Ref. Ares(2023)7053961 - 17/10/2023

EU-FarmBook

Deliverable 1.13

Data standards and knowledge graph v1

Document/Report | Public

Funded by
the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Summary

Call	HORIZON-CL6-2021-GOVERNANCE-01
Topic	HORIZON-CL6-2021-GOVERNANCE-01-24
Project	EU-FarmBook: supporting knowledge exchange between all AKIS actors in the European Union
Acronym	EU-FarmBook
Project No.	101060382
Management	Universiteit Gent
Duration	84 Months
Start date	01/08/2022
End date	31/07/2029

Deliverable	D1.13. Data Standards and Knowledge Graph v1
Type	**R (Document/Report)**
Dissemination level	**PU (Public)**
Due Date	31/07/2023
Submission Date	17/10/2023
Work Package No.	WP1
Lead Beneficiary	University of Maastricht (MU)
Authors	Hercules Panoutsopoulos (Agricultural University of Athens) | Panagiotis Stamatelopoulos (Agricultural University of Athens) | Xu Wang (University of Maastricht) | Louis Powell (University of Maastricht) | Christopher Brewster (University of Maastricht) | | Pranav Bapat (University of Maastricht)
Contributors	
Version	Version 1.0

History of Changes

Version 1	01/09/2023	Hercules Panoutsopoulos (AUA)	Initial draft
Version 1.1	01/09/2023	Xu Wang (MU)	Knowledge graph sections
Version 2	05/10/2023	Christopher Brewster (MU)	FAIR principles and semantic standards

Index

Figure index

Table index

Abbreviations

AGRO	Agronomy Ontology
AGROVOC	AGROVOC is a multilingual controlled vocabulary covering all areas of interest of the Food and Agriculture Organization of the United Nations, including food, nutrition, agriculture, fisheries, forestry and the environment.
AKIS	Agricultural Knowledge and Innovation Systems
API	Application Programming Interface
DOI	Digital Object Identifier
ENVO	Environment Ontology
EURAKNOS, EUREKA	Predecessor projects for EU-FarmBook
FAIR	Findability, Accessibility, Interoperability, and Reusability
FoodOn	A broadly scoped ontology representing entities which bear a "food role"
GACS	Global Agricultural Concept Space
JSON	JavaScript Object Notation
JSON-LD	JSON Linked Data
M[0-9]	Month (month number starting from August 2022) M1 = August 2022, M6 = January 2023, etc.
RDF	Resource Description Framework
SPARQL	SPARQL Protocol and RDF Query Language

Executive Summary

The present report ("Data standards and knowledge graph v1") is part of a series of technical reports expected to be released throughout the project to document advancements in the work on data standards and the EU-FarmBook Knowledge Graph. Updates will become available by M36 (end of July 2025) and M72 (end of July 2028). Here we present the motivation and analysis that has led to the design of the current ontologies, data standards and knowledge graph of the project.

1. Introduction

A number of requirements regarding the design and implementation of the EU-Farmbook[1], and certain design choices have meant that we will use explicit data standards and a knowledge graph in the EU-FarmBook platform. These include:

- The adoption of the FAIR data principles [1] which have been widely recognised in the scientific community, and adopted by the EC for the majority of EC funded projects. The EU-FarmBook builds here on the report written for the EUREKA project (Deliverable 3.2: “Data Repository Design and Data Pipeline Design”).
- The need for interoperability both within the project modules (as we follow a micro-services architecture) and with other projects via our API. This means that a well documented API that corresponds closely to the data standards (ontologies) that we have defined is fundamental.
- The vision of the EU-FarmBook is that each Knowledge Object should have attached to it a set of metadata (e.g. title, author/creator, topic, license etc.). This is both to conform to the FAIR data principles, and to allow for the creation of a knowledge graph capturing the relations between knowledge objects, authors/creators, projects, topics/domains.
- The planned chatbot interface of the EU-FarmBook also will benefit from the knowledge graph, enabling sophisticated querying and recommendation in the process of answering end user requests.

Good practice in the design and development of modern computer systems lays an emphasis on the use of existing data standards as much as possible to ensure interoperability, and to minimise the reinvention of the wheel. Furthermore the last ten years has seen an explosion in the use and application of knowledge graphs and these depend on an underlying data model usually expressed as an ontology.

In this deliverable, we present the design choices made concerning data standards (i.e. ontologies) and knowledge graphs, we provide an overall background and context, and lay out a roadmap for the further development and maintence of these resources throughout the life time of the EU-FarmBook project.

[1] See the features defined in the EPIC “Storage and Storage Services” defined in Deliverable 1.1 (“Platform requirements and design v1”).

2. Relation to work in other tasks and Work Packages

This deliverable stands in close association with Deliverable D1.1 ("Platform requirements and design v1") which lays out the initial rationale and design choices. Deliverable D1.6 ("Data Ingestion Pipeline and upload Interface v1") depends in part on this deliverable for the underlying attribute value pairs that populate the upload form and via which the metadata for the knowledge graph is captured. The standards and knowledge graph described here also form a core input for both the D1.6 ("Data Ingestion Pipeline and Upload Interface v1"), as well as being a crucial input for D2.4 ("Search and Recommender System"), as the knowledge graph will play a significant role in the infrastructure of the chatbot.

3. Role of data standards and knowledge graphs

3.1 Levels of Interoperability and Standards

Standards enable information system interoperability, meaning the capacity of information systems to "talk" to each other. Interoperability is about data exchange and sharing[2]. According to the Joined-Up Data Standards (JUDS) projects, interoperability is "*the ability to join up data from different sources in a standardised and contextualised way*" [2]. Core to information system interoperability is the interoperability of data and information exchanged among information systems [3]. However, interoperability does not depend on what data and information are exchanged but rather on how data and information are encoded in the interacting systems. There are a number of interoperability frameworks proposed in the literature. Interoperability can be approached and operationalised at different abstraction layers/levels.

The Data Commons Framework [4] distinguishes among: (i) the technical layer (having to do with the use of standardised interfaces to publish data on the web); (ii) the data and formats layers (concerned with the structuring of data and metadata using broadly adopted models and schemas); (iii) the organisational and institutional layers (covering aspects related to data sharing agreements and licensing, which specify responsibilities and obligations for effective cross-organisational data sharing and exchange); and the human layer (i.e., the establishment of a common understanding between producers and consumers of data regarding the meaning of terms used to describe the content of data and its potential use).

[2] https://edps.europa.eu/data-protection/our-work/subjects/interoperability_en

The European Commission's National Interoperability Framework Observatory (NIFO) proposes a four-layer framework[3] of: (i) technical interoperability (covering aspects of interface specifications, data integration and exchange services, secure communication protocols, etc.); (ii) semantic interoperability (ensuring the establishment of a common understanding of the meaning of the data and information exchanged both by human and machine agents); (iii) organisational interoperability (i.e., the alignment of adopted business processes and the data/information exchanged); and (iv) legal interoperability enabling organisations to effectively work together despite the legal and policy contexts within which they operate.

ETSI (European Telecommunications Standards Institute) identifies four interoperability levels, namely: technical, syntactic, semantic, and organisational interoperability [5]. Technical interoperability ensures between-system communication at the machine level through communication protocols and the appropriate infrastructure allowing protocols to operate. At the syntactic level, interoperability concerns the syntax used to format and encode data. From the semantics perspective, interoperability is about integrating meaning in the data and information exchanged. Semantic interoperability ensures the establishment of a common understanding of what is exchanged by humans and machines. At the organisation level, interoperability is approached at a higher level of abstraction that has to do with the ability of different organisations/entities to effectively cooperate regardless of the technologies and information systems they may use. Given this focus, organisational interoperability is a broad term encompassing interoperability at all the other levels (namely, technical, syntactic, and semantic interoperability).

There are a number of challenges related to data interoperability at the semantic level. These include the generation of data from a multitude of digital or human sources, the use of different lexical forms to denote the same concept, and barriers in data sharing due to sensitivity and privacy reasons. Interoperability helps to establish data sharing mechanisms enabling a common understanding of the data shared at the syntactic and semantic levels, as well as safeguarding the sensitivity of the information exchanged. Interoperability is considered at the foundational (having to do with machine-to-machine communication), structural (being an intermediate level of interoperability related to the structuring and syntax of the data exchanged, therefore ensuring syntactic interoperability), and semantic (namely, the common understanding of the data being exchanged by both human and machine agents) levels [6].

Despite the variations in the terminology used in the existing frameworks, the levels at which interoperability is considered are those of machine-to-machine communication, data and information structuring, data and information interpretability, as well as cross-organisational communication/cooperation based on the data/information exchanged.

Standards are specifications adopted by technology manufacturers and accepted by users, which are important for the compatibility of hardware and software[4]. They allow for the development of technology products capable of working together in a seamless way, thus enabling interoperability at different levels. Standardization helps to promote best practices in safety and quality. All these aspects are highlighted in the definition of

[3] https://joinup.ec.europa.eu/collection/nifo-national-interoperability-framework-observatory/3-interoperability-layers
[4] https://www.gartner.com/en/information-technology/glossary/standards

NIST (National Institute of Standards and Technology of the U.S.)[5], according to which standards provide a common language enabling interoperability at all levels, as well as promoting safety and product durability. Standards are established by standards development organisations and their technical committees (e.g., ISO[6], W3C[7], ETSI), a process involving the drafting of a technical description of the standard, voting by the technical committee experts, and consensus establishment[8]. Apart from this top down approach, standards may be also established in a bottom-up way driven by wide community adoption. The latter often result in *de facto* standards.

At the technical level, there are both hardware and software standards. USB (Universal Serial Bus), HDMI (High-Definition Multimedia Interface), and Bluetooth are well-known computer hardware standards. HTTP (HyperText Transfer Protocol), XML (eXtended Markup Language), and RSS (Really Simple Syndication) are examples of computer software standards. Syntactic standards allow for shaping the structure and syntax of data (i.e., the “packaging” of the data). Semantic standards enable machine-interpretable definitions of the meaning of data [7]. The proliferation of web technologies and the abundance of the data generated and needed to be exchanged among machine and human agents have driven the need for web technology standards development. At the level of semantics, W3C is a leading organisation in standards creation. RDF[9], OWL[10], SKOS[11], RDFS[12], JSON-LD[13], SPARQL[14], and SHACL[15] are widely known semantic web technology standards established by W3C.

At the syntactic level, there are several formats enabling data sharing among web tools and technologies. However, it is hard to say that any of these formats has reached the level of becoming an industry standard. Given the continuously growing number of web applications needing to “talk” to each other, using a common syntax for the sharing and exchange of data is important. Over the years, JSON (JavaScript Object Notation) has become a popular data serialisation format and it is already condisered a standard. XML and YAML are two other examples of syntactic data standards.

Several syntaxes exist for the serialisation of RDF data (the semantic standard for modelling and exchanging data in the World Wide Web; see Section 4.2 below). Such syntaxes include RDF/XML[16], N3[17], Turtle[18], TriG[19], N-Triples[20] and N-Quads[21].

[5] https://www.nist.gov/services-resources/standards-and-measurements
[6] https://www.iso.org/home.html
[7] https://www.w3.org/
[8] https://www.iso.org/developing-standards.html
[9] https://www.w3.org/TR/rdf12-concepts/
[10] https://www.w3.org/TR/owl-features/
[11] https://www.w3.org/TR/skos-reference/
[12] https://www.w3.org/TR/rdf12-schema/
[13] https://www.w3.org/TR/json-ld11/
[14] https://www.w3.org/TR/rdf-sparql-query/
[15] https://www.w3.org/TR/shacl/
[16] https://www.w3.org/TR/rdf-syntax-grammar/
[17] https://www.w3.org/DesignIssues/Notation3
[18] https://www.w3.org/TR/turtle/
[19] https://www.w3.org/TR/trig/
[20] https://www.w3.org/TR/n-triples/
[21] https://www.w3.org/TR/n-quads/

3.2 FAIR Data Principles

In the twenty-five plus years since the creation of the World Wide Web, we have moved from a Web of documents to a vision for a Web of data [8], especially open data and linked data [9]. The FAIR data principles represent the latest incarnation of a growing understanding of the implications of sharing data with other parties. While the "open data" movement represented a slightly utopian vision of the many possibilities of data sharing, the FAIR principles respect the reality of practical data sharing across different parties. For instance, the Accessible ("A") dimension of FAIR data can be seen as a parameter that varies between very open to very closed. Mons et al. [10] underline that "The "'A' in FAIR stands for 'Accessible under well-defined conditions'" and mention personal privacy, national security and competitiveness as reasons for data to be shielded. The original home territory of the FAIR principles was life science research, where there is both a strong tradition of open databases and shared metadata principles [1], as well as a strong awareness of the sensitivity of certain types of data (like health or commercial data). In the food, agriculture and forestry domains there has been a significantly slower uptake of this paradigm or philosophy. While data and data science are hailed as new economic opportunities for agricultural producers, the application of the FAIR data principles on the farm are somewhat more controvertial [11]–[13]. However, for the purposes of the EU-FarmBook our concerns are with project outputs, typically reports, academic papers and practice abstracts rather than raw data. In this arena, the advantages of the FAIR data principles as a means to ensure that knowledge objects are more easily accessible and usable, are incontrovertible.

The principles emphasise the importance of metadata about data or other research outputs, so as to allow them to Findable, Accessible, Interoperable, and Reusable. There is no longer an expectation that all data be open or directly linked to other data sets because access is recognized as an important parameter. But from a scientific point of view, the key is to ensure that the descriptive metadata is appropriate, interoperable and permanently available. There is a growing awareness that these principles are of great utility including agriculture [14], especially in research data, but we would argue the principles are equally of relevance in the forestry sector and all bio-based sectors.

3.3 Semantic Standards in the Domain of Agriculture

There now exists a plethora of data standards and ontologies for the agrifood domain. This can be seen in its broad scope by looking at the catalogue available at the Agrisemantics Map of Data Standards[22] , where over 400 vocabularies and standards are listed, or more narrowly at the Agroportal website[23] [15], which lists 159 ontologies. Most researchers in the application of ICT to the agricultural domain acknowledge the need for widely adopted data standards but as these catalogues indicate adoption is hindered less by the lack of appropriate standards than a strong tendency to reinvent the wheel for a variety of reasons, most of which can be summarised as a combination of either ignorance or "not invented here" syndrome.

[22] https://vest.agrisemantics.org/
[23] http://agroportal.lirmm.fr/

We can identify a number of different communities which have created standards (often quite independently of each other). In Table 1, we show different communities with *some example* standards/ontologies, as well as distinguishing between fundamental types of standards [16]. These different communities are neither exhaustive nor discrete with plenty of overlap between them, and if the reader looks at the VEST Agrisemantics registry, they will find many more "domains" listed than here. Equally the types of standards are not exhaustive but reflect the dominant types[24].

Table 1: Examples of dominant standards/ontologies for the agrifood sector

Sub-sector or community	Message Oriented	Vocabulary/ Thesaurus	Ontology
Agricultural Research publications		AGROVOC, GACS, CABI Thesaurus	
Agricultural research - datasets			AgrO,Plant Ontology, Environment Ontology
On Farm	AEF-ISOBUS, ICAR, AgGateway-ADAPT		
Supply Chain/Retail	GS1 family (EPCIS, GTEIN, etc.), UN/CEFACT		
Food Safety/Integrity	FOODEx2		
Food Products/ Nutrition (**Research**)			FoodOn

Here we focus exclusively on voacabularies an ontologies. The agricultural and food-related research community has been long depended on standardised vocabularies to annotate publications. This has a long tradition in library science and information retrieval, and in agrifood has found its expression in a number of such quasi-formal vocabularies or thesauri, including AGROVOC, CABI Thesaurus[25], NALT[26] and Langual[27].

AGROVOC: The UN Food and Agriculture Organisation built the multilingual AGROVOC[28] over a number of years dating back to the 1980s [19], [20]. Originally designed as a thesaurus, in the early 2000s it was formalised first in OWL then in SKOS. It is the largest concept scheme in the agrifood domain with over 39,000 concepts, and one of its key attributes is that all concepts are provided with labels in up to 40 languages. The major purpose of the AGROVOC concept scheme has been and remains to annotate the AGRIS document repository[29]. The major advantage of AGROVOC over other concept schemes and standardised vocabularies is the multiple languages, but the hierarchy and structure has tended not to find favour among ontology engineers building systems for decision support or data annotation. Nonetheless more formal ontologies, like FOODON (cf. below) and many others map their concepts to AGROVOC terms.

[24] For a more detailed taxonomy in general see [17] and for agrifood [18].
[25] https://www.cabi.org/publishing-products/cab-thesaurus/
[26] https://data.nal.usda.gov/dataset/nal-agricultural-thesaurus-and-glossary
[27] https://www.langual.org/
[28] http://www.fao.org/agrovoc/
[29] http://aims.fao.org/agris

In EU-FarmBook, we have chosen to largely base our topic annotations on AGROVOC taking adavtage of its multilinguality.

GACS: The Global Agricultural Concept Space [21] has been built selecting the most frequently used terms in AGROVOC, CABI Thesaurus[30] and NAL[31]. GACS provides a mapping between these three widely used concept schemes, and a set of permanent globally unique URIs. Furthermore, GACS inherits the multilingual labels from the other thesauri. It is built using SKOS, choosing intentionally a lightweight semantics approach. Its purpose is to be used to annotate documents and (possibly) data sets to enable “broad-brush discovery”, as opposed to formal inference. Unfortunately, GACS has not found wide uptake and in early 2020 it was announced that it would not be further maintained[32]. This is a fate frequently encountered by numerous vocabularies, concept schemes and ontologies.

Since the development of ontologies as tools for knowledge representation, in the late 90s, early 00s, there have been a succession of attempts to create formal ontologies expressed in OWL. Most significantly, AGROVOC was converted to OWL and then this was abandoned and now is available in SKOS. More recently a series of more specialised ontology or ontology frameworks have been developed partly under the influence of parallel activities in the Life Sciences/Bioinformatics. Here, we briefly list some of these which are the most influential in agriculture and forestry research.

ENVO: The Environment Ontology [22], [23] originates in the bioinformatics community’s desire to describe consistently the environmental origins of tissues, pathogens, and metagenomic samples. Built on rigorous foundations, it followed the Open Biomedical and Biological Ontologies (OBO) Foundry principles [24], and is intended to facilitate the annotation and retrieval of a broad range of environmental data. This covers biomes, habitats, environmental features, and environmental materials. The ontology has grown and has been extended to cover environmental processes, anthropogenic environments, and entities relevant to environmental health initiatives and the global Sustainable Development Agenda for 2030. The ENVO has approximately 6500 concepts (many inherited from elsewhere such as the CHEBI ontology) and has a large proportion of terms relevant to the agrifood sector. The ontology has been quite successful in being widely adopted by different projects and with its terms reused in other ontologies. It is maintained as a community effort largely from the bioinformatics community.

FOODON: The Food Ontology [25][33] is a relative newcomer to the landscape of agrifood related ontologies, although it already contains over 30,000 classes. It is partly an outgrowth of work on ENVO, and attempts to provide a comprehensive systematic and rigorous ontology focussing initially on *food products* although the overall ambition is to provide an ontology for all aspects of the food system from farm to fork. Its origins lie in part with the desire to provide systematic descriptions of food borne pathogens. FOODON is built upon the Langual system[34], a food classification system originally built

[30] https://www.cabi.org/cabthesaurus
[31] https://agclass.nal.usda.gov/
[32] http://www.fao.org/agrovoc/news/gacs-core-not-currently-being-updated-and-will-not-be-actively-maintained-future
[33] https://foodon.org/
[34] https://www.langual.org/default.asp

for the US FDA in the 1970s [26] focussing on food products and their ingredients. It follows the OBO Foundry principles like ENVO, and provides a systematic means to describe raw food sources ingredients, food processing, packaging and preservation. FOODON is mapped (via LanguaL) to EFSA Foodex2 vocabulary. The major difference between FOODON and AGROVOC or GACS lies primarily in the ontological rigour, but also in that (for example) GACS describes a "domain of discourse" while FOODON tries to formally represent the relationship between a raw ingredient, derived food products and production processes. Furthermore, FOODON is largely monolingual (only in English) with little attempt to provide labels in other languages. There is considerable uptake of FOODON in work related to nutrition, although the intention is that this ontology would be especially useful in track and trace as well. Overall, FOODON is actively maintained and positions itself as part of a family of OBO based ontologies related to food.

AGRO: The Agronomy Ontology, like ENVO and FOODON, also forms part of the OBO ecosystem and follows those principles concerning formal ontology design. Its focus is the agronomy domain, providing terms to describe agricultural practices, cropping systems, field management, soil, weather and crop phenotypes, building on a number of existing ontologies including ENVO and PATO[35]. The ontology was built and is maintained by the CGIAR network of research centres, and forms a core part of the Agronomy Field Information Management System, a CGIAR system to capture fieldbooks describing crop management practices[36], as well the GARDIAN data management system[37]. It is a relatively small ontology with just over 2000 classes and is extensively used in a number of crop specific ontologies built as part of the "Crop Ontology Curation Tool" initiative[38].

There are many other agriculture and forestry related ontologies but here we have focussed on the most influential.

3.4 Knowledge Graphs

A semantic conceptualisation of a domain (such as AGROVOC) can be instantiated and thus provide a knowledge graph. Knowledge graphs have become extremely widely used ever since Google introduced the term to a wider audience in the early 2010s. There ia growing academic literature on how to construct knowledge graphs, how to use them for a variety of applications, and how they can provide essential knowledge representation infrastructure for many applications [27], [28]. The EU-FarmBook knowledge graph provides a key part of the metadata infrastructure. With each Knowledge Object that is uploaded a variety of metadata is captured both manually and automatically. This is mapped to the core ontologies that the project has developed and maintains. From this, a semantically enabled knowledge graph is constructed and subsequently available for a variety of purposes. Further details are provided in Section 5.

[35] https://www.ebi.ac.uk/ols/ontologies/pato
[36] https://agrofims.org/about
[37] https://bigdata.cgiar.org/resources/gardian/
[38] http://www.cropontology.org/

4. Data Standards

4.1 Syntactic Standards for the Data Ingestion Pipeline

At the syntax level, the standard considered for building the EU-FarmBook's data ingestion pipeline (details are available in the report of Deliverable 1.6 "Data Ingestion Pipeline and upload Interface v1") is JSON (JavaScript Object Notation). JSON is a lightweight data exchange syntax that can be easily generated and parsed by machines and written and read by humans[39]. Many libraries are available in different programming languages, which enable the creation and parsing of JSON objects. In JSON, data is modelled in the form of key-value pairs. It is the format used by REST APIs[40] for the exchange of data between software applications or application components. It is the format used by web APIs for pushing data into and pulling data from web applications. It is a text-based, language-independent data interchange format[41] which packages data as collections of key/value pairs. Although not required by REST[42] (API format used for web API development), it makes a good practice to use JSON for the development of REST APIs given its text-based and compact data serialisation [30]. The development of the EU-FarmBook platform is based on the micro-services approach, providing more flexibility in maintenance and updates. The use of REST APIs is at the core of the micro-service approach.

The syntax for serialising the EU-FarmBook's Knowledge Graph is JSON-LD[43], where LD is an abbreviation for Linked Data. Linked Data is a term tightly related to the Semantic Web concept. It refers to a specific way of data publishing and consumption using unique and machine-readable identifiers, allowing for the embodiment of meaning in data and creating relations between different data elements. Linked Data and Semantic Web technologies are at the core of knowledge graphs. JSON-LD is a syntax used to serialise Linked Data in JSON. It is designed to integrate into deployed systems already using JSON easily and allows to store Linked Data in JSON-based data stores. These features make JSON-LD a flexible format for data exchange within the platform system, considering MongoDB as our data store (details are available in the report of Deliverable 1.10 "System infrastructure and API v1") and REST APIs to enable communication between system components at various levels.

[39] https://www.json.org/json-en.html
[40] For details about what RESTful web services are and how they are implemented see [29].
[41] https://datatracker.ietf.org/doc/html/rfc4627
[42] https://blog.postman.com/rest-api-examples/
[43] https://www.w3.org/TR/json-ld/

4.2 Semantic Standards for the EU-FarmBook Ontology and Knowledge Graph

4.2.1 Previous Work on the EU-FarmBook Ontology

The EU-FarmBook ontology presented next (see Sub-section 4.2.3) is the result of work that has been done during the first year of the project, building upon legacy work from the EURAKNOS and EUREKA projects, the predecessors of the EU-FarmBook project. Here, we provide some details of previous work to link to the work currently in progress. The ontology creation has been driven by the need to define metadata for the description of the Knowledge Objects hosted in the platform. The metadata aims to capture as much information as possible regardless of how and why the Knowledge Objects were created. Ultimately, a structured description of the variety of Knowledge Objects was targeted. Responding to questions about the kinds/types of Knowledge Objects, the projects from which Knowledge Objects become available, and the persons involved in their creation was at the core of the ontology development process.

In EURAKNOS, the first version of the ontology was built. Design conventions proposed in Ontology UML Profiles (namely, extensions of the set of design elements used in UML class diagrams[44]) were used to model the ontology classes (i.e., abstractions that denote collections of individuals, concepts, or objects with similar attributes), relations between classes, and cardinalities showing the number of the instances of classes participating in a relation.

The work on the ontology in EURAKNOS was continued and improved in EUREKA. The ontology was associated with the functional and non-functional requirements identified and defined as the result of the user needs analysis in the project. UML was substituted by RDF[45] (Resource Description Framework), a flexible language that describes any resource available on the World Wide Web. Modelling Knowledge Object metadata as RDF triples[46] was the primary motivation for adopting RDF. A significant improvement compared to the work in EURAKNOS was the use of classes and properties defined in well-known and widely adopted ontologies and vocabularies (e.g., schema.org[47] and Dublin Core[48]) to model information related to Knowledge Objects and their relations (instead of custom definitions). Modelling Knowledge Object metadata as triples embodying semantic information (i.e., formal definitions of concepts and relations from external ontologies and/ or vocabularies) was a first step towards FAIRness starting from the metadata level. AGROVOC was the domain vocabulary that maps the topics in Knowledge Objects to agriculture- and forestry-related concepts.

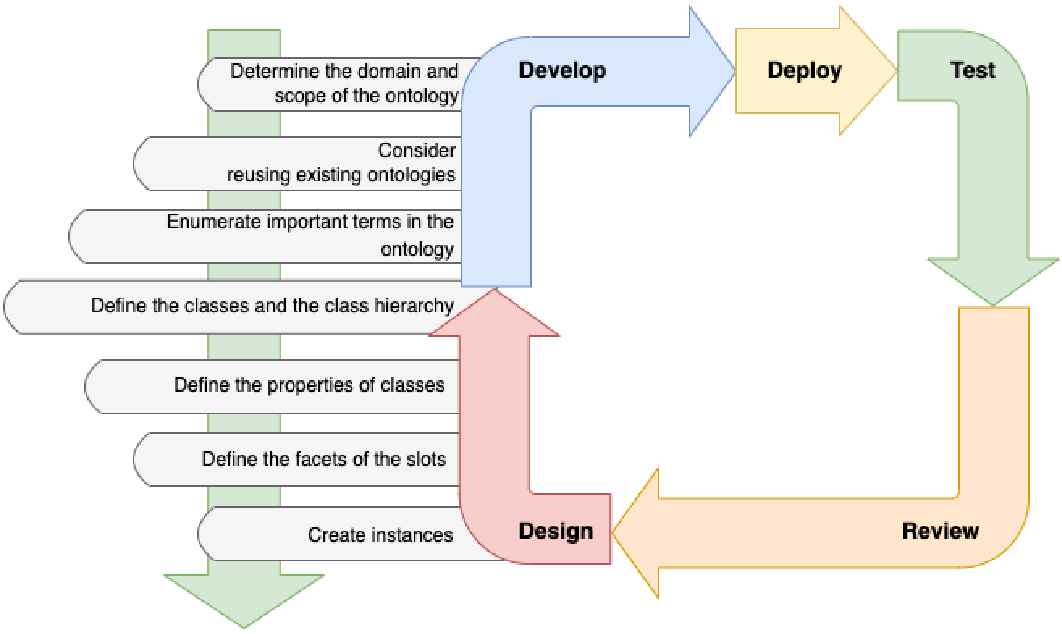
---

[44] UML stands for Unified Modeling Language. Resources related to what UML is and how, and why, is used are available at https://www.uml.org/.
[45] https://www.w3.org/TR/rdf-primer/
[46] https://www.w3.org/TR/PR-rdf-syntax/
[47] https://schema.org/
[48] https://www.dublincore.org/

4.2.2 Methodology for the EU-FarmBook Ontology Development and Update

The development and update of the ontology are grounded on the methodology of Noy and McGuiness [31]. Details about the Noy and McGuiness methodology can be found in Section 8.1 (Annex 1). The Noy and McGuiness methodology (presented in Annex I) for ontology creation is part of the broader process followed in the EU-FarmBook project for developing and updating the EU-FarmBook ontology. Based on the agile approach, the ontology building and update are approached in an iterative manner. This ensures that evolving user needs, as well as the functional and non-functional requirements and domain developments in the projects producing and providing Knowledge Objects, will be adequately captured and considered. The overall methodology for the development and update of the EU-FarmBook ontology is illustrated in Figure 1 below.

Figure 1: EU-FarmBook ontology development and update methodology



During the first year of the project, some physical and online events took place to communicate the EU-FarmBook platform's vision and the work to be done throughout the seven-year duration of the project. Such an event was the Platform Day, held in Brussels on 21 February 2023. In those events, interactions with actors involved in Research and Innovation projects provided insights about future work directions and potential ontology updates. Reconsidering how we capture the geographic context of Knowledge Objects' content, the values of some class properties (e.g., the property related to a Knowledge Object's purpose), or information encoded about the projects producing Knowledge Objects and Knowledge Object creators are indicative examples of future update work. Such updates will be addressed in future ontology versions and documented in the next versions of the report, expected by the end of M36 and M72. The reason for not considering them in the current ontology version is that by the time of receiving stakeholder insights, work on the 1st release of the EU-FarmBook platform was already in progress. To that end, ontology update details are beyond the present report's scope.

4.2.3 Current version of the EU-FarmBook ontology

The ontology constitutes the schema layer of the EU-FarmBook Knowledge Graph of the Knowledge Objects and their metadata. It models abstractions of entities, artefacts, and concepts associated with Knowledge Objects and the projects that Knowledge Objects come from in the form of classes. For instance, the class `frapo:CreativeWork` is used to denote any Knowledge Object resulting from a project. A project is an instance of the `foaf:Project` class. The ontology also models the relations between instances of classes and the properties that the instances of a class have (such as the title and keywords of a Knowledge Object). The schema layer allows for making logical inferences. This means that there is no need to model all the possible relations between the ontology classes explicitly. Based on the visual representation of the EU-FarmBook ontology shown in Figure 2 below, a practice abstract is a Knowledge Object constituting an instance of the `schema:TextDigitalDocument` class. Through logical inference, it is also an instance of the `schema:DigitalDocument` class given that `schema:TextDigitalDocument` is a subclass of the `schema:DigitalDocument` class. Therefore, there is no need to explicitly model a practice abstract as an instance of the `schema:DigitalDocument` class. The EU-FarmBook ontology comprises 15 classes and 18 properties (specifically, five object properties and 13 data properties). The development of the ontology has been based on well-acknowledged and broadly adopted semantic standards such as RDF and OWL[49], as well as the RDFS vocabulary for RDF data[50].

Figure 2: The EU-FarmBook knowledge object ontology

The ontology development has been grounded on the following competency questions:

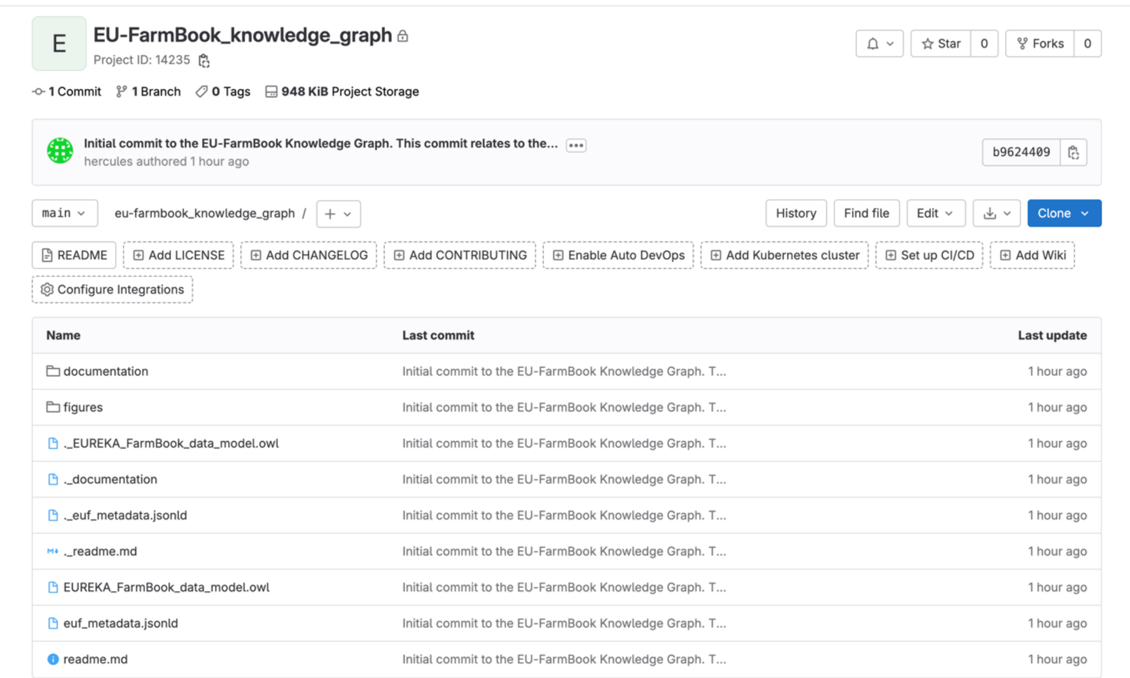
- What types of Knowledge Objects are available from Project X?
- What are the subjects (topics) of the Knowledge Objects available from Project X?

[49] https://www.w3.org/TR/owl-features/
[50] https://www.w3.org/TR/rdf12-schema/

- Which geographic locations relate to the content of the Knowledge Objects of type X?
- What are the Knowledge Objects of type X on subject Y related to the geographic location Z?
- What are the Knowledge Objects of type X available from project Y concerning the geographic location Z?

Components of existing ontologies and vocabularies (classes and properties) have been reused to develop the ontology. Schema.org, Dublin Core, FRAPO (Funding, Research Administration and Projects Ontology), FOAF (Friend-of-a-Friend), and the Ontopic ontology have been considered and used for capturing generic properties and modelling relevant entities and relations. Definitions of the ontology classes and properties can be found in Annex 2 (Section 8.2). Technical details about the EU-FarmBook ontology and the ontology itself are available in the project's GitLab repository (https://ci.tno.nl/gitlab/eu-farmbook/eu-farmbook_knowledge_graph). An overview of the information available in GitLab for the EU-FarmBook Knowledge Graph is provided in Figure 3 below.

Figure 3: Overview of the EU-FarmBook Knowledge Graph project in GitLab



5. Knowledge Graph

5.1 Current version of the EU-FarmBook Knowledge Graph and updates

The current version of the EU-FarmBook Knowledge Graph is sourced from the metadata of uploaded Knowledge Objects. The metadata of Knowledge Objects contains descriptive metadata (textual descriptive information of knowledge objects), structural metadata (such as type of knowledge object), etc. The EU-FarmBook Knowledge Graph is stored in JSON-LD format, a Linked Data format which is easy for humans to read and write (see Section 4.2). Details about how JSON-LD works for storing the EU-FarmBook Knowledge Graph are available in Section 5.2. The EU-FarmBook Knowledge Graph's update relies on updates of uploaded Knowledge Objects. Knowledge Object updates (creation or modification) will be captured in the Knowledge Graph after periodic updates. The current version of the Knowledge Graph serialised in a JSON-LD format is available in the GitLab repository of the project (https://ci.tno.nl/gitlab/eu-farmbook/eu-farmbook_knowledge_graph).

5.2 Knowledge Graph Repository

The Knowledge Graph repository uses JSON-LD for storing RDF data in MongoDB. As a valid JSON format, JSON-LD can be stored in MongoDB and represent the information and knowledge in the metadata of uploaded Knowledge Objects in the RDF data model. Also, as a human-understandable data format, JSON-LD is web-friendly (namely, it could easily be linked to open web data) and data-friendly (namely, it could easily be converted to other types of structural data or be integrated with other RDF data).

MongoDB has several advantages which could support the Knowledge Graph repository, including schema flexibility, optimised performance, and scalability. Schema flexibility means a non-fixed, non-predefined schema for data in MongoDB, which makes RDF data exceptionally adaptive and accommodating to changes. Furthermore, MongoDB is optimised for performance, allowing rapid read and write operations for the EU-FarmBook Knowledge Graph. Lastly, MongoDB is designed with scalability, meaning it can easily scale horizontally, distributing data across multiple servers using its automatic sharding feature.

5.3 Knowledge Graph Construction

The EU-FarmBook Knowledge Graph is built based on the EU-FarmBook ontology. Due to the advantage of JSON-LD and MongoDB, Knowledge Object metadata (serialised in JSON) can be easily converted into JSON-LD by adding context. Context provides the mapping between JSON terms and linked data semantics. With the help of JSON-LD context, the EU-FarmBook Knowledge Graph can be linked to other open data, following the FAIR data principles.

5.4 Knowledge Graph Validation

Knowledge Graph validation aims to check the correctness and consistency (correspond with “real-world”) of knowledge in the Knowledge Graph [32]. SHACL (Shapes Constraint Language), a data modelling language developed by the W3C Consortium, is used to describe and validate the graph. SHACL brings four kinds of constraints for Knowledge Graph validation: (i) node constraints; (ii) property constraints; (iii) inverse property constraints; and (iv) general constraints (general mechanism based on SPARQL query language). The use of SHACL needs expert help to create all kinds of constraints for domain-specific knowledge to ensure the knowledge objects would not bring any errors in common sense domain knowledge. For the EU-FarmBook platform, the knowledge objects are validated regarding the general uploading guidelines for the metadata model and specific agricultural domain knowledge guidelines.

5.5 EU-FarmBook Knowledge Graph Endpoint

The EU-FarmBook Knowledge Graph endpoint allows querying the RDF graph using the SPARQL query language. This endpoint provides interfaces for other APIs that use the EU-FarmBook Knowledge Graph, including the exploration of the EU-FarmBook Knowledge Graph and complex knowledge/data retrieval for a specific purpose. The EU-FarmBook Knowledge Graph provides an open SPARQL endpoint, which allows to query thwe graph using HTML-based query editors or programming languages. The SPARQL endpoint brings several benefits, including a wide variety of answer content types (e.g., RDF, HTML, JSON, CSV), compliance with user agents or servers, and ease of query because of the user-friendliness of SPARQL.

6. Conclusions and next steps

This report has presented the foundation of data standards used for the development of the EU-FarmBook platform. We have also discussed the construction and use of the Knowledge Graph derived from Knowledge Object metadata. Data standards will be updated and developed further as the platform grows and as the world and the language used to describe it changes. Subsequent versions of this deliverable will highlight the specific changes that are being adopted.

The EU-FarmBook ontology maintenance is work in progress. Updates of the ontology (to be reported in the follow-up releases of the present report) will focus, among others, on the more thorough capture of details related to the Research and Innovation projects producing Knowledge Objects, the topics covered in the Knowledge Objects housed in the EU-FarmBook platform, as well as user-related information. As part of our research activities in the context of the EU-FarmBook, we will also be developing tools for ontology and knowledge graph maintenance, visualisation, and evolution. The building and use of these tools will be reported on in subsequent versions of this report.

7. References

[1] M. D. Wilkinson *et al.*, 'The FAIR Guiding Principles for scientific data management and stewardship', *Scientific data*, vol. 3, p. 160018, Mar. 2016, doi: 10.1038/sdata.2016.18.

[2] L. Steele and T. Orrell, 'The frontiers of data interoperability for sustainable development', Publish What You Fund and Development Initiatives, Nov. 2017. Accessed: Oct. 05, 2023. [Online]. Available: https://devinit.org/resources/data-interoperability-sustainable-development/

[3] L. González Morales and T. Orrell, 'Data Interoperability: A Practitioner's Guide to Joining Up Data in the Development Sector', United Natons Statistics Division, 2018. [Online]. Available: https://unstats.un.org/wiki/display/InteropGuide/Home

[4] E. Goldstein, U. Gasser, and R. Budish, 'Data Commons Version 1.0: A Framework to Build Toward AI for Good', Berkman Klein Center Collection. Accessed: Oct. 05, 2023. [Online]. Available: https://medium.com/berkman-klein-center/data-commons-version-1-0-a-framework-to-build-toward-ai-for-good-73414d7e72be

[5] H. van der Veer and A. Wiles, 'Achieving Technical Interoperability - the ETSI Approach', ETSI, Sophia Antipolis, White Paper 3, Apr. 2008. [Online]. Available: https://www.etsi.org/images/files/ETSIWhitePapers/IOP%20whitepaper%20Edition%203%20final.pdf

[6] S. Dasgupta, 'The Semantic Problem in healthcare | LinkedIn', LinkedIn Pulse. Accessed: Oct. 05, 2023. [Online]. Available: https://www.linkedin.com/pulse/semantic-problem-healthcare-shaonli-dasgupta/

[7] E. E. Umberfield, C. J. Staes, T. P. Morgan, R. W. Grout, B. W. Mamlin, and B. E. Dixon, 'Chapter 9 - Syntactic interoperability and the role of syntactic standards in health information exchange', in *Health Information Exchange (Second Edition)*, B. E. Dixon, Ed., Academic Press, 2023, pp. 217–236. doi: 10.1016/B978-0-323-90802-3.00004-6.

[8] T. Berners-Lee, J. Hendler, and O. Lassila, 'The Semantic Web', *Scientific American*, pp. 30–37, May 2001.

[9] C. Bizer, T. Heath, and T. Berners-Lee, 'Linked Data - The Story So Far', *International journal on Semantic Web and information systems*, 2009, doi: 10.4018/jswis.2009081901.

[10] B. Mons, C. Neylon, J. Velterop, M. Dumontier, L. O. B. da Silva Santos, and M. D. Wilkinson, 'Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud', *Information Services & Use*, vol. 37, no. 1, pp. 49–56, Jan. 2017, doi: 10.3233/ISU-170824.

[11] S. Wolfert, L. Ge, C. Verdouw, and M.-J. Bogaardt, 'Big Data in Smart Farming--A review', *Agricultural systems*, vol. 153, pp. 69–80, 2017, doi: 10.1016/j.agsy.2017.01.023.

[12] B. Ali and P. Dahlhaus, 'The Role of FAIR Data towards Sustainable Agricultural Performance: A Systematic Literature Review', *Agriculture*, vol. 12, no. 2, Art. no. 2, Feb. 2022, doi: 10.3390/agriculture12020309.

[13] K. Kosior, 'Towards a New Data Economy for EU Agriculture', *Studia Europejskie - Studies in European Affairs*, vol. 23, no. 4, pp. 91–107, 2019.

[14] C. Caracciolo *et al.*, '39 Hints to Facilitate the Use of Semantics for Data on Agriculture and Nutrition', *CODATA*, vol. 19, no. 1, Dec. 2020, doi: 10.5334/dsj-2020-047.

[15] C. Jonquet *et al.*, 'AgroPortal: A vocabulary and ontology repository for agronomy', *Computers and Electronics in Agriculture*, vol. 144, pp. 126–143, Jan. 2018, doi: 10.1016/j.compag.2017.10.012.

[16] C. Brewster, 'The landscape of agrifood data standards: From ontologies to messages', in *EFITA WCCA 2017 Conference*, Jul. 2017. [Online]. Available: http://cbrewster.com/papers/Brewster_EFITA17.pdf

[17] M. L. Zeng, 'Knowledge Organization Systems (KOS)', *KO*, vol. 35, no. 2–3, pp. 160–182, 2008, doi: 10.5771/0943-7444-2008-2-3-160.

[18] S. Aubin *et al.*, 'Landscaping the use of semantics to enhance the interoperability of agricultural data', Research Data Alliance, report, Sep. 2017. Accessed: Apr. 15, 2021. [Online]. Available: https://hal.inrae.fr/hal-02607815

[19] S. Rajbhandari and J. Keizer, 'The AGROVOC concept scheme--a walkthrough', *Journal of integrative agriculture*, vol. 11, no. 5, pp. 694–699, 2012, [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2095311912600586

[20] D. Soergel, B. Lauser, A. Liang, F. Fisseha, J. Keizer, and S. Katz, 'Reengineering Thesauri for New Applications: the AGROVOC Example', Journal of Digital Information. Accessed: May 30, 2021. [Online]. Available: http://eprints.rclis.org/15694/

[21] T. Baker, B. Whitehead, R. Musker, and J. Keizer, 'Global agricultural concept space: lightweight semantics for pragmatic interoperability', *npj Science of Food*, vol. 3, no. 1, Art. no. 1, Sep. 2019, doi: 10.1038/s41538-019-0048-6.

[22] P. L. Buttigieg, N. Morrison, B. Smith, C. J. Mungall, S. E. Lewis, and the ENVO Consortium, 'The environment ontology: contextualising biological and biomedical entities', *Journal of Biomedical Semantics*, vol. 4, no. 1, p. 43, Dec. 2013, doi: 10.1186/2041-1480-4-43.

[23] P. L. Buttigieg, E. Pafilis, S. E. Lewis, M. P. Schildhauer, R. L. Walls, and C. J. Mungall, 'The environment ontology in 2016: bridging domains with increased scope, semantic density, and interoperation', *Journal of Biomedical Semantics*, vol. 7, no. 1, p. 57, Sep. 2016, doi: 10.1186/s13326-016-0097-6.

[24] B. Smith *et al.*, 'The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration', *Nature biotechnology*, vol. 25, no. 11, pp. 1251–1255, Nov. 2007, doi: 10.1038/nbt1346.

[25] D. M. Dooley *et al.*, 'FoodOn: a harmonized food ontology to increase global food traceability, quality control and data integration', *npj Science of Food*, vol. 2, no. 1, p. 23, Dec. 2018, doi: 10.1038/s41538-018-0032-6.

[26] L. G. Saldanha *et al.*, 'A structured vocabulary for indexing dietary supplements in databases in the United States', *Journal of Food Composition and Analysis*, vol. 25, no. 2, pp. 226–233, Mar. 2012, doi: 10.1016/j.jfca.2011.10.003.

[27] A. Hogan *et al.*, 'Knowledge Graphs', *arXiv:2003.02320 [cs]*, Jan. 2021, Accessed: May 28, 2021. [Online]. Available: http://arxiv.org/abs/2003.02320

[28] J. Sequeda and O. Lassila, *Designing and building enterprise knowledge graphs*. in Synthesis lectures on data, semantics, and knowledge, no. # 20. San Rafael: Morgan & Claypool Publishers, 2021.

[29] L. Richardson and S. Ruby, *RESTful Web Services*. Sebastopol: O'Reilly Media, 2008.

[30] S. Robbins, 'A beginner's guide to JSON, the data format for the internet - Stack Overflow', Stackoverflow Blog. Accessed: Oct. 05, 2023. [Online]. Available: https://stackoverflow.blog/2022/06/02/a-beginners-guide-to-json-the-data-format-for-the-internet/

[31] N. F. Noy and D. L. McGuiness, 'Ontology development 101 A guide to creating your first ontology', Stanford Knowledge Systems Laboratory, Stanford University, Stanford, CA, 2001. [Online]. Available: http://www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness-abstract.html

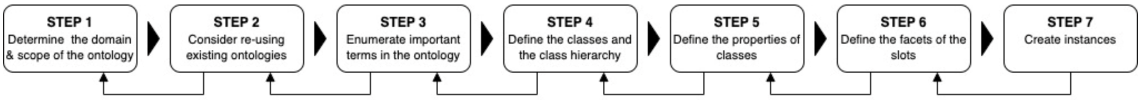
[32] E. Huaman, E. Kärle, and D. Fensel, D., 2020. Knowledge graph validation. *arXiv preprint arXiv:2005.01389*.

8. Annexes

Annex 1: Noy and McGuiness methodology for ontology development

The Noy and McGuiness methodology, which has been adopted for the development of the EU-FarmBook ontology is shown in Figure 4 below. This Annex section provides a brief description of each of the methodology steps.

Figure 4: Noy and McGuiness methodology for ontology development



Step 1: **Determine the ontology domain and scope**

The first step in the ontology building process is about framing the context of the task. It encompasses the identification of the ontology's purpose, as well as key questions (the so-called competency questions) that the ontology should provide answers for.

Step 2: **Consider re-using existing ontologies**

Noy and McGuiness [31] explicitly mention that "*reusing existing ontologies may be a requirement if our system needs to interact with other applications that have already committed to particular ontologies or controlled vocabularies*." (p. 6). Re-using concepts and relations already defined in existing ontologies helps towards data and information interoperability and exchange. This is important when considering machine agents as consumers of data. The need for enabling and enhancing the interoperability of data and information is emphasised in the FAIR data principles [1], which have gained increased attention in the last years as a good data sharing practice.

Step 3: **Enumerate important terms in the ontology**

The aim of this step is to identify a list of key terms relating to the purpose of the ontology, which help define the ontology classes (i.e., collections of entities with similar properties), relations between classes, and class properties.

Step 4: **Define the classes and class hierarchy**

A critical step in the ontology development process is the definition of classes and class hierarchies to: (i) group entities (i.e., the real-world objects or abstract concepts captured in our model) with the same properties into self-contained collections; and (ii) describe class-related subset-superset dependencies.

Step 5: **Define the properties of classes**

In the terminology introduced by the Web Ontology Language (OWL) specification there are two kinds of properties: (i) object properties; and (ii) datatype properties. Object properties describe relationships between class instances in a formal manner. To render relationships explicit, a list of relations (or better say, relation types) is introduced.

An object property has a domain and a range. Both the domain and range of an object property are (distinct) ontology classes. Datatype properties are properties receiving data values (e.g., integers, strings, floating point numbers, Boolean values).

Step 6: **Define the facets of slots**

Facets of the slots refer to the values that the datatype properties can take.

Step 7: **Create instances**

The final step in an ontology building process is that of creating class instances. Defining an instance requires to: (i) select a class; (ii) create an individual instance of that class; and (iii) provide values to the instance's properties.

Annex 2: EU-FarmBook Ontology

The classes of the EU-FarmBook ontology, their URIs, source ontologies/vocabularies, and short descriptions of them are available in Table 2 below.

Table 2: Classes defined in the EU-FarmBook ontology

Ontology class	URI	Ontology/ vocabulary	Description
`frapo:Output`	http://purl.org/cerif/ frapo/Output	FRAPO	A Knowledge Object (i.e., a project output).
`foaf:Project`	http://xmlns.com/foaf/0.1/	FOAF	The project in which the Knowledge Object was created.
`dct:Agent`	http://purl.org/dc/ terms/Agent	Dublin Core	The entity/-ties involved in the creation of a Knowledge Object (normally, one or more persons).
`ontopic:Subject`	http://www.ontology designpatterns.org/ ont/dul/ontopic.owl#Subject	Ontopic ontology	The subject of a Knowledge Object (i.e., what the Knowledge Object is about).
`schema: CreativeWork`	https://schema.org/ CreativeWork	schema.org	A generic kind of creative work, including documents, images, software programs, etc.
`schema: DigitalDocument`	https://schema.org/ DigitalDocument		An electronic file or document.
`schema: MediaObject`	https://schema.org/ MediaObject		A media object, such as an image, video, or audio object.
`schema:Place`	https://schema.org/ Place		The geographic location associated with a Knowledge Object's content.
`schema:Dataset`	https://schema.org/ Dataset		Structured information describing some topic(s) of interest.
`schema:Text DigitalDocument`	https://schema.org/ TextDigitalDocument		A file composed primarily of text.
`schema: Presentation DigitalDocument`	https://schema.org/ PresentationDigital Document		A file containing slides or used for a presentation.
`schema: AudioObject`	https://schema.org/ AudioObject		An audio file.
`schema: ImageObject`	https://schema.org/ ImageObject		An image file.

schema: VideoObject	https://schema.org/VideoObject		A video file.
schema:SoftwareApplication	https://schema.org/SoftwareApplication		A software application.

The object properties of the EU-FarmBook ontology, their URIs, source ontologies/vocabularies, and short descriptions of them are available in Table 3 below.

Table 3: Object properties defined in the EU-FarmBook ontology

Object property	URI	ontology/vocabulary	Description
frapo:isOutputOf	http://purl.org/cerif/frapo/isOutputOf	FRAPO	Property linking the Knowledge Object to the project it comes from.
dct:creator	http://purl.org/dc/terms/creator	Dublin Core	Property linking the Knowledge to its creator.
dct:subject	http://purl.org/dc/terms/subject	Dublin Core	Property links the Knowledge Object to the subject(s) it addresses.
schema:contentLocation	https://schema.org/contentLocation	schema.org	Property linking the content of the Knowledge Object to the geographic location the content relates to.
rdfs:subClassOf	https://www.w3.org/TR/rdf12-schema/ -ch_subclassof	RDF Schema	Property modelling the hierarchical relations between ontology classes.

The data properties of the EU-FarmBook ontology, their URIs, their source ontologies/vocabularies, and short descriptions of them are available in Table 4 below.

Table 4: Data properties defined in the EU-FarmBook ontology

Data property	URI	ontology/vocabulary	Description
dct:title	http://purl.org/dc/terms/title	Dublin Core	Property linking a Knowledge Object to its title.
dct:format	http://purl.org/dc/terms/format		Property linking a Knowledge Object to its file format.
dct:type	http://purl.org/dc/terms/type		Property linking a Knowledge Object to its type.
schema:description	https://schema.org/description	schema.org	Property linking a Knowledge Object to a short textual description.
schema:keywords	https://schema.org/keywords		Property linking a Knowledge Object to a list of keywords.
schema:inLanguage	https://schema.org/inLanguage		Property linking a Knowledge Object to the language of its content.

`schema:dateCreated`	https://schema.org/dateCreated		Property linking a Knowledge Object to the date of its creation.
`schema:fileSize`	https://schema.org/fileSize		Property linking a Knowledge Object to its file size.
`schema:license`	https://schema.org/license		Property linking a Knowledge Object to a license specifying reuse terms and conditions.
`schema:potentialAction`	https://schema.org/potentialAction		Property linking a Knowledge Object to the purpose it is meant to be used for.
`schema:name`	https://schema.org/name		Property linking a project, Knowledge Object creator, or geographic location (related to a Knowledge Object's content) to its name.
`schema:alternateName`	https://schema.org/alternateName		Property linking a project to its acronym.
`schema:url`	https://schema.org/url		Property linking a Knowledge Object to its URL.

In the above tables, prefixes are used to provide readable class and property definitions. The links between prefixes and the namespaces they are used for are shown in Table 5 below.

Table 5: Prefixes used in class and property definitions, and their namespaces

Prefix	Namespace
schema	http://schema.org/
frapo	http://purl.org/cerif/frapo/
foaf	http://xmlns.com/foaf/0.1/
ontopic	http://www.ontologydesignpatterns.org/ont/dul/ontopic.owl#
dct	http://purl.org/dc/terms/