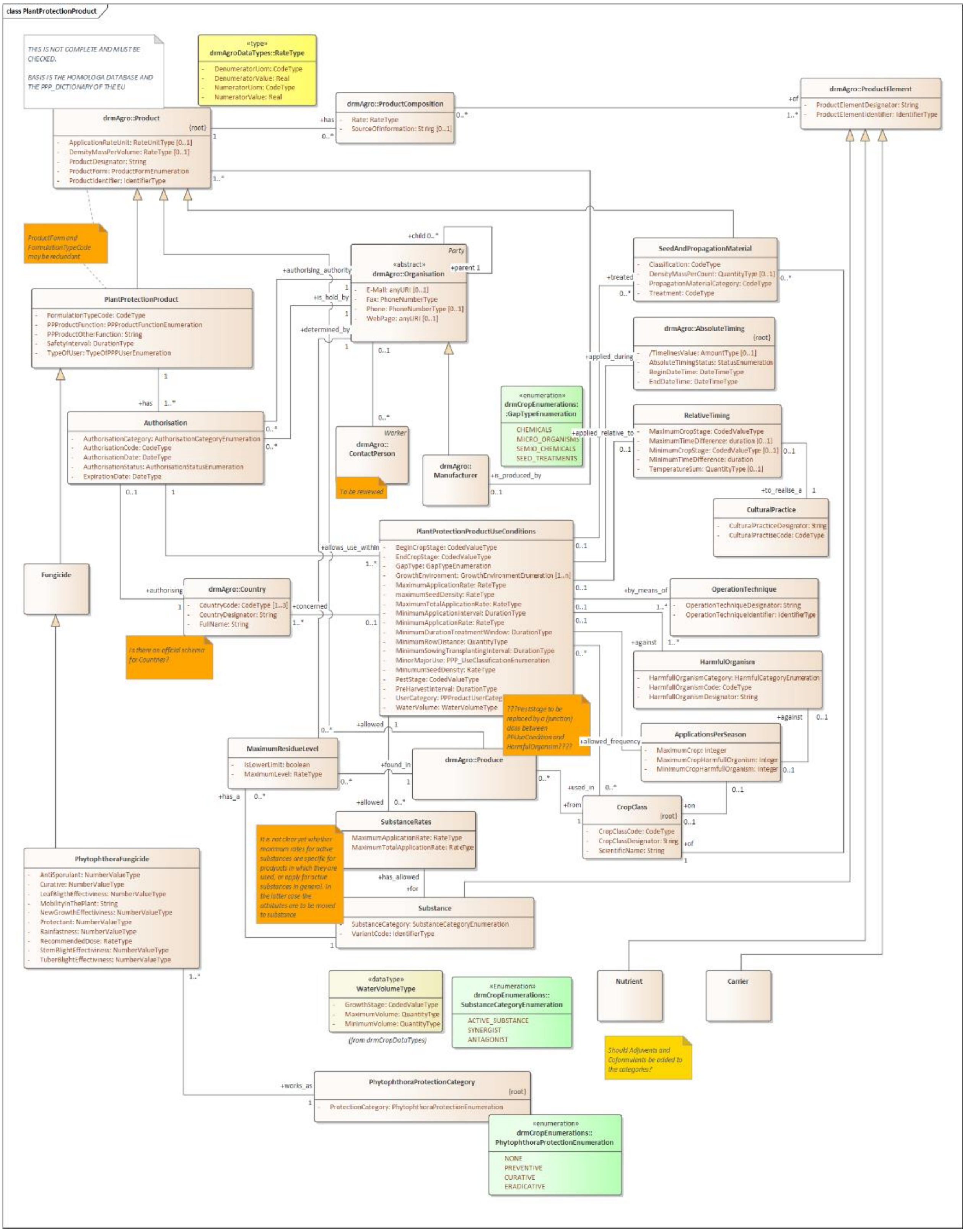

Figure 11 *PPP Reference data model.*

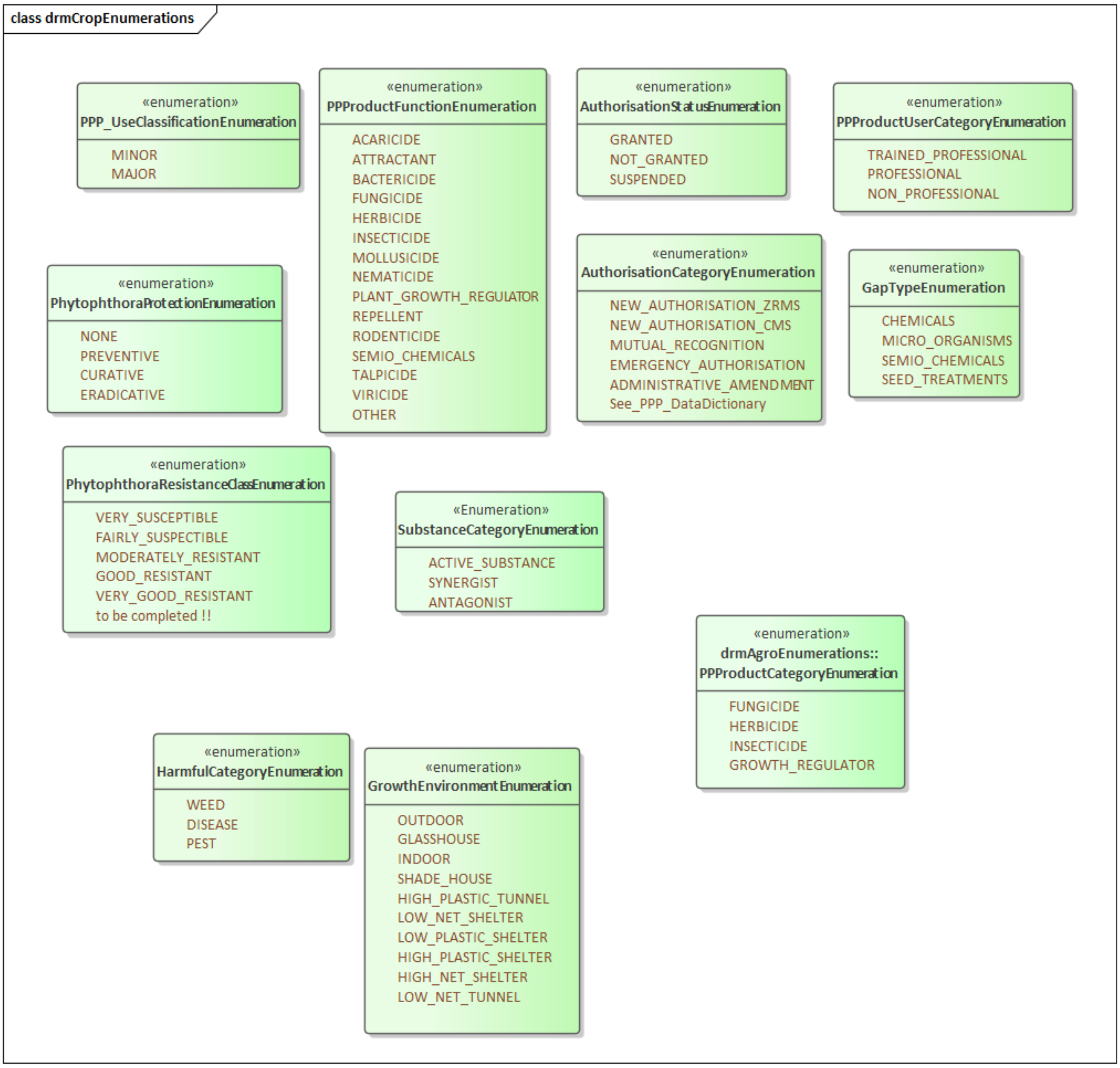

Figure 12 *Enumerations of the PPP Reference data model.*

Appendix 2 Example Reference Data

Table 5 *Example reference data for operationTechnique and culturalPractice including two levels of semantics of these classes that represent two different tables.*

CL291	**CL293**	**Definition of CulturalPractise**	**CL292**	**Definition of Operation Technique**	**remarks**
Cultural Practice	**Cultural Practice**	**Level 2**	**Operation Technique**		
Level 1	**Level 2**	This is not complete		This is not complete	
Administration					
Administration	Evaluation		Farm management information system		
	Evaluation		Manually		
Administration	Financial and technical administration		Farm management information system		
	Financial and technical administration		Manually		
Administration	Planning		Decision support system		
	Planning		Manually		
Crop care/conditioning			-		
Crop conditioning	Chemical foliage removal		Broadcast spraying		
Crop conditioning	Chemical foliage removal		vision based selective spraying		
Crop conditioning	Flower removal		manual flower picking		
Crop conditioning	Flower removal		manual		
Crop conditioning	Foliage burning		haulm burner		
Crop conditioning	Mechanical foliage removal		flail haulm pulveriser		A validation is needed whether it is haulm or foliage.
Crop conditioning	Mechanical foliage removal		haulm puller		
Crop conditioning	Mechanical foliage removal		manual pulling/picking		
	Pruning	*removing loose, infected or dead plant parts.*			
Crop conditioning	Pruning		hand secateur		
Crop conditioning	Pruning		hand pruning saw		
Crop conditioning	Pruning		hand pruning shear		
Crop conditioning	Pruning		hand pulling/picking		
Crop conditioning	Ridging		Ridger		
Crop conditioning	Ridging		rotary ridging		
Crop conditioning	Root shaping		manual by spade		

CL291	CL293	Definition of CulturalPractise	CL292	Definition of Operation Technique	remarks
Cultural Practice	**Cultural Practice**	**Level 2**	**Operation Technique**		
Level 1	**Level 2**	This is not complete		This is not complete	
Crop conditioning	Tendril		stick		
Crop conditioning	Tendril		wire		
Crop conditioning	Thinning		manual thinning		A validation is needed whether it is weeding or thinning plants?
Crop conditioning	Thinning		vision based mechanical thinning		
	Trimming	*cutting back overgrown plant parts*			
Crop conditioning	Trimming		hand secateur		
Crop conditioning	Trimming		hand pruning saw		
Crop conditioning	Trimming		hand pruning shear		
Crop conditioning	shooter removal		manual by scythe		scythe is in Dutch a 'zeis'. It is also known as 'snit' a zeis with an extended stem.
Plant protection	-	*The practice of managing pests, plant diseases, weeds and other pest organisms that damage agricultural crops and forestry*	-		
	Chemical weed control	*weed control by the use of plant protection products.*			We need a separate list of drift reduction techniques.
Plant protection	Chemical weed control		Band spraying	*spraying which is restricted to a band over the plant row.*	
Plant protection	Chemical weed control		Band spraying with shields		
Plant protection	Chemical weed control		Bed spraying with shields	*spraying which is restricted to the planted beds.*	
Plant protection	Chemical weed control		Bed spraying with shields and air suspension		
Plant protection	Chemical weed control		Broadcast spraying	*spraying with a field crop sprayer and a boom with nozzles which covers the*	

CL291	CL293	Definition of CulturalPractise	CL292	Definition of Operation Technique	remarks
Cultural Practice	**Cultural Practice**	**Level 2**	**Operation Technique**		
Level 1	**Level 2**	This is not complete		This is not complete	
				whole sprayed surface	
Plant protection	Chemical weed control		Broadcast spraying		
Plant protection	Chemical weed control		Broadcast spraying with aircraft		
Plant protection	Chemical weed control		Broadcast spraying with helicopter		
Plant protection	Chemical weed control		Broadcast spraying with knapsack sprayer		
Plant protection	Chemical weed control		Broadcast spraying with portable boom with air suspension		
Plant protection	Chemical weed control		Broadcast spraying with shielded spray boom		
Plant protection	Chemical weed control		Broadcast spraying with shielded spray boom and air suspension		
Plant protection	Chemical weed control		Chemical hoe		
Plant protection	Chemical weed control		Controlled droplet application	*spray heads consisting of rotating discs for low volume, controlled droplet, application*	
Plant protection	Chemical weed control		Drip irrigation		
Plant protection	Chemical weed control		Inter row spraying	*spraying which is restricted to a band between the plant rows.*	
Plant protection	Chemical weed control		Inter row spraying with shields	*spraying which is restricted to a band between the plant rows and where the nozzles and space below it is shielded to prevent drift to the plant row*	
Plant protection	Chemical weed control		Spot spraying by hand		

CL291	CL293	Definition of CulturalPractise	CL292	Definition of Operation Technique	remarks
Cultural Practice	**Cultural Practice**	**Level 2**	**Operation Technique**		
Level 1	**Level 2**	This is not complete		This is not complete	
Plant protection	Chemical weed control		Spot spraying by vision control		
Plant protection	Chemical weed control		Sprinkler irrigation		
Plant protection	Chemical weed control		Weed wiper		
Plant protection	Chemical weed control		Wheel track spraying	*spraying which is restricted to the surface of the wheel tracks between plant beds.*	
	Disease control				
Plant protection	Disease control		Axial fan sprayer		
Plant protection	Disease control		Broadcast spraying	*spraying with a field crop sprayer and a boom with nozzles which covers the whole sprayed surface*	
Plant protection	Disease control		Cross flow sprayer		
Plant protection	Disease control		Manual duster (and other dusters)		Validation is needed if it is in the field ?
	Mechanical weed control	*Weed control by hand, tools or machines, without the use of plant protection products.*			
Plant protection	Mechanical weed control		Draw hoe (Collinear hoe)	*manual hoe with sharpened side towards the person using the hoe*	A validation is needed
Plant protection	Mechanical weed control		Finger weeder		
Plant protection	Mechanical weed control		Hand weeding	*Weed removal by pulling/removing by hand.*	
Plant protection	Mechanical weed control		Inter row brush weeder		
Plant protection	Mechanical weed control		Inter row thermic weed control		
Plant protection	Mechanical weed control		Inter row rolling cultivator		
Plant protection	Mechanical weed control		Intra row burning		
Plant protection	Mechanical weed control		Intra row pneumatic weed control		
Plant protection	Mechanical weed control		Intra row thermic weed control		

CL291	CL293	Definition of CulturalPractise	CL292	Definition of Operation Technique	remarks
Cultural Practice	**Cultural Practice**	**Level 2**	**Operation Technique**		
Level 1	**Level 2**	This is not complete		This is not complete	
Plant protection	Mechanical weed control		Intra row weeder with plant recognision		
Plant protection	Mechanical weed control		Manual hoe	*hand held hoe which must be pushed*	
Plant protection	Mechanical weed control		Mechanical hoe	*one or more hoes on a toolbar*	
Plant protection	Mechanical weed control		Mechanical hoe with RTK GNNS control		
Plant protection	Mechanical weed control		Mechanical hoe with vision controlled		
Plant protection	Mechanical weed control		Mowing		
Plant protection	Mechanical weed control		Rotative weed harrow		
Plant protection	Mechanical weed control		Tined harrow		
Plant protection	Mechanical weed control		Torsion weeder		
Plant protection	Mechanical weed control		Whole field burning		
Plant protection	Mechanical weed control		Whole field thermic weed control		
Plant protection	Mechanical weed control		Zig zag harrow		
	Nematode control				
Plant protection	Nematode control		banding	*applying granulate in the plant row*	
Plant protection	Nematode control		broadcasting	*broadcasting of granulate*	
Plant protection	Nematode control		Injection with plough		
Plant protection	Nematode control		Injection with shank		
Plant protection	Nematode control		side dressing	*side dressing of granulate*	
Plant protection	Nematode control		Soil fumigation		
Plant protection	Nematode control		Water lodgingInundatio		A validation is needed
	Pest control				
Plant protection	Pest control		Axial fan sprayer		
Plant protection	Pest control		Broadcast spraying	*spraying with a field crop sprayer and a boom with nozzles which covers the whole sprayed surface*	
Plant protection	Pest control		Cross flow sprayer		
Plant protection	Pest control		Fogging		

CL291	CL293	Definition of CulturalPractise	CL292	Definition of Operation Technique	remarks
Cultural Practice	**Cultural Practice**	**Level 2**	**Operation Technique**		
Level 1	**Level 2**	This is not complete		This is not complete	
Plant protection	Pest control		Fumigate		
Plant protection	Pest control		Granulate banding		
Plant protection	Pest control		Granulate broadcasting		
Plant protection	Pest control		Low volume fogging		
Plant protection	Pest control		Manual anointing		
Plant protection	Pest control		Manual removal		
Plant protection	Pest control		Mole clamp		
Plant protection	Pest control		Spray can		
	Not specified which level 2				
Plant protection	Not specified which level 2		Portable thermal fogger		In the field or glasshouse
Plant protection	Not specified which level 2		Manual (Other techniques)		A validation is needed
Plant protection	Not specified which level 2				A validation is needed whether this is in the field.
Plant protection	Sorting		manual sorting		

Appendix 3 Knowledge Graph Plant Protection Products

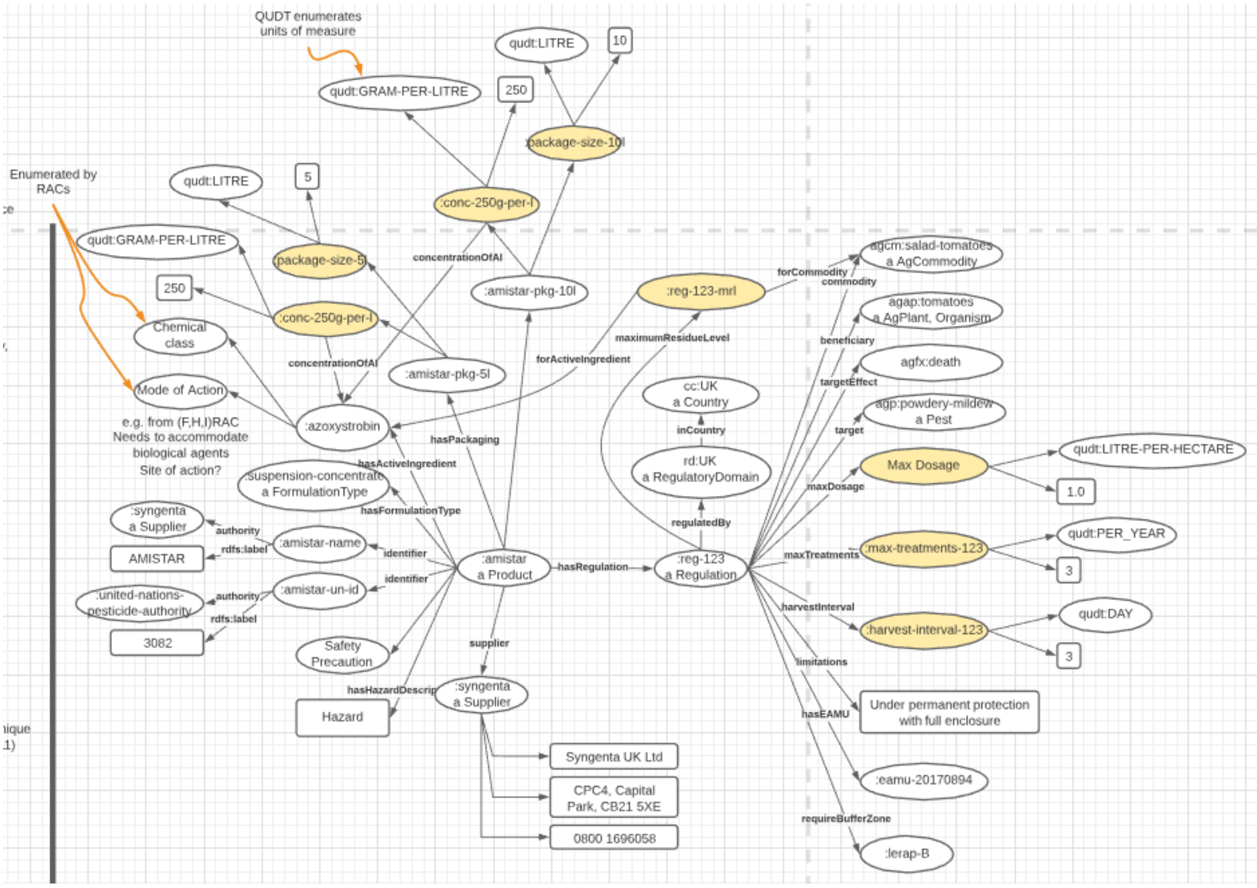

Figure 13 *PPP Knowledge graph – example of a specific product, represented using the RDF model.*

Appendix 4 Collaboration methods and requirements

This chapter is provided as an attachment to this report to ensure reproducibility of this study to some extent. It provides insights on a how a possible shared development and maintenance environment could be realised. It is merely formative and was written during the execution of the project.

Required software
Maintenance of the data and ontology files are done with TopBraid Composer. It is possible to use the free edition of TopBraid (https://www.topquadrant.com/topbraid-composer-install/, Free Edition) although this is a slightly old version. Also, local Git client software is needed. At WUR, Git 2.18.0 is available which is working fine. The latest version can be downloaded here: https://git-scm.com/downloads

Install a merge tool to facilitate the resolution of merge conflicts when pulling or pushing changes. You could use Meld (https://meldmerge.org/). To integrate in your Git Client (Git Bash in this case) enter the following statements:
- git config --global merge.tool meld.
- git config --global mergetool.meld.path 'C:\Program Files (x86)\Meld\Meld.exe'.

After this configuration you should be able to start the merge tool from within Git Bash using the following statement:
- git mergetool.

Connecting to the ag-gateway repository
We didn't succeed in setting up the link to the Git repository directly from within TopBraid Composer, probably because of the two factor authentication in place for the ag-gateway GitHub repository.

To get a local copy of the repository you could use Git Bash (this software is available when you installed the Git client as mentioned above). Clone the repository (https://github.com/agrimetrics/ag-gateway - Connect to preview). A cheat sheet can be found here:

https://www.atlassian.com/git/tutorials/atlassian-git-cheatsheet

After a successful clone you can import the cloned project in TopBraid Composer. When this all is successful you should see the small orange icons on the cloned project files in Composer which indicate the link to the GitHub repository

Procedure to collaborate on the Git repository
- Before you start your changes, pull the latest version of the repository:
  In Git Bash git pull --rebase.
  If this causes merge conflicts you can resolve them with e.g. 'Meld' (see above in required software).
  In Git Bash git mergetool.
- Maintain the repository using TopBraid and save the changes.
- After saving you can check your local changes using Git GUI. Alternatively you can use git diff in Git Bash.
- When you are satisfied with your changes you should do a git pull --rebase again and when this is successful you can do a git add (Stage in Git GUI) and git commit -m '<message>' (Commit Message and Commit in Git GUI). Always add a comprehensive comment on what you did and why.
- Finally push your change to the master: git push (Push in Git GUI).
- You can check the status with command git status to check if your local git folders are up to date with the remote repository, which should be the case after a successful commit.

Miscellaneous:

- Your local.project file should not be admitted to the GitHub repository. It is added to the.gitignore file in the repository, so this should be OK if you cloned or pulled the latest version.
- If you have other local files in your local git folder which should not be added to the git repository you can add them with the command: echo filename >>.gitignore.
- If you by accident added a file to the remote repository you can remove it using this command: git rm <filename>.
- Use git pull --rebase instead of git pull. The --rebase option will cause a fetch of the remote's copy of current branch and rebases the local changes onto it to integrate them.
- To have a look at the local directory, using Git Bash: lsor ls -a to see hidden files as well (such as.project and.gitignore).
- To have a look at the commit history (single line per commit), using Git Bash:
  git log --oneline, or use the gitk interactive tool (gitk & in bash).

Wageningen Economic Research
P.O. Box 29703
2502 LS The Hague
The Netherlands
T +31 (0)70 335 83 30
E communications.ssg@wur.nl
wur.eu/economic-research

WHITE PAPER 2022-051

To explore the potential of nature to improve the quality of life

Wageningen Economic Research
P.O. Box 29703
2502 LS Den Haag
The Netherlands
T +31 (0) 70 335 83 30
E communications.ssg@wur.nl
wur.eu/economic-research

White Paper 2022-051

The mission of Wageningen University & Research is “To explore the potential of nature to improve the quality of life”. Under the banner Wageningen University & Research, Wageningen University and the specialised research institutes of the Wageningen Research Foundation have joined forces in contributing to finding solutions to important questions in the domain of healthy food and living environment. With its roughly 30 branches, 6,800 employees (6,000 fte) and 12,900 students, Wageningen University & Research is one of the leading organisations in its domain. The unique Wageningen approach lies in its integrated approach to issues and the collaboration between different disciplines.