H2020-SFS-2018-2020

DECIDE

Data-driven control and prioritisation of non-EU-regulated contagious animal diseases

Deliverable 1.2

Species-specific ontologies

WP1 – Data identification, characterisation and acquisition

Authors	Saba Noor (UGent), Miel Hostens (UGent)
Lead participant	UGent
Delivery date	v1.0: 29 December 2022 v2.0: 27 December 2023 v3.0: 14 May 2025
Dissemination level	Public
Type	Report & Online published ontologies

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101000494.

Revision History

Author Name (Partner short name)	Description	Date
Miel Hostens (UGent), Saba Noor (UGent)	Draft deliverable	13.12.2022
Gerdien van Schaik (UU)	Revision 1	14.12.2022
Miel Hostens (UGent), Saba Noor (UGent)	Updated deliverable	20.12.2022
Fernanda Dorea (SVA)	Revision 2	22.12.2022
Miel Hostens (UGent), Saba Noor (UGent)	Updated deliverable	29.12.2022
Johannes Ripperger (accelCH)	Quality check and layout	29.12.2022
Miel Hostens (UGent)	Final version (v1.0)	29.12.2022
Emmanouil Liandris (REA) *(Project Officer)*	Request for revision *[Ref. Ares(2023)3313496 – 11/05/2023]*	11.05.2023
Miel Hostens (UGent), Saba Noor (UGent)	Draft update of v1.0	11.12.2023
Jeroen Degroote (UGent)	Revision	12.12.2023
Fernanda Dórea (SVA)	Revision	15.12.2023
Saba Noor (UGent)	Updated deliverable (v2.0)	19.12.2023
Johannes Ripperger (accelCH)	Quality check and layout	27.12.2022
Saba Noor (UGent)	Updated deliverable (v3.0)	02.05.2025
Johannes Ripperger (accelCH)	Quality check and layout	14.05.2025

Content

Abbreviations

Abbreviation	Description
ABSMSA	Agent-Based Semantic Model for Smart Agriculture
ADO	Animal Disease Ontology
AHOL	Animal Health Ontology for Livestock
AHSO	Animal Health Surveillance Ontology
ANDO	Animal Disease Ontology
ASReview	Active Learning for Systematic Reviews
ATOL	Animal Trait Ontology for Livestock
BCV	Bovine coronavirus
BRSV	Bovine respiratory syncytial virus
CIDO	Coronavirus Infectious Disease Ontology
DCPO	Dairy Cattle Performance Ontology
DFO	Dairy Farming Ontology
DTRA	Defense Threat Reduction Agency
EU	European Union
FAIR	Findable, Accessible, Interoperable, and Reusable
FAO	Food and Agriculture Organization
GDPR	General Data Protection Regulation
H2020	Horizon 2020
HS	Histophilus somni
IB	Infectious bronchitis
IBIS	International Biosurveillance System
IDO	Infectious Disease Ontology
LHO	Livestock Health Ontology
M	Milestones
MB	Mycoplasma bovis
MH	Mannheimia haemolytica
NCIT	National Cancer Institute Thesaurus
O	Objectives
ODKFADM	Ontology-driven knowledge-based framework of Farm Animal Data Management

OWL	Ontology web language
PH	Pulmonary hypertension
PLF	Precision livestock farming
PM	Pasteurella multocida
PRRS	Porcine reproductive and respiratory syndrom
REA	Research Executive Agency
RDF	Resource Description Framework
SET	Surveillance Evaluation Tool
SI	Swine influenza
SPARQL	SPARQL Protocol and RDF Query Language (recursive acronym)
T	Task
VIDO	Virus Infectious Disease Ontology
WOAH	World Organization for Animal Health
WP	Work Package

Partner short names

Short name	Organisation
UU	Universiteit Utrecht
UGent	Universiteit Gent
SVA	Swedish Veterinary Agency
accelCH	accelopment Schweiz AG

Executive Summary

This deliverable serves as a comprehensive guide for the Livestock Health Ontology (LHO), a species-specific ontology developed for terrestrial (Cattle, Pig, Poultry) and aquatic species adhering to FAIR principles and GDPR.

Objectives of the deliverable

The main objective of this deliverable is:

- Developing a species-specific ontology called Livestock Health Ontology.
- Generating RDF data for Livestock use cases.
- Mapping RDF data to LHO for semantic interoperability.

Current activities:

- **Systematic literature review utilising ASReview:** Utilising the ASReview tool for a systematic literature review, distinguishing active and non-active ontologies and exploring potential extensions.
- **Livestock Health Ontology (LHO) Development:** The DECIDE ontology was extended to create the Livestock Health Ontology (LHO), expanding coverage to pigs, poultry, and cattle while continuing to support the salmon use case. LHO is openly accessible on BioPortal and AgroPortal, with the full code available on GitHub to promote transparency, reuse, and community feedback.
- **Ontology Tutorials and documentation:** Develop easy-to-follow documentation and tutorials to simplify the understanding and usage of the ontology
- **Organizing Workshop & Training at ECPLF**
- **Publication:** Published a book chapter (*Agri Semantics*) and conference papers showcasing ontology development and its application in livestock health use cases.

Outcome

- **The LHO Ontology**, which covers cattle, pigs, poultry, and salmon, is accessible on AgroPortal and BioPortal, with supporting R, PySpark, and pandas notebooks available on GitHub for code generation across use cases (GitHub repository).
- **Ontology and RDF data implementation documentation and tutorials can be accessed at**: GitHub Repository and the DECIDE project tutorials page.
- **Workshop** on "Understanding Semantic Web and Ontologies for Precision Livestock Farming," organized by Ghent University, was held at ECPLF 2024. Program | Agenda
- **Book Chapter Publication:** https://biblio.ugent.be/publication/01HZ08G35DJNW77HZP2X1P6H6B
- **Conference publication:** https://biblio.ugent.be/publication/01JB4AJ4WCTQK85038GAQEEPFH
- **DECIDE case studies website** showcases the transformation of cattle, poultry, pig, and salmon data from European labs. View case studies.

Next steps

- Test three data access strategies: Direct Data Sharing - Manual exchange via email, and process in Python scripts, Centralized Data Exchange using cloud-based pipelines, and decentralized access using Solid Pods (Federated).
- After that, the focus is on applying a federated learning prototype that uses ontology-driven knowledge discovery to predict future trends.

1 Introduction

Livestock farming is an essential contributor to global food production and rural economies, but it faces increasing challenges due to the spread of contagious diseases. These diseases, caused by multiple pathogens, affect both terrestrial and aquatic species and can significantly impact animal welfare, food security, and economic stability. Effective disease surveillance is essential for early detection and response, but it depends on large volumes of data from diverse sources, including farm sensors, laboratory tests, and national databases. Within the DECIDE project, the focus is on improving disease monitoring for cattle, poultry, pigs, and salmon, species that are central to meat production and aquaculture in Europe. The project began with species for which data was already available, particularly cattle, poultry, pig, and salmon, to support the development of a decision support tool. However, the data used in disease surveillance is often stored in different formats, uses inconsistent terminology (data heterogeneity), and is scattered across multiple systems. This creates challenges in interoperability (how well systems can communicate), integration (combining data from various sources), and scalability (managing large and growing datasets). These limitations restrict efficient data sharing, reusing, and cross-species comparisons. Addressing these challenges requires better ways to structure, link, and interpret disease data across species and systems.

To address these challenges, we need to design a cost-effective, accessible, reusable, and lightweight ontology-driven knowledge-based domain-specific framework for farm animal health management systems. This report focuses on 4 use cases that serve as representative examples for the application of the proposed ontology-driven knowledge-based framework. The Livestock Health Ontology (LHO) was developed as an extension of the DECIDE core ontology to meet this need, LHO focuses on key terrestrial species (cattle, pigs, poultry) and maintains coverage for salmon, offering a harmonized model for animal health terminology, disease states, diagnostics, and management practices.

The key contributions of this report include:

- **Conducting a systematic literature review** to identify existing ontologies related to animal health and disease surveillance, with a focus on potential reuse, extension, and alignment with established resources.

- **Developing and publishing the Livestock Health Ontology (LHO)** as an extension of the core DECIDE ontology, covering key species including cattle, poultry, pigs, and salmon.

- **Designing a standardized ontology-driven data management framework (ODKFADM)** to organize, map, and analyze livestock data based on FAIR principles. This framework balances data availability, accessibility, connectivity, and value generation.

- **Evaluating the ODKFADM framework** across four real-world use cases as a proof of concept, demonstrating its application in ontology-driven decision support for livestock health management.

- **Supporting the adoption and reuse of LHO** through detailed documentation, training workshops, and public code repositories.

The remaining sections of the report are organised as follows: Section 2 provides a systematic literature review on existing ontologies for animal health and disease surveillance. Section 3 presents the methodology in detail. In Section 4, several case studies are described, integrating and analysing data from multiple animal health sources. Section 5 concludes the report, discussing future research goals.

2 Systematic literature review with ASReview

This section outlines the systematic literature review methodology by integrating the PRISMA 2020 framework with the ASReview tool. ASReview utilizes machine learning (ML) to streamline study selection by suggesting the most relevant articles from a large dataset (Quan et al., 2024). Figure 1 illustrates how PRISMA steps (identification, screening, and inclusion) integrate with the ASReview process, enhancing efficiency by prioritizing key studies and streamlining manual screening for faster, more accurate paper selection.

This review focused on ontology-based approaches to animal health management for disease surveillance. A literature search was conducted across multiple databases, including PubMed, IEEE Xplore, and Google Scholar, covering research published between 2011 and 2025. The search strategy used a comprehensive set of keywords related to ontology and ontology-based systems (e.g., "ontology," "ontology-based," "livestock health ontology"), livestock species (e.g., "cattle," "pigs," "poultry," and "salmon"), animal health and infectious diseases (e.g., "disease surveillance," "health surveillance," "infectious diseases," "early detection"), farming, agriculture farming, precision livestock farming and farm management, as well as data-related challenges (e.g., "data sharing," "data governance," "data integration").

The search yielded 294 articles, which were reduced to 286 unique articles after removing duplicates. However, it's worth noting that such search criteria often result in a mix of papers, some directly related to our field of study, while others less. To address this issue and ensure the highest relevance level, we employed the ASReview tool, an automated systematic review software, which allowed us to further refine our selection. The next step involved organizing the gathered papers in structured formats (e.g., CSV, RIS), ensuring that bibliographic information, abstracts, and keywords were captured. After installing ASReview, we set up a project with a dataset. The ASReview software utilized ML to suggest papers for review, which we classified as either "relevant" or "not relevant." Initially, 28 relevant and 10 irrelevant prior knowledge records were used to train the AI model. These prior records were excluded from the active screening pool. The remaining 248 records were then screened using ASReview's active learning interface. The AI tool excluded 54 records as irrelevant based on the model's predictions. As more papers were reviewed, ASReview adapted and improved its recommendations. From the remaining 194 records (248 - 54), reports were sought for retrieval. Of these, 20 records were manually reviewed during full-text screening, and a total of 122 records were excluded by humans because they did not focus on animal disease surveillance or health management. Once predefined criteria were met, the tool generated a final list of relevant articles, allowing us to efficiently review 286 records and identify 100 papers for further analysis. Citation searching was not performed as part of this review. Details of the search strategy, including database queries, specific keywords, and a description of the ASReview project setup, are provided in the supplementary materials. The supplementary materials also include instructions for replicating the review process. These resources are available on Zenodo DOI 10.5281/zenodo.14604329.

Following the ASReview process, we developed a taxonomy of animal health management approaches, categorizing the selected papers into ontology-based and non-ontology-based approaches, as shown in Figure 2. Both are essential for disease surveillance and early detection. Ontology-based approaches use structured frameworks to enhance data integration, analysis, and management by conceptualizing domain knowledge (Kang et al., 2018). These frameworks support data organization, sharing, and standardization for improved decision-making (Hu et al., 2020).

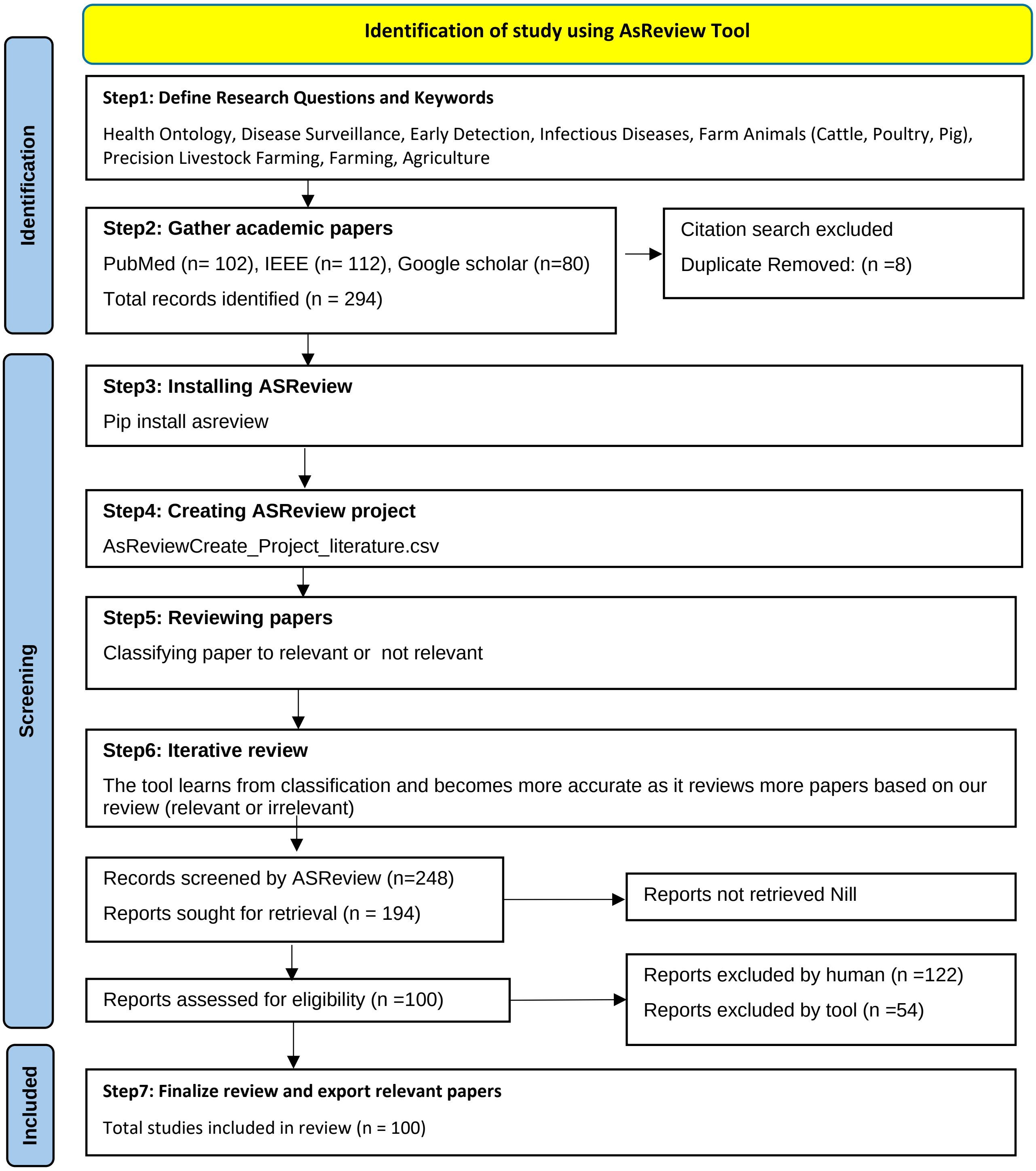


Figure 1: Methodology for integrating the state-of-the-art PRISMA 2020 framework into ASReview for systematic literature review.

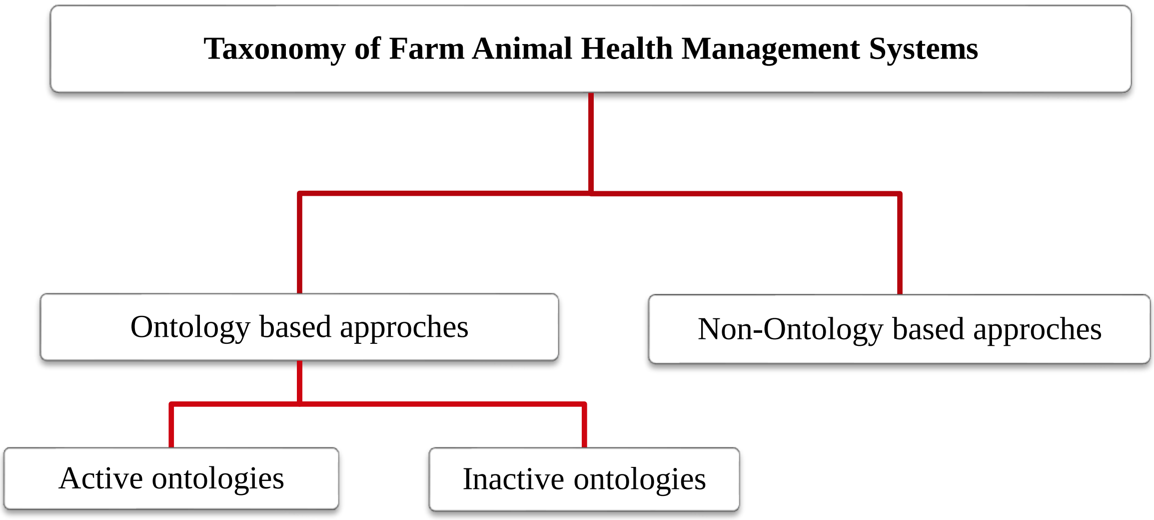


Figure 2: Taxonomy of farm animal health management approaches.

A comparative analysis of active and inactive ontologies is presented in Table 1 and Table 2. An active ontology is publicly accessible, documented, and actively maintained for reuse in its intended domain. In contrast, we categorized ontologies as inactive in our study due to incomplete citations, proprietary designs, access restrictions, or uncertain public availability; however, this classification highlights challenges in locating and utilizing these ontologies rather than definitively confirming their inactivity, ultimately limiting their reuse, broader adoption, integration into other systems, and contribution to livestock disease management and agricultural advancements. To further refine active ontologies, we categorize them based on their focus areas, specifically targeting infectious diseases and surveillance ontologies in farm animals to enhance animal health management strategies. On the other hand, non-ontology-based approaches include traditional data-driven decision support tools and ML models, such as those for disease surveillance and disease prediction, which are widely used in livestock farming and management systems (Nadar et al., 2023).

2.1 Ontology-based approaches

In ontology-based approaches, Table 1 reveals that out of 32 identified ontologies, only 18 remain active, while the other 14 are inactive. This categorization provides insight into the current usability and adoption of ontologies in terrestrial livestock management systems. The high number of inactive ontologies suggests challenges in long-term sustainability, interoperability, or practical application in real-world settings.

2.1.1 Active ontologies

SKOS Vocabularies in Agriculture and Animal Health

In the field of livestock farming and their management, numerous controlled vocabularies and ontological sources have been developed to support interoperability and standardization (Kwok et al., 2024). Among these, AGROVOC agriculture thesaurus discussed by Caracciolo et al. (2013), is one of the most widely adopted. It is maintained by the Food and Agriculture Organization (FAO) of the United Nations. AGROVOC aligns with the FAIR principles and is linked to open data source for agriculture, consisting of over 3200 concepts and available in 20 different languages. It serves as a valuable resource for capturing general concepts regarding agriculture and animal health.

Animal health ontologies

These ontologies focus on animal diseases, health conditions, and management. For example, Animal Disease Ontology (ANDO) discussed by Fauré et al. (2018), available in both French and English, focuses on representing knowledge of animal diseases specifically focusing on 500 pairs of genetic disorders and infectious and metabolic diseases across different animal species. It addresses concepts regarding disease treatment, pathogens, diagnostic tests, clinical signs, and progression. The disease ontology discussed by Schriml et al. (2022) serves as a standardized and open-source system that facilitates integration with various biomedical vocabulary, supporting multiscale data integration and analysis. The ontology includes entries for various animal diseases, including respiratory illnesses such as swine influenza in pigs, providing detailed information on their characteristics, transmission methods, and associated clinical signs. The Animal Health Ontology for Livestock (AHOL) described by Salaun et al. (2019), describes production diseases. It is structured around disease types (communicable, non-communicable, genetic, metabolic, physical, and psychological) and classifies diseases based on associated clinical signs, affected organisms (livestock, poultry, mammals, fishes), and causative agents (virus, bacteria, fungus, parasite).

Animal trait ontologies

These ontologies address phenotypic traits in animals, particularly regarding production and welfare. For example, Animal Trait Ontology for Livestock (ATOL) documented by Golik et al. (2012) is an ontology that models a multi-species ontology for animal phenotype traits. It focuses on animal products e.g. meat, milk quality and quantity, reproduction, and welfare. Expanding on this direction, Hulsegge et al. (2012) presented a domain-specific framework and developed two first-generation ontologies: the Reproductive Trait and Phenotype Ontology (REPO) and the Host Pathogen Interactions Ontology. Their work focused on female fertility in cattle and pig-Salmonella interactions, illustrating how ontologies can support genomic studies by enabling the identification of candidate genes for traits like fertility. These ontologies, along with others like those described by Tian et al. (2020), build on and reuse existing structures such as the Plant Trait Ontology (PTO) and ATOL, further advancing the integration of trait and genomic data to support animal breeding and health research. The Animals in Context Ontology (ACO) discussed by Santamaria et al. (2012) is a formal ontology developed by adapting a subset of Systematized Nomenclature of Medicine - Clinical Terms (SNOMED) to align with Open Biological and Biomedical Ontologies (OBO) Foundry Principles. It has practical uses in science, medicine, and agriculture, and integrates with other ontologies like Phenotypic trait ontology (PATO), National Center for Biotechnology Information (NCBI) Taxonomy, Environment Ontology (EnvO) , and Genome Ontology (GO).

Disease Surveillance ontologies and tools

The surveillance ontologies, such as the Animal Health Surveillance Ontology (AHSO) developed by Dórea et al. (2019) exemplify the use of ontology in decision-making based on diverse data sources, such as clinical records and laboratory findings. AHSO accommodates the entire animal production cycle, capturing information beyond disease events, and focuses on surveillance analytics utilizing observations like disease occurrences, births, and product yield. The ontology comprises three main levels: sample, observations, and observational context, modelling health events as abstract concepts. The International Biosurveillance System (IBIS) introduced by Valentin et al. (2021) and developed by the Defense Threat Reduction Agency (DTRA) in the United States, focuses on monitoring biological events and activities worldwide. It aims to detect and assess potential biological threats to national security and facilitating data integration and behavioural studies. As a platform for automated extraction of disease information, the system was specifically designed to monitor animal disease outbreaks using web-based sources, such as online news media.

Infectious disease ontologies

These ontologies are designed to provide standardized frameworks for capturing diverse data related to infectious agents, disease processes, and their effects on various species, including humans and animals. Notable ontologies such as Infectious Disease Ontology (IDO) discussed by Cowell and Smith, (2010), the

Coronavirus Infectious Disease Ontology (CIDO) described by He et al. (2022, 2020) and Zheng et al. (2023) and the Virus Infectious Disease Ontology (VIDO) discussed by Babcock et al. (2021), focus on identifying and categorizing pathogens, their hosts, and associated disease processes across different species, including humans, animals, and plants. Additionally, specialized ontologies like the BioCaster Ontology (BCO) documented by Doan et al. (2008), and Infectious Transmission Ontology (ITO) developed by Slavco et al. (2022) support surveillance by mapping outbreaks, transmission routes, and geographic distribution, while the DECIDE ontology emphasizes the impact of diseases on livestock and aquaculture systems. These ontologies enhance our ability to track, prevent, and respond to infectious disease outbreaks across species and environments. Thus, overcoming format and structure challenges and enabling effective analysis in global disease monitoring and prevention (Zeginis et al., 2019).

Other ontologies:

For livestock behavioural studies, breed classification, health, and tracking, several active ontologies contribute to structured data representation and integration. The Dairy Cattle Performance Ontology (DCPO) proposed by Fuentes et al. (2022) aims to optimize dairy farming through computerized data analysis. Although the ontology integrates diverse information on animal health, nutrition, yield, and genetics, it does not focus on infectious diseases. The MoonCAB ontology described by N. Hammouda et al. (2023) is a modular ontology for analysing livestock behaviour (sheep and goats) in pastures, capturing activity duration, seasonal patterns, and interpretations. The Vertebrate Breed Ontology (VBO) proposed by Mullen et al. (2025) focuses on standardizing breed classification for livestock and other domestic animals. While these ontologies enhance livestock research and management, they do not specifically address infectious disease surveillance or disease management, which is the focus of Section 3.1.1, exploring relevant ontologies for infectious disease surveillance and livestock health management.

Table 1: Comparative analysis of the active ontological sources along with Domains of knowledge covered include species, pathogens, traits, behaviour, and environmental factors.

Reference	Ontology Name	Status	Domain Knowledge covered
SKOS Vocabularies in Agriculture and Animal Health:			
(Caracciolo et al., 2013)	AGROVOC	Active	Covering various aspects of crops, livestock, and broad terms related to livestock diseases and pathogens
Animal health ontologies:			
(Fauré et al., 2018)	ANDO	Active	Encompasses a variety of production animals, including livestock species such as cattle, pigs, poultry, and other animals relevant to agricultural systems. Pathogens for viruses, bacteria, fungi).
(Schriml et al., 2012)	DO	Active	Focused on categories like neoplasms, infectious diseases, and genetic diseases and causes of diseases, such as genetic mutations, treatments, risk factors, and preventive measures like vaccination and screening.
(Salaun et al., 2019)	AHOL	Active	Covers mammals, and fish; traits for dairy, meat; and pathogens (viruses, bacteria, fungi, parasites) as causative agents
Animal trait ontologies:			
(Golik et al., 2012)	ATOL	Active	Includes multiple livestock species (cattle, pigs, chickens) for cross-species trait analysis. Focuses on traits related to growth, behaviour, stress, carcass quality, and animal welfare, providing a standardized resource for trait descriptions.
(Hulsegge et al., 2012)	Reproductive Trait and Phenotype Ontology (REPO) and the Host Pathogen Interactions Ontology	Active	cover key domains including livestock species (such as cattle and pigs), pathogens (like *Salmonella*), reproductive traits (with a focus on female fertility), behaviour related to animal health, and environmental factors that influence host-pathogen interactions.
(Santamaria et al., 2012)	Animals in Context Ontology (ACO)	Active	Upper-level ontology that captures various characteristics of animals, including physiological features, age, diet, and living environment.
Disease Surveillance ontologies and tools			
(Dórea et al., 2019)	AHSO	Active	The upper-level ontology covers Mammals, birds, and fish. Pathogens include viruses, bacteria, Dairy, and meat production systems that facilitate comprehensive surveillance and detailed tracking of disease outbreaks across different animal populations.
(Valentin et al., 2021)	IBIS	Active	The focus is on livestock, aquaculture and wildlife species, tracking pathogens such as viruses, bacteria, fungi, and parasites. It covers traits like disease clinical

			signs while incorporating animal behaviour patterns and environmental factors influencing disease spread.
Infectious disease ontologies:			
(Cowell and Smith, 2010)	IDO	Active	It covers humans, animals, and plants. Pathogens include viruses, bacteria, and fungi. Traits focus on disease processes, infections, and immunity.
(He et al. 2022)	CIDO	Active	Primarily focuses on human diseases, particularly those caused by coronaviruses
(Babcock et al., 2021)	VIDO	Active	encompasses viral infectious diseases across various species, including humans, animals, and plants
(Doan et al., 2008)	BCO	Active	A broad range of species, including humans, and animals, includes information on various infectious agents, such as viruses, bacteria, fungi,
(Slavco et al., 2022)	Infectious transmission ontology (ITO)	Active	Publicly not available
(Slavco and Noor, 2022)	DECIDE	Active	Focuses on livestock and aquaculture species. analyse the impact of animal diseases on livestock productivity, tracking pathogens such as viruses, covers traits like disease clinical signs, and diagnostic markers, Addresses various livestock production systems (e.g., dairy, meat, aquaculture)
Other ontologies			
(Fuentes et al., 2022)	DCPO	Active	Focus on dairy cattle, traits like milk production (volume, composition, quality), reproductive performance (fertility rates, calving intervals), health indicators (disease incidences, somatic cell counts), behavioural traits (temperament, social interactions) and Environmental factors (housing conditions, climate impacts) included.
(Hammouda et al., 2023)	MoonCAB Ontology	Active	Primarily focuses on livestock species, including sheep and goats with pastures behaviour. Considering environmental factors influencing animal behaviour, such as habitat conditions and management practices.
(Mullen et al., 2025)	VBO	Active	VBO standardizes breed classification across vertebrate species, including cattle, horses, chickens, dogs, and cats. Each breed is uniquely represented and linked to its genus and species based on NCBI Taxonomy. It maps connections between foundation stock and derived breeds, aiding genetic lineage analysis and supporting advanced data interoperability in livestock research.

2.1.2 Inactive Ontologies

Despite advancements in ontology development for livestock management, several ontologies are inactive or no longer maintained as shown in Table 2. Some of these ontologies were specifically designed for disease surveillance, while others focused on disease management or broader health management in livestock.

In the context of disease surveillance and disease management, a rule-based ontology model proposed by Kang and Choi (2011) utilized clinical signs and sensor data for predicting livestock diseases. Later, Kang et al. (2018) developed an ontology for predicting livestock diseases and addressing a national foot-and-mouth disease outbreak, demonstrating the role of ontologies in tracking and managing disease events, though it did not progress beyond initial development. Similarly, the Animal Disease and Symptom Ontology by Da Silva et al. (2019) linked animal diseases to symptoms, facilitating improved classification and diagnosis. The SE-DiagEnf ontology-based expert system, developed by Alarcón-Salvatierra et al. (2018), was specifically designed for cattle disease diagnosis based on clinical signs, further supporting veterinary decision-making by providing treatment recommendations. Beyond disease-focused ontologies, some frameworks have contributed to general livestock health management. SAAONT, as discussed by Alreshidi (2020), was developed for plant and animal health management in Saudi Arabia, incorporating broader health-related concepts beyond infectious disease tracking. The Monarch Initiative by Shefchek et al. (2020) aimed to integrate genetic and disease-related data across species, bridging the gap between genetic research and disease ontology applications. Meanwhile, context-based ontology models by Sivamani et al. (2015, 2016a, 2016b) focused on livestock nutrition monitoring and diet planning using sensor data and fuzzy logic, though they remain unavailable for public use.

Some ontologies were designed to address interoperability challenges, precision farming, agricultural data management, and general farming systems, but they did not specifically focus on disease surveillance or disease management. For example, AgroRDF was developed by Martini et al. (2013), AgriOpenLink was proposed by Tomic et al. (2013), Dairy Farming Ontology (DFO) was created by Tomic et al. (2015), Common Dairy Ontology (CDO) was developed by Fuentes et al. (2021), APODOCSIIS and its domain ontology were discussed by Zhu et al. (2009), and Agent-Based Semantic Model for Smart Agriculture (ABSMSA), incorporating SAGRO-Lite, a lightweight ontology, was proposed by Shankhdhar et al. (2021). These ontologies were developed to support agricultural decision-making, knowledge graphs, farm data integration, and policy management. However, their scope did not extend to structured disease tracking, pathogen monitoring, or surveillance analytics.

Table 2: Comparative analysis of inactive ontological sources along with Domains of knowledge covered.

Reference	Ontology Name	Status	Domain Knowledge covered
(Kang and Choi, 2011)	Rule Based ontology	Inactive	Claims to predict livestock disease
(Kang et al., 2018)	Foot-and-mouth disease outbreak by using ontology	Inactive	Claims Ontological model for foot-and-mouth disease outbreak.
(Da Silva et al., 2019)	Animal Disease and Symptom Ontology	------	Claims Ontological model for classification of diseases in bovines and buffaloes
(Alarcón-Salvatierra et al., 2018)	SE-DiagEnf: An ontology-based expert system	Inactive	Claims ontology-based expert system that is used for cattle disease diagnosis and treatment recommendations
(Alreshidi, 2020)	SAAONT	---------	focuses on crops and livestock, addresses plant diseases, and environmental factors like climate conditions and soil types are considered
(Shefchek et al., 2020)	Ontology-based approach along with fuzzy logic for livestock species	Inactive	Claims to integrate genetic and disease data by using ontologies across species.
Sivamani et al. (2015, 2016a, 2016b)	Context-based ontology model to detect the amount of nutrition supply in cows	Inactive	Claims to use ontology for livestock nutrition monitoring and diet planning
(Martini et al., 2013)	AgroRDF	Inactive	Focuses on improving semantic interoperability in the agrifood sector by enhancing data exchange across agricultural systems.
(Tomic et al., 2013)	Agri Openlink	Inactive	Claims a plug-and-play architecture for adaptive optimization of agricultural processes through open interfaces, linked data, and semantic services.
(Tomic et al., 2015)	DFO	Inactive	Claims that DFO serves as a knowledge graph, enabling the encoding and manipulation of dairy farming domain knowledge.
(Fuentes et al., 2021)	CDO	Inactive	ÇDO is designed to support decision-making in dairy farming by enhancing semantic alignment across diverse data sources
(Symeonaki et al., 2022)	An ontology-based IoT middleware approach for smart livestock farming	Inactive	Claims to use ontology as IOT middleware for smart livestock farming
(H. Zhu et al., 2009)	APODOCSIIS system and its domain ontology	Inactive	Facilitates the automatic semantic integration of agricultural policy texts, covering tasks and semantic alignment of agriculture policy documents
(Shankhdhar et al., 2021)	SAGRO lite	Inactive	A lightweight ontology designed for farming in developing countries

2.1.3 Relevant ontologies for animal health and disease

This section focuses on active and relevant ontologies coverage related to infectious diseases and surveillance systems.

Table 3: Ontology coverage for contagious respiratory diseases in farm animals.

Ontology Name	Pig Diseases			Poultry Diseases	Cattle Diseases							Salmon Disease	
	PRRS[1]	SI[2]	M. hyopneumoniae[3]	IB[4]	PM[5]	PH[6]	HS[7]	MH[8]	MB[9]	BCV[10]	BRSV[11]	ISAV[12]	L. salmonis[13]
AGROVOC	✓	✓	✓	✕	✓	✓	✕	✕	✓	✕	✓	✓	✓
AHOL	✓	✓	✓	✕	✓	✕	✕	✕	✓	✕	✕	✕	✓
IDO	✕	✓	✕	✓	✕	✕	✕	✕	✕	✕	✕	✕	✕
CIDO	✕	✓	✕	✓	✕	✕	✕	✕	✕	✕	✕	✕	✕
VIDO	✕	✓	✕	✕	✕	✕	✕	✕	✕	✕	✕	✕	✕
AHSO	✕	✓	✕	✕	✕	✕	✕	✕	✕	✕	✕	✕	✕
DECIDE	✕	✕	✕	✓	✓	✕	✕	✕	✓	✓	✓	✕	✓
ANDO	✓	✓	✓	✓	✓	✕	✓	✓	✓	✓	✓	✓	✓
DO	✕	✓	✕	✕	✕	✕	✕	✕	✕	✕	✕	✕	✕

[1] PRRS – Porcine Reproductive and Respiratory Syndrome
[2] SI – Swine Influenza
[3] M. hyopneumoniae –Mycoplasma hyopneumoniae
[4] IB – Infectious Bronchitis
[5] PM – Pasteurella multocida
[6] PH – Pasteurella haemolytica
[7] HS – Histophilus somni
[8] MH – Mannheimia haemolytica
[9] MB – Mycoplasma bovis
[10] BCV – Bovine Coronavirus
[11] BRSV – Bovine Respiratory Syncytial Virus
[12] Infectious Salmon Anaemia Virus (ISAV)
[13] Lepeophtheirus salmonis (L. salmonis)
[14] Checkmarks (✓) denote the presence of information for the corresponding disease in the ontology, while crosses (✕) indicate its absence.

Table 3 demonstrates a higher convergence of infectious disease data for cattle, pigs, poultry, and salmon, highlighting the relevance of certain ontologies and emphasizing the importance of specific ontologies in addressing the global burden of contagious diseases. These diseases were selected based on their high economic impact, global prevalence, and significant role in livestock mortality, such as orcine Reproductive and Respiratory Syndrome (PRRS) virus, Swine Influenza (SI), and Mycoplasma hyopneumoniae (M. hyopneumoniae) in pigs; Infectious Bronchitis (IB) in poultry; and a range of bovine respiratory diseases in cattle including Pasteurella multocida (PM), Pasteurella haemolytica (PH), Histophilus somni (HS), Mannheimia haemolytica (MH), Mycoplasma bovis (MB), Bovine Coronavirus (BCV), and Bovine Respiratory Syncytial Virus (BRSV) and Infectious Salmon Anaemia Virus (ISAV) in salmon highlight the need for up-to-date ontologies that focus on these critical livestock health issues. The table explores the details of disease coverage within each ontology, revealing their strengths and limitations in understanding various health conditions.

In our assessment, AGROVOC, AHOL, AHSO, IDO, DECIDE, and ANDO are focusing primarily on infectious diseases and pathogens. They were identified as pivotal for extending or reusing within a data-driven decision-support system for managing infectious diseases. However, these ontologies differ in scope, structure, and species coverage. For example, AHSO, AHOL, and DECIDE, as upper-level ontologies, establish foundational concepts crucial for structuring domain-specific ontologies. AHSO captures individual health event data, ensuring interoperability in surveillance. DECIDE structures pathogen tracking in livestock and aquaculture species, integrating disease, clinical signs, sample type, location, and diagnostic details, thus supporting global disease burden assessment by standardizing pathogen, host, and clinical sign data for infectious disease surveillance. AHOL organizes phenotypic traits, linking disease groups, symptoms, and affected species for structured data management. The selective reuse of AGROVOC facilitates the incorporation of well-established agricultural and infectious disease concepts. While not a formal ontology, its hierarchical thesaurus structure supports semantic consistency and can complement ontology development by providing standardized terminology. ANDO directly addresses disease management, making it a valuable component in the broader ontology ecosystem for animal health.

2.2 Non-ontology-based approaches

During the ASReview process, we reviewed traditional methods for animal disease surveillance. Llaria et al., (2024); Shukla et al., (2023) highlight the role of ML models in monitoring animal behaviour, while Heller et al., (2016) emphasize the value of disease surveillance tools for outbreak insights. Moreover, surveillance evaluation tools such as RiskSUR, discussed by Peyre et al. (2019), and the SET tool, discussed by Vasconcelos Gioia et al. (2021), play a crucial role in assessing national surveillance capacities for animal health threats. However, these systems often lack real-time analysis and data interoperability, limiting their scalability and adaptability. The Southern African Center for Infectious Disease Surveillance (SACIDS) initiative was evaluated under the One Health (OH) framework developed by the Network for Evaluation of One Health (NEOH), showed strengths in planning and collaboration but struggled with data sharing and resource allocation (Hanin et al., 2018). Social factors such as public awareness, governance, and cultural norms also affect the adoption to One Health strategies (Brandão et al., 2021; Acharya et al., 2019; Anthony and De Paula Vieira, 2022).

3 Methodology

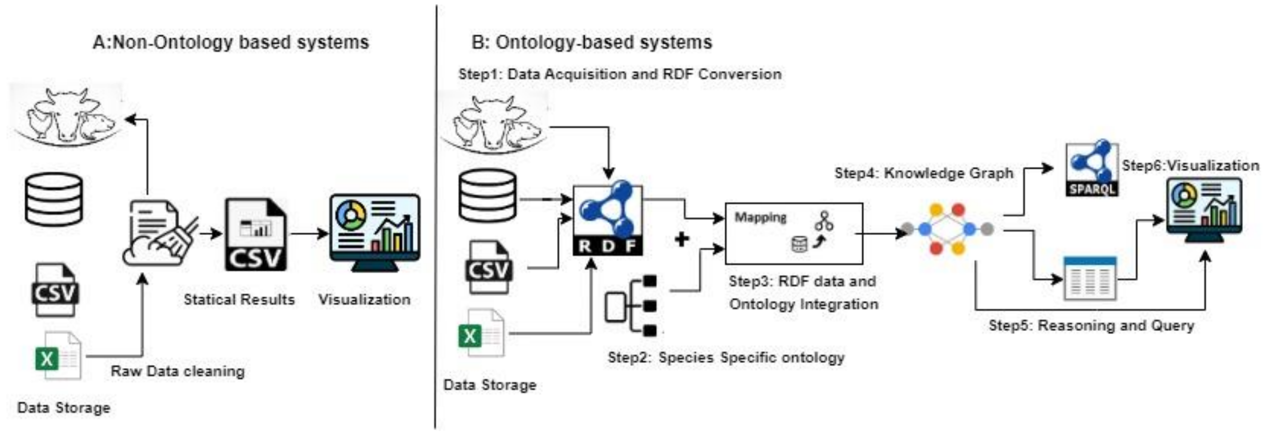


Figure 3: (A) Non-Ontology based vs. (B) Ontology-Driven Knowledge-based framework of Farm Animal Data Management (ODKFADM).

Figure 3 illustrates the data flow in a generic Non-Ontology-Based System compared to the Ontology-Driven Knowledge-Based Framework for Farm Animal Data Management (ODKFADM).

Traditional, non-ontology-based systems typically rely on relational databases or ad hoc data processing pipelines. These systems require repetitive and time-consuming work each time a new dataset is introduced, due to variations in data formats, structures, and naming conventions. Data cleaning and integration steps must be adapted manually, which limits the scalability, reusability, and semantic consistency of the system. Furthermore, such systems lack reasoning capabilities that would enable the inference of new knowledge from existing relationships in the data (Ren et al., 2020).

However, ODKFADM uses ontologies and RDF concepts to structure heterogeneous data (Figure 3). It consists of six core steps: data acquisition and RDF conversion, development of the species-specific Livestock Health Ontology (LHO), ontology mapping and integration, knowledge graph creation, SPARQL querying, and data visualization.

3.1 Step 1:data acquisition and RDF conversions

In the first step, we collected data from animal health laboratories in different European countries. These datasets came in various formats, such as CSV, Excel, and PDF. To prepare this data for analysis, we used tools like pandas, PySpark, and R to read and process it. We then converted the cleaned data into RDF (Resource Description Framework) format, which is a standard way to structure data so that it can be easily shared and linked. For this conversion, we used the RDFLib library in Python and RDF packages in R.

3.2 Step 2: Species-specific livestock health ontology

Ontology provides a formal and explicit description of a data model within an application domain, where relationships between individuals facilitate the reuse and sharing of knowledge. Livestock Health Ontology (LHO) is an extension of the DECIDE core upper-level ontology, developed following OBO Foundry guidelines to support semantic integration and reasoning across livestock species. It captures structured information such as farm identifiers, geolocation, breed, diagnostic tests, and pathogen results, represented through defined classes, subclasses, and relationships. LHO initially focused on cattle but has extended to include pigs, poultry, and salmon. The Livestock Health Ontology (LHO) is designed to integrate and represent both pathogenic diagnostic data (e.g., test results, pathogens, farm identifiers, geolocation, breed, sample metadata) and production performance data (e.g., mortality, growth rate, biomass, hatch date, thinned) across major livestock species: cattle, pigs, poultry, and salmon. Covered pathogens include MS, PM, MB, BCV, BRSV, PRRS, Swine Influenza (H1N1, H1N2, H2N3), MH, ISAV, and IB strains. It supports farm-level and sample-level modelling to improve health surveillance and decision support in animal farming systems.

3.2.1 Formal representation of the LHO ontology

The LHO model can be formally represented as:

$\text{LHO} = (C, P, A, I)$

Where LHO stands for the Livestock Health Ontology model. Here, C represents the concepts and instances conceptualisation, P denotes the properties and their relationships, A represents the rules or axioms and I represent the model interpretation. In the LHO, object properties (P) and relationships (Re) can be formally represented by the following symbols:

$P = (D\text{: Data type; } O\text{: Object; } T\text{: Transitive; } F\text{: Functional})$

$Re = (\equiv\text{: equivalence; } \subset\text{: subsumed; } \cap\text{: Disjoint})$

The domain of a property is formally represented as follows:

$[\forall b; b: p\ (a; b) \rightarrow \text{Domain}\ (a)];\ p \in P;\ \text{Domain} \in C$

The range of a property is formally represented as follows:

$[\forall b; b: p\ (a; b) \rightarrow \text{Range}\ (b)];\ p \in P;\ \text{Range} \in C$

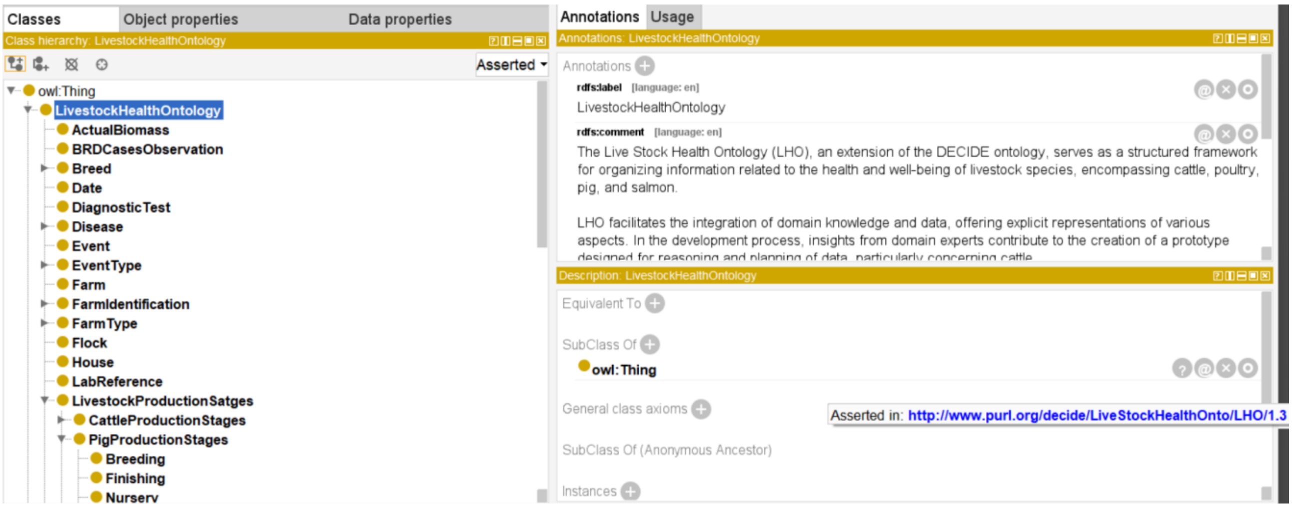
The interpretation of the ontology model, represented as LHO, indicates that classes and properties have various many relationships (Re), such as subsuming, equivalence, and disjoint relationships. The symbol ≡ denotes the equivalence of classes; the symbol ∩ indicates conceptually disjoint classes and is symmetric, reflexive, and transitive. Examples of these relationships in the LHO ontology model include:

$\text{CattleSample} \cap \text{PoultrySample} \cap \text{PigSample} \cap \text{SalmonSample} \subset \text{LiveStockHealthOntology}$

In the LHO model, all (⊂) subsume relations are asymmetric, irreflexive, and transitive. However, ∩ is conceptually disjoint and maintains its properties of being transitive (relationships can be inherited), symmetric (relationships are bidirectional), and reflexive (entities can have a relation with themselves).

3.2.1.1 Define classes and subclasses of LHO

In LHO, LiveStockHealthOntology is represented as the top-level concept. SampleType, SampleNumber, FarmIdentification, Breed, Province, Country, DiagnosticTest, Pathogen, etc are the subclasses of LiveStockHealthOntology class as shown in Figure 4.



rdfs:label [language: en]
LivestockHealthOntology

rdfs:comment [language: en]
The Live Stock Health Ontology (LHO), an extension of the DECIDE ontology, serves as a structured framework for organizing information related to the health and well-being of livestock species, encompassing cattle, poultry, pig, and salmon.

LHO facilitates the integration of domain knowledge and data, offering explicit representations of various aspects. In the development process, insights from domain experts contribute to the creation of a prototype designed for reasoning and planning of data, particularly concerning cattle

Asserted in: http://www.purl.org/decide/LiveStockHealthOnto/LHO/1.3

Figure 4. Define classes and subclasses of LHO.

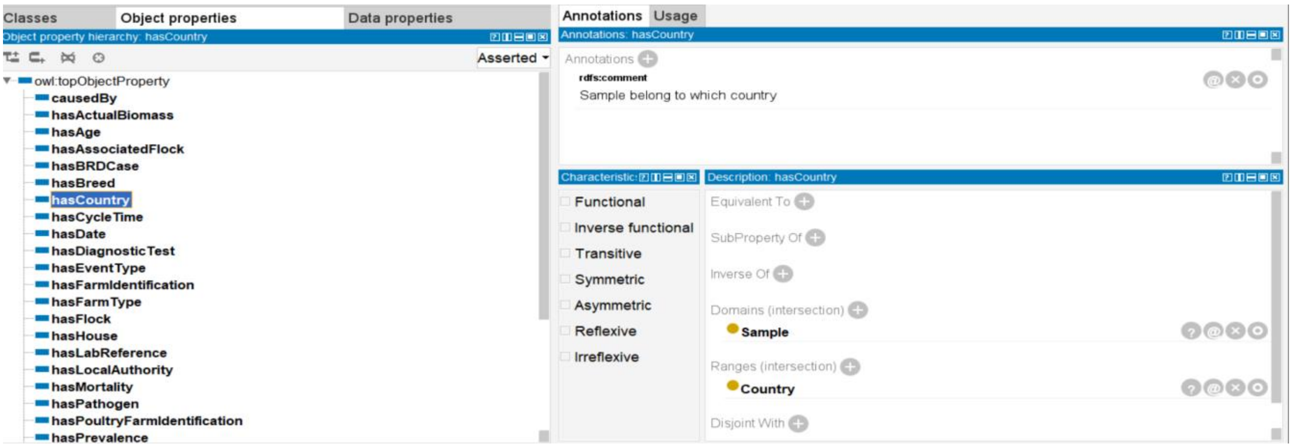
3.2.1.2 Define object properties of LHO

In the ontology, object properties are used to define semantic relationships between classes. Figure 5 shows the relationships of object properties such as hasBreed, hasCountry, hasPathoge, hasSampleResults, etc. Each property includes a domain (the class it applies to) and a range (the class it points to). For example: Object Property: hasCountry

Domain: Sample

Range: Country

Description: Indicates which country a given sample originates from.



rdfs:comment
Sample belong to which country

☐ Functional
☐ Inverse functional
☐ Transitive
☐ Symmetric
☐ Asymmetric
☐ Reflexive
☐ Irreflexive

Domains (intersection): Sample
Ranges (intersection): Country

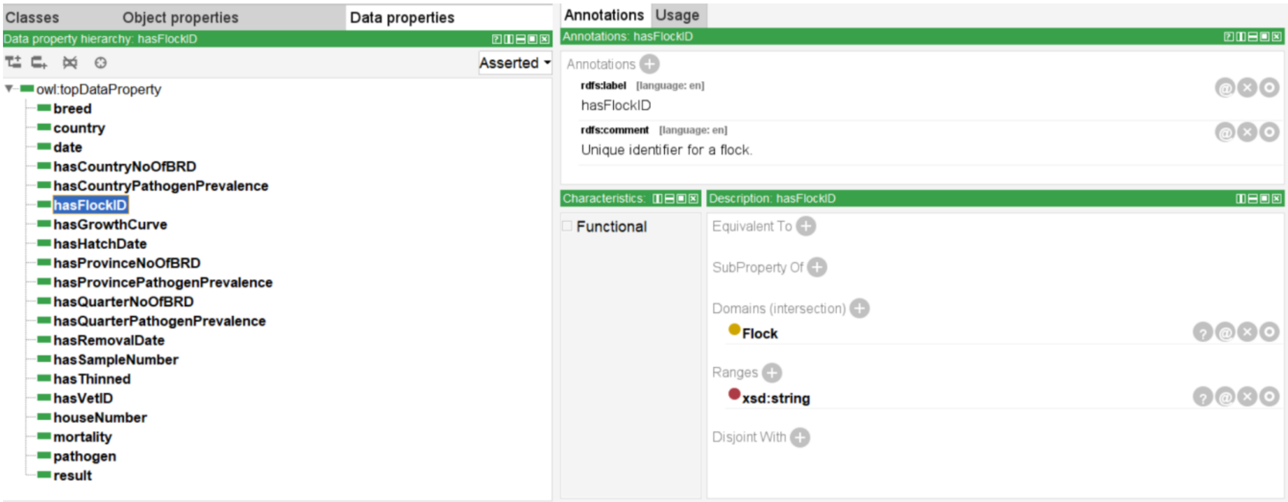
Figure 5. Define the object properties of LHO.

3.2.1.3 Define data properties of LHO

In the ontology, data properties are used to link individuals (instances of classes) to literal values (such as strings, numbers, or dates). Each data property is defined with a domain (the class it applies to) and a range (the type of data it links to). For example, Figure 6 shows:

Data Property: hasFlockID
Domain: Flock
Range: xsd:string
Description: Represents a unique string identifier for a flock.

This means that hasFlockID applies to individuals of the class Flock and assigns them a string value as an identifier.



rdfs:label [language: en]
hasFlockID

rdfs:comment [language: en]
Unique identifier for a flock.

☐ Functional

Domains (intersection): Flock

Ranges: xsd:string

Figure 6. Define data properties of LHO.

3.2.1.4 Define LHO axioms

LHO axioms represent the relationship between classes, instances, and attributes. For example, Figure 7 shows class axioms describing the semantic structure and meaning such as owl: disjoint class, rdfs: Subclassof, rdf: Description, and rdf: type. After defining these axioms, we can provide the disjoint and equivalence relationships among classes.

```
    <!--
    ///////////////////////////////////////////////////////////////////////////////////////
    //
    // General axioms
    //
    ///////////////////////////////////////////////////////////////////////////////////////
     -->

    <rdf:Description>
        <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AllDisjointClasses"/>
        <owl:members rdf:parseType="Collection">
            <rdf:Description rdf:about="https://www.purl.org/decide/LiveStockHealthOnto/LHO#Finishing"/>
            <rdf:Description rdf:about="https://www.purl.org/decide/LiveStockHealthOnto/LHO#Nursery"/>
            <rdf:Description rdf:about="https://www.purl.org/decide/LiveStockHealthOnto/LHO#Suckling"/>
            <rdf:Description rdf:about="https://www.purl.org/decide/LiveStockHealthOnto/LHO#Weaning"/>
        </owl:members>
    </rdf:Description>
</rdf:RDF>
```

Figure 7: Define axioms of LHO.

3.2.1.5 Graphical representation of ontology

Figure 8 shows the graphical representation of ontology in Onto Graph representing classes, subclasses, and their relationships.

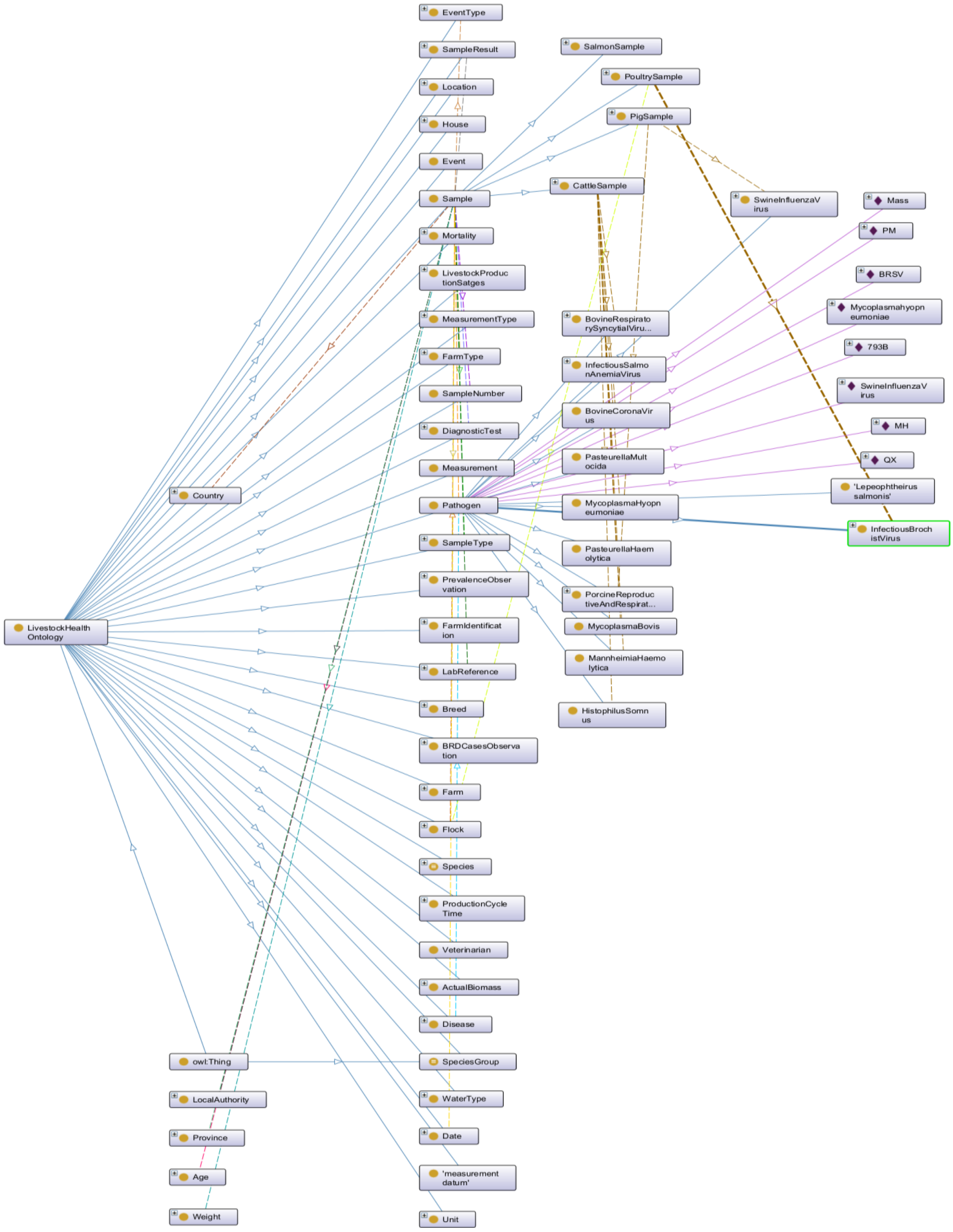


Figure 8. Graphical representation of LHO in protege.

3.3 Step 3: RDF data and ontology integration (mapping)

The algorithm in Table 4, automates the process of transforming tabular health data (CSV, Excel or pdf) into RDF triples and aligns the data with Livestock Health Ontology (LHO). This step is species-agnostic; it works for cattle, poultry, pig, and salmon datasets because it follows the same structure but maps species-specific attributes and instances into the shared ontology framework.

Table 4: RDF data and ontology integration (Mapping).

Algorithm: Create RDF graph from dataset and aligned with ontology

```
Input:
OntologyFilePath = "path/to/your/ontology.rdf"
df = pd.read_csv("path/to/your/dataset.csv")   # Data Frame containing the dataset
Output:
RDFoutputFile = " path/to/your/RDFoutput.rdf"
Procedure:

1. Initialize an RDF graph:
g = rdflib.Graph()
1.   Define namespaces for RDF entities:
Namespace = Namespace("http://www.example.com/ontology#")
2.   Load the ontology into the RDF graph.
g.parse(OntologyFilePath, format="xml"
4.   Iterate through each row in the dataset:
for row_index, row in df.iterrows():
   a. Create a unique URI for the sample based on the row index.
   sample_uri = URIRef("http://www.example.com/ontology#Sample_" + str(row_index))
   b. Add assertions for the sample, specifying its type and description.
   g.add((sample_uri, RDF.type, ontology.SampleClass))
   description = "An individual representing a sample."
   g.add((sample_uri, RDFS.comment, Literal(description, lang="en")))
   # Similar steps for other columns
   c. For each attribute in the dataset:
   for attribute in df.columns:
      i. Create a unique URI for the attribute/Object property value.
      attribute_value_uri = URIRef("http://www.example.com/ontology#" + attribute + "_" + str(row[attribute]))

      ii. Add assertions for the sample having the attribute with the corresponding value.
      g.add((sample_uri, namespace.hasAttribute, attribute_value_uri))
      g.add((attribute_value_uri, RDF.type, namespace.AttributeValueClass))
      description = f"An individual representing the value {row[attribute]} for the attribute {attribute}."
      g.add((attribute_value_uri, RDFS.comment, Literal(description, lang="en")))
5. Serialize the RDF graph to an output RDF file:
g.serialize(rdf_output_file, format="xml")
6. Output the path to the generated RDF file:
RDFoutputFile = " path/to/your/RDFoutput.rdf"
End Algorithm
```

3.4 Step 4: Knowledge graph (ontology update)

The knowledge graph is a well-unified, structured, and integrated knowledge repository. When the RDF data is integrated with the Ontology, it may require updating the existing ontology, consequently contributing to the development of a knowledge graph. Such updating could be done by adding new classes, relationships, and properties for a better understanding of integrated data. This would result in a knowledge graph that enhances more advanced reasoning and querying capabilities (Hogan et al., 2021).

3.5 Step 5: Reasoning and query

To explore the knowledge graph generated from RDF data and the Livestock Health Ontology (LHO), we used reasoning and SPARQL queries (Kollia et al., 2011). Reasoning allows us to infer implicit knowledge from explicitly defined classes, relationships, and axioms in ontology. Each use case (cattle, poultry, pig, salmon) reuses the same ontology-driven framework but tailors the queries to its specific data attributes, pathogens, and health indicators. In section 4, we provide detailed queries for each species to illustrate how ontology alignment enables species-level querying and analysis.

3.6 Step 6: Visualisation and analysis

To analyse and interpret the knowledge graph across species, we selected Tableau as the preferred visualization tool due to its compatibility with RDF-based data sources and its interactive capabilities setup supports SPARQL-based visual exploration of integrated livestock health data. Each use case, cattle, poultry, pig, and salmon, has its own tailored visual dashboard (referred to as barometers), highlighting species-specific disease trends, pathogen distribution, and key indicators such as location, breed, or age. These dashboards are designed to support decision-making for veterinarians, researchers, and industry stakeholders, allowing them to monitor, compare, and respond to infectious disease patterns and mortality rates more effectively.

The ODKFADM framework follows a structured multi-step approach designed to support semantic data integration, reasoning, and visualization across multiple livestock species. To support transparency and reuse, the full implementation, including R, PySpark, and pandas notebook code, is publicly available on GitHub. This enables researchers, data scientists, and animal health stakeholders to apply and adapt the framework to their own datasets and use cases. By openly sharing these resources, the project aims to foster collaboration and knowledge exchange, ultimately contributing to improved disease surveillance and livestock health management. The code is available at the GitHub repository decide-project-eu/LiveStockHealthOntologyWp1, and the ontology can be accessed on BioPortal and AgroPortal.

4 Case studies as proof of concept

4.1 Cattle case study

In this case study, we analyse and visualize pathogen prevalence across the cattle production system using diagnostic sample data collected from multiple animal health laboratories across Europe.

4.1.1 Data set and test bed

We evaluate the ODKFADM framework using diagnostic data focused on cattle health. The datasets span from 2016 to 2024 and include information on farms, geolocation, breed type (beef, dairy, or mixed), diagnostic tests, and confirmed results for bacterial pathogens such as HS, MH, PM, MB, as well as viral agents like BCV and BRSV. After aggregation at the herd level, a total of 11,838 records remained, comprising 8,678 from Belgium, 1,404 from the Netherlands, and 1,756 from Ireland.

4.1.2 LHO ontology

Classes added: Cattle, CattleSample, CattleProductionStages, Breed, DiagnosticTest, and Infectious Disease were defined or reused to accurately model the cattle health domain and support semantic integration of the data. Introduce subclasses HistophilusSomnus, MannheimiaHaemolytica, MycoplasmaBovis, PasteurellaMultocida, PasteurellaHaemolytica and viruses such as BovineCoronaVirus, and BovineRespiratorySyncytialVirus under "Pathogen".

Attributes defined: "hasProvince", "hasCountry", "hasBreed", "hasSampleType", "hasPathogen", "hasSampleNumber", "hasDate", "hasResult".

4.1.3 Integrate cattle health datasets

Cattle health data was integrated into the ontology by aligning it with the LHO structure using the RDF mapping algorithm described in the ODKFADM framework. This process, demonstrated in Table 4, converts tabular cattle data into RDF format and maps it to the appropriate ontology classes and properties for seamless semantic integration.

4.1.4 Formulated queries

4.1.4.1 SPARQL Query 1 for filtering positive cattle samples for MB pathogen via Swab tests

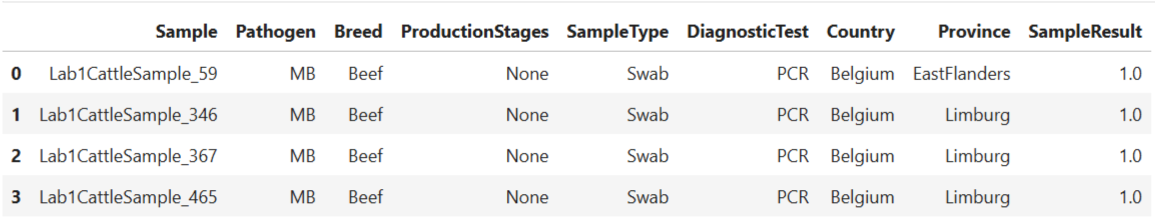
Table 5 shows the SPARQL query retrieves detailed information about cattle health samples from an RDF dataset using the Livestock Health Ontology (LHO). It specifically selects samples classified as CattleSample that are associated with the pathogen MB, have a sample type of Swab, and a test result of 1.0 (indicating a positive outcome). For each matching sample, it extracts attributes such as the sample name, pathogen, breed, sample type, diagnostic test, country, and result as shown in Figure 9.

Table 5: Cattle query filtering positive cattle samples for MB pathogen via Swab tests.

```
PREFIX decide: <http://www.purl.org/decide#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX LHO: <http://www.purl.org/decide/LiveStockHealthOnto/LHO#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT
  (strafter(str(?Sample), "#") AS ?SampleName)
  (strafter(str(?Pathogen), "#") AS ?PathogenName)
  (strafter(str(?Breed), "#") AS ?BreedName)
  (strafter(str(?LivestockProductionStages), "#") AS ?StagesName)
  (strafter(str(?SampleType), "#") AS ?SampleTypeName)
  (strafter(str(?DiagnosticTest), "#") AS ?DiagnosticTestName)
  (strafter(str(?Country), "#") AS ?CountryName)
```

```
 (strafter(str(?Province), "#") AS ?provinceName)
  (strafter(str(?SampleResult), "#") AS ?SampleResultName)

WHERE {{
    ?Sample rdf:type LHO:CattleSample .
    ?Sample LHO:hasPathogen ?Pathogen .
    FILTER (?Pathogen = LHO:MB)
    ?Sample LHO:hasSampleType ?SampleType .
    FILTER (?SampleType = LHO:Swab)
    ?Sample LHO:hasCountry ?Country .
    ?Sample LHO:hasProvince ?Province .
    ?Sample LHO:hasBreed ?Breed .
    ?Sample LHO:hasResult ?SampleResult .
    FILTER (?SampleResult = LHO:1.0)
    ?Sample LHO:hasDiagnosticTest ?DiagnosticTest .
  }}
```



	Sample	Pathogen	Breed	ProductionStages	SampleType	DiagnosticTest	Country	Province	SampleResult
0	Lab1CattleSample_59	MB	Beef	None	Swab	PCR	Belgium	EastFlanders	1.0
1	Lab1CattleSample_346	MB	Beef	None	Swab	PCR	Belgium	Limburg	1.0
2	Lab1CattleSample_367	MB	Beef	None	Swab	PCR	Belgium	Limburg	1.0
3	Lab1CattleSample_465	MB	Beef	None	Swab	PCR	Belgium	Limburg	1.0

Figure 9: Cattle query filtering positive cattle samples for MB pathogen via Swab tests.

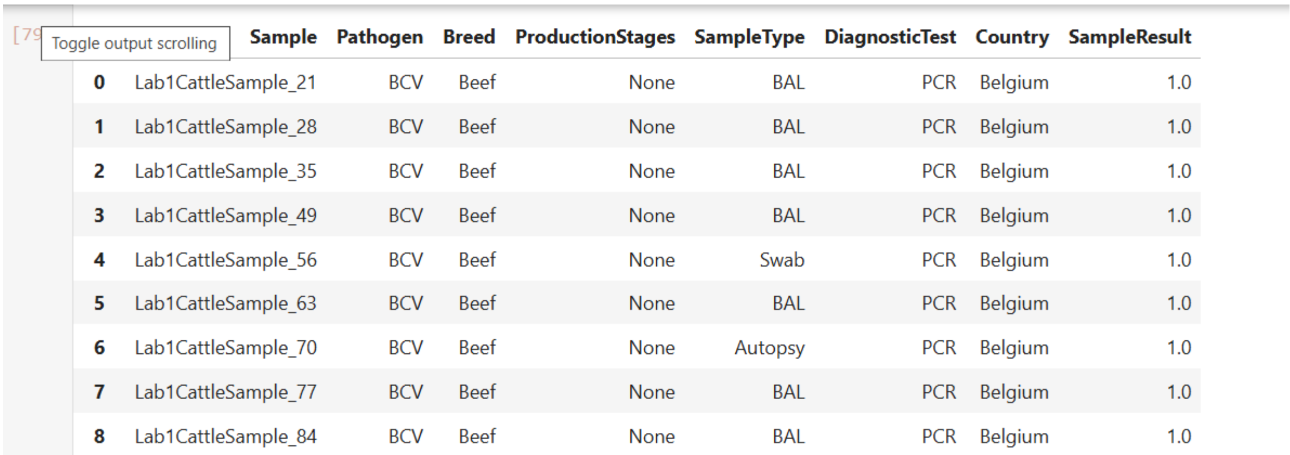
4.1.4.2 Query 2: filtering the results based on the breed being "Beef" associated with cattle sample

Breed has three types, beef, dairy, or mixed. The SPARQL query 2 in Table 6 retrieves the results based on Breed begin beef for pathogen BCV as shown in Figure 10.

Table 6: Cattle query for breed being «Beef».

```
PREFIX decide: <http://www.purl.org/decide#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX LHO: <http://www.purl.org/decide/LiveStockHealthOnto/LHO#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT
  (strafter(str(?Sample), "#") AS ?SampleName)
  (strafter(str(?Pathogen), "#") AS ?PathogenName)
  (strafter(str(?Breed), "#") AS ?BreedName)
  (strafter(str(?LivestockProductionStages), "#") AS ?StagesName)
  (strafter(str(?SampleType), "#") AS ?SampleTypeName)
  (strafter(str(?DiagnosticTest), "#") AS ?DiagnosticTestName)
  (strafter(str(?Country), "#") AS ?CountryName)
  (strafter(str(?SampleResult), "#") AS ?SampleResultName)
WHERE {
  {
    ?Sample rdf:type LHO:CattleSample .
    ?Sample LHO:hasPathogen ?Pathogen .
    FILTER (?Pathogen = LHO:BCV)
```

```
?Sample LHO:hasSampleType ?SampleType .
?Sample LHO:hasCountry ?Country .
?Sample LHO:hasBreed ?Breed .
FILTER (?Breed = LHO:Beef)
?Sample LHO:hasResult ?SampleResult .
?Sample LHO:hasDiagnosticTest ?DiagnosticTest .
FILTER (?DiagnosticTest = LHO:PCR)}}
```



	Sample	Pathogen	Breed	ProductionStages	SampleType	DiagnosticTest	Country	SampleResult
0	Lab1CattleSample_21	BCV	Beef	None	BAL	PCR	Belgium	1.0
1	Lab1CattleSample_28	BCV	Beef	None	BAL	PCR	Belgium	1.0
2	Lab1CattleSample_35	BCV	Beef	None	BAL	PCR	Belgium	1.0
3	Lab1CattleSample_49	BCV	Beef	None	BAL	PCR	Belgium	1.0
4	Lab1CattleSample_56	BCV	Beef	None	Swab	PCR	Belgium	1.0
5	Lab1CattleSample_63	BCV	Beef	None	BAL	PCR	Belgium	1.0
6	Lab1CattleSample_70	BCV	Beef	None	Autopsy	PCR	Belgium	1.0
7	Lab1CattleSample_77	BCV	Beef	None	BAL	PCR	Belgium	1.0
8	Lab1CattleSample_84	BCV	Beef	None	BAL	PCR	Belgium	1.0

Figure 10: Query 2, filtering the results based on the breed being "Beef" associated with cattle sample.

4.1.5 Step 6: Visualisation and analysis

The SPARQL query results are processed and then integrated into Tableau for visualization. This visualisation, known as the cattle barometer, is a powerful visualisation tool designed to provide meaningful insights and understanding from data related to respiratory infectious diseases in cattle. This tool is intended for use by veterinary professionals, researchers, and livestock industry stakeholders, such as farmers and producers, who are keen on monitoring and mitigating the impact of diseases on the cattle population, as shown in Figure 11.

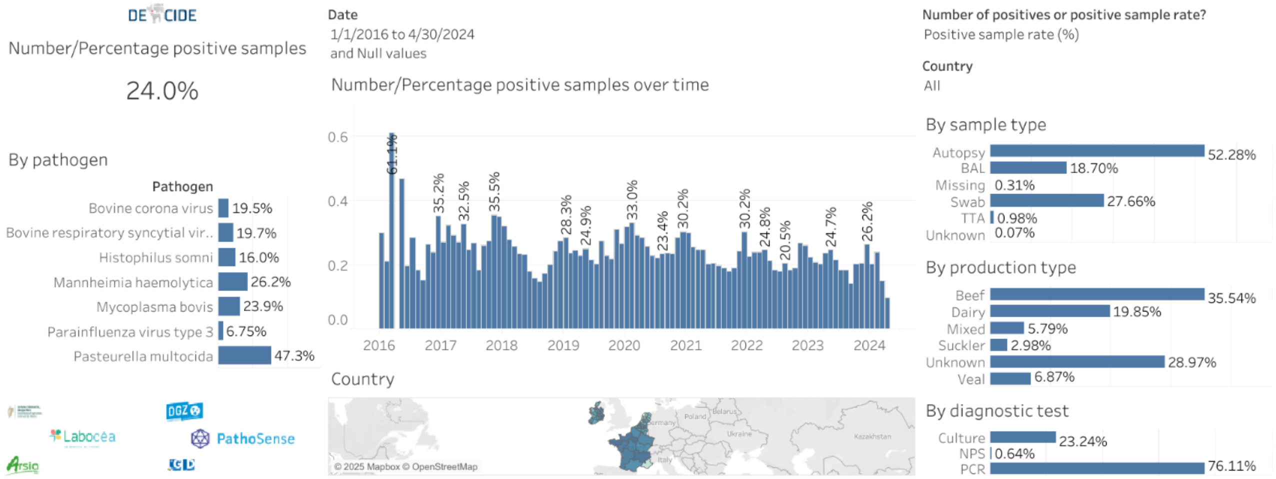


Figure 11: Cattle barometer use case, focusing on respiratory infectious diseases.

To learn more about case studies related to the transformation of the cattle barometer for multiple labs across Europe in the DECIDE project, you can visit the website link: https://decide-project-eu.github.io/case-studies-website/case-studies/cattle-barometer.html. This tool provides a comprehensive view of the historical trends, dynamics, and current presence of specific pathogens in different regions. Its capabilities extend to offering early warnings and supporting decision-making on farms to control infectious diseases. Users can leverage the insights gained from the barometer to make informed decisions related to antimicrobial use, vaccination strategies, and the acquisition of animals.

4.2 Poultry case study

This use case builds on the existing cattle use case methodology and LHO ontology, requiring only minor adjustments to the classes and attributes.

4.2.1 Data set and test bed

The poultry use case includes two main datasets. The first consists of sample-level diagnostic data collected from broiler chickens. Each record contains key details such as the detected pathogen strain (e.g., 491, 793B, Mass, D181, D1486, D274, D2860, QX), breed type (meat-producing or egg-laying), production stage (broiler), and sample type. It also includes the country of origin and a diagnostic result, where a value of 1 indicates the presence of a pathogen. The second dataset provides flock-level production data that complements the diagnostic results. It includes information like the flock ID, the veterinarian responsible, hatch date, and removal date, marking the start and end of the flock's life cycle. The growth curve shows the expected growth pattern for managing the flock. The "thinned" value tells whether the flock was partially removed before final slaughter: 0 means no thinning occurred, while 1 means thinning occurred.

4.2.2 Extend LHO ontology

Classes added: Poultry, PoultrySample, PoultryProductionStages, Flock, House, and subclasses have been introduced under the "Pathogen" class to represent Infectious Bronchitis (IB). Individual instances have been added for specific pathogen strains, including 491, 793B, Maas, D181, D1486, D274, D2860, and QX. For class PoultryProductionStages, the defined stages include Breeder, Broiler, Hatching, Layer, and Processing.

Attributes defined: "hasAssociatedFlock," "hasHouse," "hasFlock ," and reuse hasCountry, hasBreed, hasSampleType, hasResult

4.2.3 Integrate poultry health datasets

Poultry health data was integrated into the ontology by aligning it with the LHO structure using the RDF mapping algorithm described in the ODKFADM framework. This process, demonstrated in Table 4, converts tabular cattle data into RDF format and maps it to the appropriate ontology classes and properties for seamless semantic integration.

4.2.4 Formulated queries

4.2.4.1 SPARQL Query 1 for sample-level diagnostic data

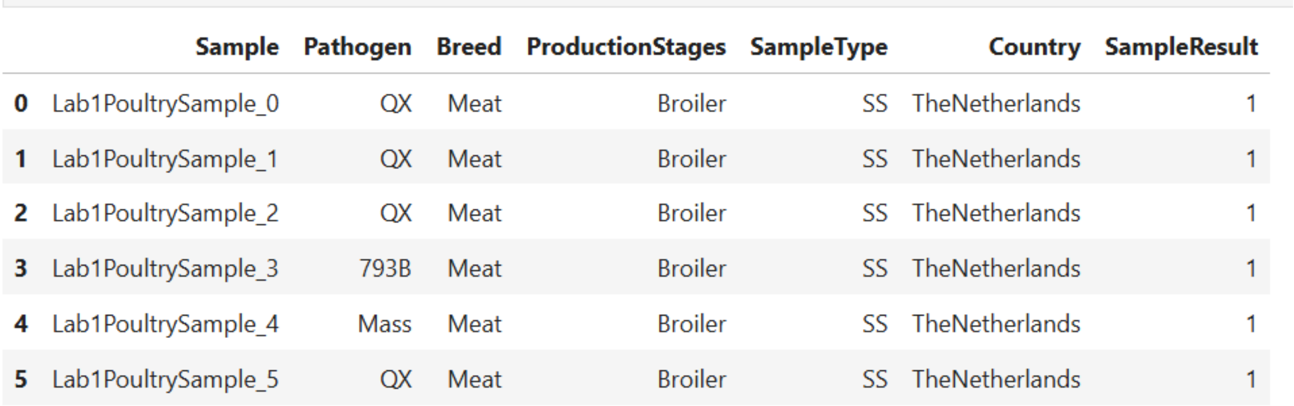
Table 7, the SPARQL SELECT query retrieves specific information about poultry samples, pathogens, sample types, and provinces; the aliases (?SampleName, ?PathogenName, ?SampleTypeName, ?CountryName) indicate that these values are extracted from respective URIs, providing clear names for the poultry sample, pathogen, breed, production stage, sample type, country, and sample result as shown in Figure 12.

Table 7: Poultry sample-level diagnostic data extraction query.

```
PREFIX decide: <http://www.purl.org/decide#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX LHO: <https://www.purl.org/decide/LiveStockHealthOnto/LHO#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT
  (strafter(str(?Sample), "#") AS ?SampleName)
  (strafter(str(?Pathogen), "#") AS ?PathogenName)
  (strafter(str(?Breed), "#") AS ?BreedName)
```

```
  (strafter(str(?LivestockProductionStages), "#") AS ?StagesName)
  (strafter(str(?SampleType), "#") AS ?SampleTypeName)
  (strafter(str(?Country), "#") AS ?CountryName)
  (strafter(str(?SampleResult), "#") AS ?SampleResultName)

WHERE {{
   ?Sample rdf:type LHO:PoultrySample .?Sample LHO:hasPathogen ?Pathogen .?Sample
LHO:hasSampleType ?SampleType .
   ?Sample LHO:hasCountry ?Country .
   ?Sample LHO:hasProductionStages ?LivestockProductionStages .
   ?Sample LHO:hasBreed ?Breed .
   ?Sample LHO:hasResult ?SampleResult .
  }}
```



	Sample	Pathogen	Breed	ProductionStages	SampleType	Country	SampleResult
0	Lab1PoultrySample_0	QX	Meat	Broiler	SS	TheNetherlands	1
1	Lab1PoultrySample_1	QX	Meat	Broiler	SS	TheNetherlands	1
2	Lab1PoultrySample_2	QX	Meat	Broiler	SS	TheNetherlands	1
3	Lab1PoultrySample_3	793B	Meat	Broiler	SS	TheNetherlands	1
4	Lab1PoultrySample_4	Mass	Meat	Broiler	SS	TheNetherlands	1
5	Lab1PoultrySample_5	QX	Meat	Broiler	SS	TheNetherlands	1

Figure 12: Poultry sample-level diagnostic data extraction query result.

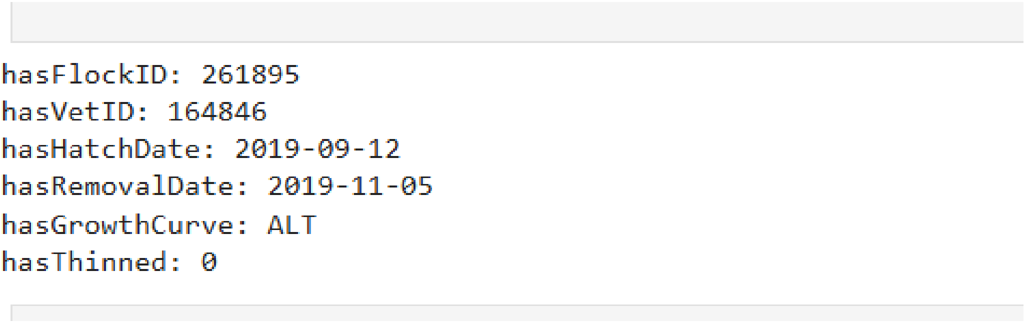
4.2.4.2 SPARQL Query 2 for flock-level production data

This SPARQL query in Table 8 retrieves all literal data properties and their values for a specific flock instance (lho:FlockID_261895) in the ontology. The properties shown in the query result (e.g., hasHatchDate, hasRemovalDate, hasGrowthCurve, hasThinned) describe various attributes of the flock, such as its hatch date, removal date, veterinary contact, growth pattern, and whether any birds were removed (thinned) during the production cycle (Figure 13).

Table 8: flock-level production data extraction query.

```
PREFIX lho: <http://www.purl.org/decide/LiveStockHealthOnto/LHO#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
SELECT ?property ?label ?value
WHERE {
   lho:FlockID_261895 ?property ?value .
   FILTER (isLiteral(?value)) .

   OPTIONAL {
      ?property rdfs:label ?label .
      FILTER (lang(?label) = "" || lang(?label) = "en")
   }
}
"""
```



```
hasFlockID: 261895
hasVetID: 164846
hasHatchDate: 2019-09-12
hasRemovalDate: 2019-11-05
hasGrowthCurve: ALT
hasThinned: 0
```

Figure 13: Poultry flock-level production data query result.

4.2.5 Visualisation and analysis

With the support of the Poultry case group in DECIDE, the SPARQL query results are processed and then integrated into Tableau for visualization, as shown in Figure 14. This visualisation, the poultry barometer, offers vital insights for veterinarians, researchers, and industry stakeholders. It captures pathogen trends, enabling early warnings, informed vaccination decisions, and effective disease control on farms. Explore the transformation details in the DECIDE project's Case Studies https://decide-project-eu.github.io/case-studies-website/case-studies/poultry-barometer.html.

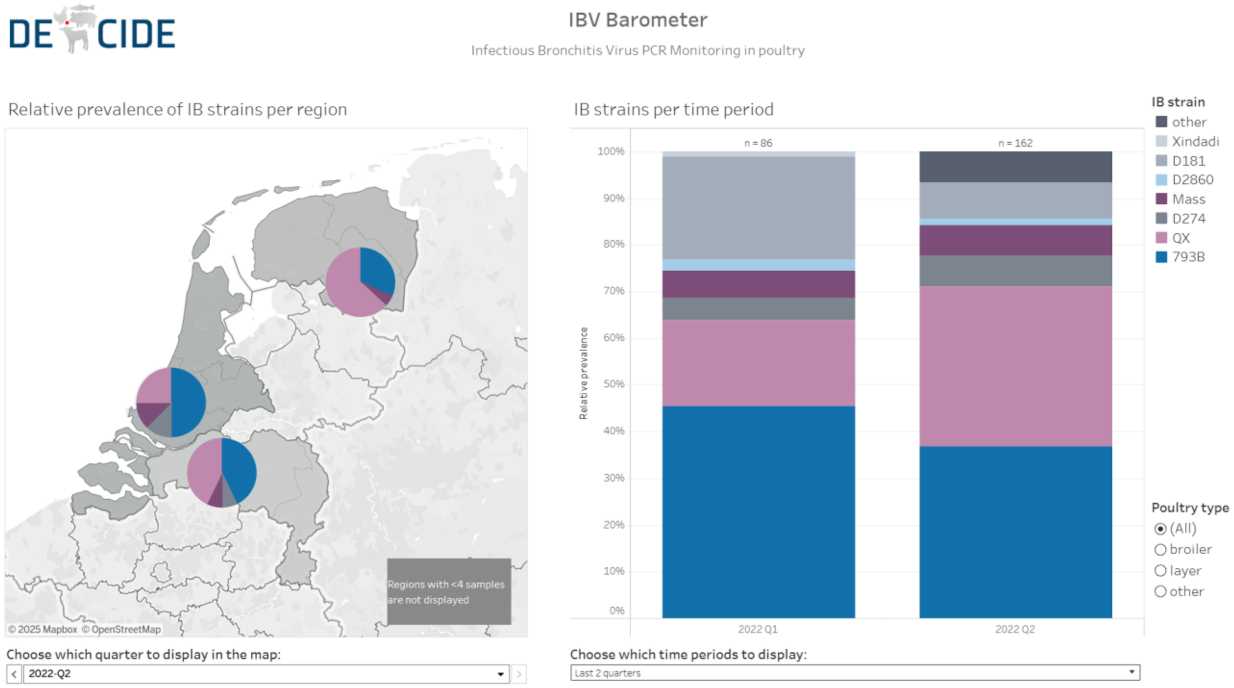


Figure 14. Poultry barometer , focusing on respiratory infectious diseases.

4.3 Pig case study

Similar to the poultry case, this use case builds on the existing cattle use case methodology and the LHO ontology, requiring only minor class and attribute adjustments. In this case study, we analyze and visualize pathogen prevalence across different pig production stages using diagnostic sample data.

4.3.1 Data set and test bed

The dataset originates from a veterinary diagnostic test bed focused on pathogen surveillance in pig populations, particularly within the Spanish swine industry. It comprises laboratory test records covering

22,013 cases from year 2020 to 2024, each uniquely identified by a sample number, and includes metadata such as sampling date, province, detected pathogen (e.g., swine influenza virus, Porcine Reproductive and Respiratory Syndrome, Mycoplasma hyopneumoniae, and Lawsonia intracellularis), diagnostic result (positive or negative), and sample type (e.g., tissue, fluid, blood, swab, feces, mixed). Supplementary fields like pig age, pig weight, and farm type provide crucial biological and management context. These attributes are used in combination to assign each sample to a corresponding pig production stage, following internationally accepted life cycle phases, breeding, suckling, weaning, nursery, and finishing.

The breeding stage refers to sexually mature pigs such as sows, boars, and gilts. This stage is identified based on farm type, reproductive status, or breed. This includes sows in gestation or farrowing, boars used for mating, and gilts—young females eligible to breed but not yet farrowed. The suckling stage includes piglets from birth to around 4 weeks of age, typically weighing 2 to 7 kg. During this phase, piglets are nursed by the sow and begin early developmental milestones such as teething and sensory development. The weaning stage includes pigs aged 4 to 6 weeks, transitioning from milk to solid feed. These pigs usually weigh 6 to 10 kg, begin socializing, and often receive their first health interventions. The nursery stage includes pigs aged 6 to 10 weeks, weighing roughly 7 to 25 kg, and housed under nursery-specific farm conditions. During this period, they grow rapidly, adapt to group housing, and undergo vaccinations or treatments. The finishing stage refers to pigs aged over 10 weeks and weighing more than 25 kg, reared until they reach market weight, typically over 100 kg, before slaughter. This stage-based classification supports structured epidemiological analysis and enables targeted disease monitoring across the full spectrum of swine development.

4.3.2 LHO ontology extension

The Livestock Health Ontology (LHO) serves as the foundation for pig use case. To meet the specific needs of pig health prediction, the ontology is extended with the following key classes and relationships:

Classes added: "Pig," "PigSample," "PigProductionstages" with specific subclasses for "Pathogen," such as "PorcineReproductiveAndRespiratorySyndrome (PRRS)," "SwineInfluenza" (further divided into subtypes like H1N1, H1N2, and H2N3), representing distinct strains of the influenza virus, "LawsoniaIntracellularis", and "MycoplasmaHyopneumoniae".

Attributes defined: "hasWeight," "hasAge," "hasProductionstages ," "hasFarmType," and reuse the attributes "hasProvince", "hasCountry", "hasBreed", "hasSampleType", "hasPathogen", "hasResult".

4.3.3 Integration of datasets

Pig data was integrated into the ontology by aligning it with the LHO structure using the RDF mapping algorithm described in the ODKFADM framework. This process, demonstrated in Table 4, converts tabular cattle data into RDF format and maps it to the appropriate ontology classes and properties for

4.3.4 Queries for insight extraction

4.3.4.1 Query 1 for positive pathogen results in pig tissue samples during the finishing stage

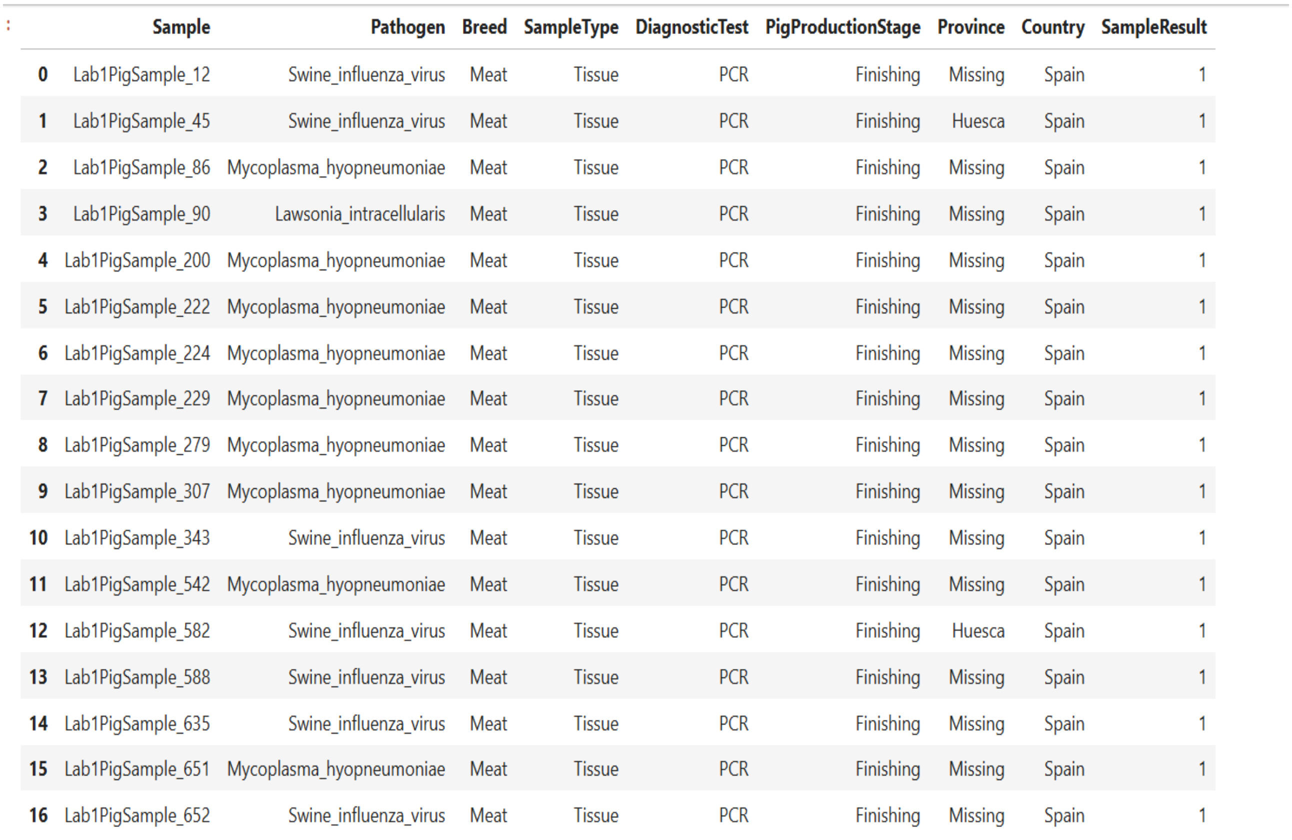
The query in Table 9 retrieves pig tissue samples that tested positive for a pathogen, taken during the finishing production stage. It filters results by sample type, production stage, and test result, returning related data like breed, country, province, and diagnostic test used. The results shown in Figure 15.

Table 9: For positive pathogen results in pig tissue samples during the finishing stage

```
PREFIX decide: <http://www.purl.org/decide#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX LHO: <http://www.purl.org/decide/LiveStockHealthOnto/LHO#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT
  (strafter(str(?Sample), "#") AS ?SampleName)
  (strafter(str(?Pathogen), "#") AS ?PathogenName)
```

```
  (strafter(str(?Breed), "#") AS ?BreedName)
  (strafter(str(?SampleType), "#") AS ?SampleTypeName)
  (strafter(str(?DiagnosticTest), "#") AS ?DiagnosticTestName)
  (strafter(str(?PigProductionStages), "#") AS ?ProductionStage)
  (strafter(str(?Province), "#") AS ?ProvinceName)
  (strafter(str(?Country), "#") AS ?CountryName)
  (strafter(str(?SampleResult), "#") AS ?SampleResultName)

 WHERE {{
  ?Sample rdf:type LHO:PigSample .
  ?Sample LHO:hasPathogen ?Pathogen .
  ?Sample LHO:hasBreed ?Breed .
  ?Sample LHO:hasSampleType ?SampleType .
  FILTER (?SampleType = LHO:Tissue)
  ?Sample LHO:hasProvince ?Province .
  ?Sample LHO:hasCountry ?Country .
  ?Sample LHO:hasProductionStages ?PigProductionStages .
  FILTER (?PigProductionStages = LHO:Finishing)
  ?Sample LHO:hasResult ?SampleResult .
  FILTER (?SampleResult = LHO:1)
  ?Sample LHO:hasDiagnosticTest ?DiagnosticTest .
  }}
```



	Sample	Pathogen	Breed	SampleType	DiagnosticTest	PigProductionStage	Province	Country	SampleResult
0	Lab1PigSample_12	Swine_influenza_virus	Meat	Tissue	PCR	Finishing	Missing	Spain	1
1	Lab1PigSample_45	Swine_influenza_virus	Meat	Tissue	PCR	Finishing	Huesca	Spain	1
2	Lab1PigSample_86	Mycoplasma_hyopneumoniae	Meat	Tissue	PCR	Finishing	Missing	Spain	1
3	Lab1PigSample_90	Lawsonia_intracellularis	Meat	Tissue	PCR	Finishing	Missing	Spain	1
4	Lab1PigSample_200	Mycoplasma_hyopneumoniae	Meat	Tissue	PCR	Finishing	Missing	Spain	1
5	Lab1PigSample_222	Mycoplasma_hyopneumoniae	Meat	Tissue	PCR	Finishing	Missing	Spain	1
6	Lab1PigSample_224	Mycoplasma_hyopneumoniae	Meat	Tissue	PCR	Finishing	Missing	Spain	1
7	Lab1PigSample_229	Mycoplasma_hyopneumoniae	Meat	Tissue	PCR	Finishing	Missing	Spain	1
8	Lab1PigSample_279	Mycoplasma_hyopneumoniae	Meat	Tissue	PCR	Finishing	Missing	Spain	1
9	Lab1PigSample_307	Mycoplasma_hyopneumoniae	Meat	Tissue	PCR	Finishing	Missing	Spain	1
10	Lab1PigSample_343	Swine_influenza_virus	Meat	Tissue	PCR	Finishing	Missing	Spain	1
11	Lab1PigSample_542	Mycoplasma_hyopneumoniae	Meat	Tissue	PCR	Finishing	Missing	Spain	1
12	Lab1PigSample_582	Swine_influenza_virus	Meat	Tissue	PCR	Finishing	Huesca	Spain	1
13	Lab1PigSample_588	Swine_influenza_virus	Meat	Tissue	PCR	Finishing	Missing	Spain	1
14	Lab1PigSample_635	Swine_influenza_virus	Meat	Tissue	PCR	Finishing	Missing	Spain	1
15	Lab1PigSample_651	Mycoplasma_hyopneumoniae	Meat	Tissue	PCR	Finishing	Missing	Spain	1
16	Lab1PigSample_652	Swine_influenza_virus	Meat	Tissue	PCR	Finishing	Missing	Spain	1

Figure 15. Query result for positive pathogen results in pig tissue samples during the finishing stage.

4.3.4.2 Query 2 for positive pathogen detection in pig fluid samples during weaning stage

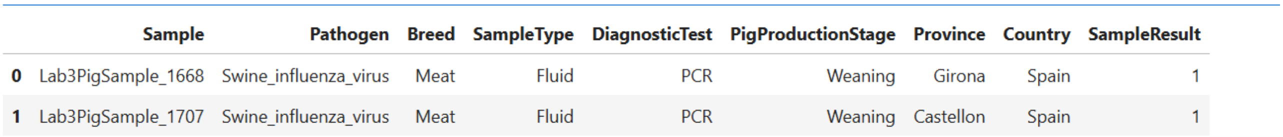
This SPARQL query in Table 9 retrieves pig fluid samples that tested positive for a pathogen during the weaning production stage. It returns information such as breed, pathogen detected, sample location (country and province), and the diagnostic test used. The results shown in Figure 16.

Table 9: For positive pathogen detection in pig fluid samples during the weaning stage.

```
PREFIX decide: <http://www.purl.org/decide#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX LHO: <http://www.purl.org/decide/LiveStockHealthOnto/LHO#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT
  (strafter(str(?Sample), "#") AS ?SampleName)
  (strafter(str(?Pathogen), "#") AS ?PathogenName)
  (strafter(str(?Breed), "#") AS ?BreedName)
  (strafter(str(?SampleType), "#") AS ?SampleTypeName)
  (strafter(str(?DiagnosticTest), "#") AS ?DiagnosticTestName)
  (strafter(str(?PigProductionStages), "#") AS ?ProductionStage)
  (strafter(str(?Province), "#") AS ?ProvinceName)
  (strafter(str(?Country), "#") AS ?CountryName)
  (strafter(str(?SampleResult), "#") AS ?SampleResultName)

WHERE {{
  ?Sample rdf:type LHO:PigSample .
  ?Sample LHO:hasPathogen ?Pathogen .
```

```
  ?Sample LHO:hasBreed ?Breed .
  ?Sample LHO:hasSampleType ?SampleType .
  FILTER (?SampleType = LHO:Fluid)
  ?Sample LHO:hasProvince ?Province .
  ?Sample LHO:hasCountry ?Country .
  ?Sample LHO:hasProductionStages ?PigProductionStages .
  FILTER (?PigProductionStages = LHO:Weaning)
  ?Sample LHO:hasResult ?SampleResult .
  FILTER (?SampleResult = LHO:1)
  ?Sample LHO:hasDiagnosticTest ?DiagnosticTest .
 }}
```



	Sample	Pathogen	Breed	SampleType	DiagnosticTest	PigProductionStage	Province	Country	SampleResult
0	Lab3PigSample_1668	Swine_influenza_virus	Meat	Fluid	PCR	Weaning	Girona	Spain	1
1	Lab3PigSample_1707	Swine_influenza_virus	Meat	Fluid	PCR	Weaning	Castellon	Spain	1

Figure 16: Query result for positive pathogen detection in pig fluid samples during weaning stage.

4.3.5 Visualisation and analysis

For the pig use case, we adopted the same visualisation approach used in the cattle case, using Tableau connected via ODBC to Virtuoso instances. With support from the pig group in DECIDE, the Pig Barometer was developed, as shown in Figure 17. An interactive dashboard that provides key insights into pathogen trends affecting pigs. More details on the data transformation and visualisation process can be explored at the pig barometer case study page.

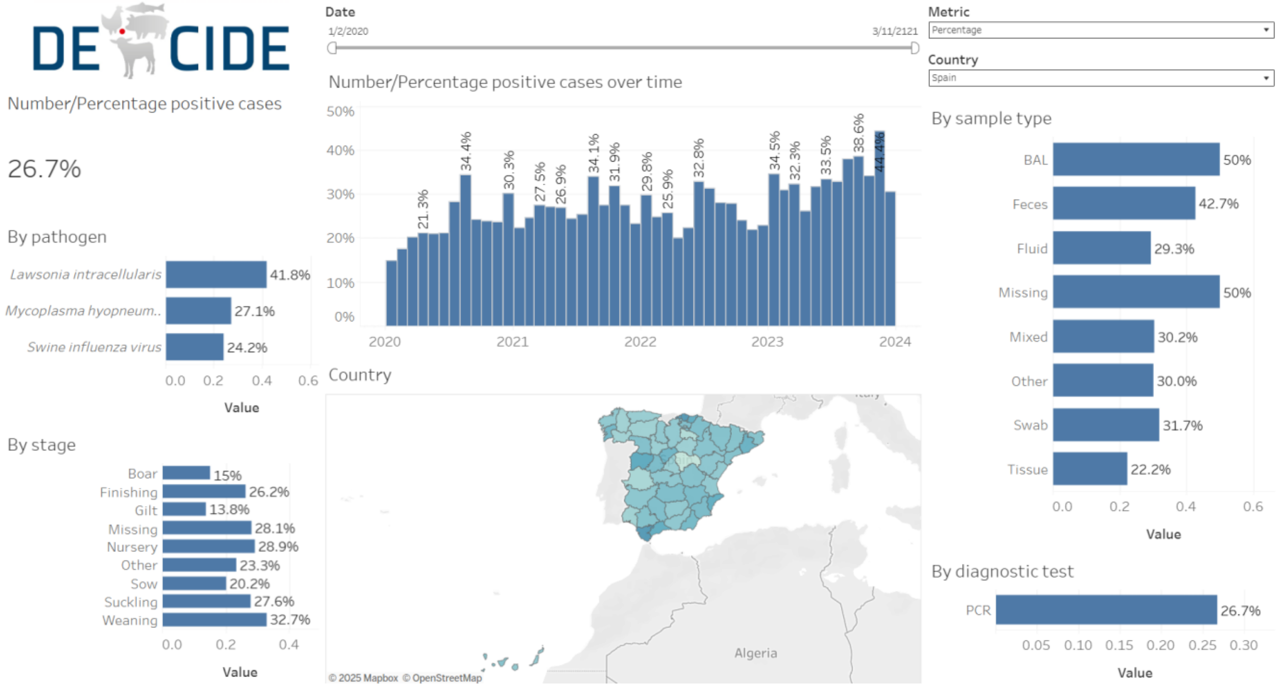


Figure 17: Pig barometer, focusing on porcine infectious diseases.

4.5 Cross-Species Query Use Case: Cattle, Pig, Poultry

For demonstration purposes, a Python notebook with a small dataset of labs has been uploaded to a dedicated GitHub repository under DECIDE at https://github.com/decide-project-eu/LivestockHealthOntologyIntegration-CattlePigPoultryUsecase-. This notebook includes a SPARQL query designed to retrieve diagnostic samples and associated metadata from cattle, pig, and poultry datasets. The query leverages the UNION operator to combine patterns across the three terrestrial livestock species, enabling unified access to diverse data sources while preserving species-specific distinctions. By aligning sample data through shared ontology terms such as pathogens, diagnostic tests, and geographical attributes, this cross-species query exemplifies the ontology's ability to support integrated health surveillance and data-driven decision-making across multiple farms and regions.

This use case has also been published as part of a conference paper, available via UGent Biblio and as an open-access resource on Zenodo, ensuring broad accessibility for both academic and practical use.

4.6 Salmon case study (aquatic livestock)

In this case study, we analyse and visualize mortality trends across the production cycle of farmed Atlantic salmon in Scotland using structured diagnostic and farm-level data.

4.6.1 Data set and test bed

The dataset spans from 2002 to 2024 and includes key parameters such as actual and allowed biomass, mortality in kilograms, water type (freshwater or seawater), feed input, and local authority location.

4.6.2 LHO ontology extension

The Livestock Health Ontology (LHO), developed as an extension of the broader DECIDE ontology, provides a semantic framework for standardizing and integrating animal health data across species. By aligning with the Livestock Health Ontology (LHO), this case study demonstrates the semantic modeling and integration of aquaculture-specific attributes such as hasMortality, hasActualBiomass, hasWaterType, and hasLocalAuthority. The data supports epidemiological analysis of fish health by mapping samples and measurements to environmental and production factors.

Classes added: Classes like "Salmon", "SalmonSample", "ActualBiomass", "WaterType", "LocalAuthority", and "Mortality" are added

Attributes defined: "hasWeight," "hasActualBiomass", "hasMortality", "hasLocalAuthority", "hasWaterType", and reuse the attributes "hasCountry", "hasSampleNumber", "hasDate".

4.6.3 Integration of datasets

Salmon health data is converted into RDF format and mapped to LHO, enabling integration into semantic knowledge graphs.

4.6.4 Formulated Query

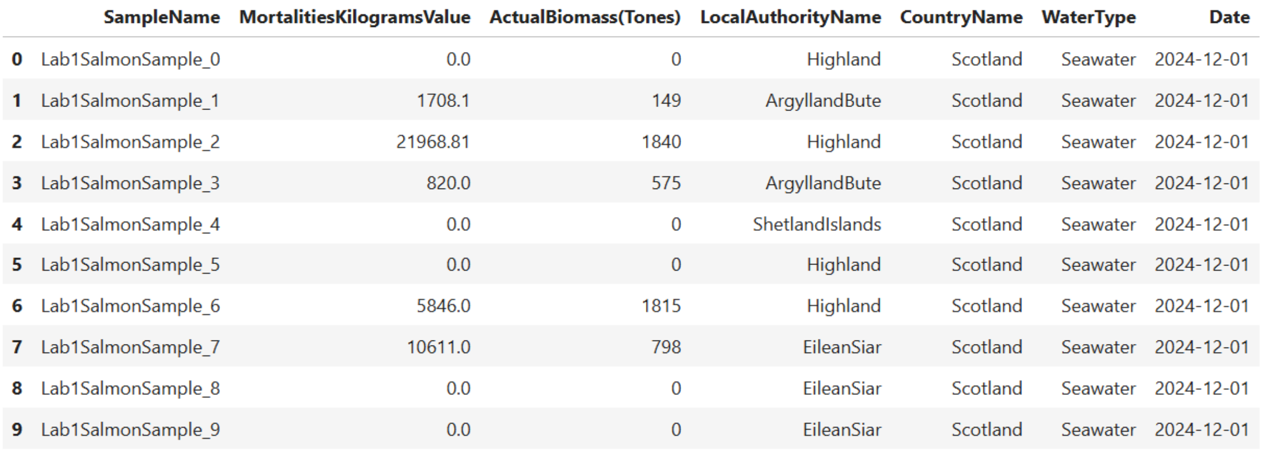
4.6.4.1 Query 1: Retrieve metadata from salmon samples (mortality, biomass, location, water type, and date)

The query in Table 10 retrieves metadata about salmon farming samples, including mortalities (kg), actual biomass, water type, country, local authority, and date. The results shown in Figure 18.

Table 10: Retrieve metadata from salmon samples (mortality, biomass, location, water type, and date).

```
PREFIX decide: <http://www.purl.org/decide#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX LHO: <http://www.purl.org/decide/LiveStockHealthOnto/LHO#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
SELECT
```

```
  (strafter(str(?Sample), "#") AS ?SampleName)
  (strafter(str(?MortalitiesKilograms), "#") AS ?MortalitiesKilogramsValue)
  (strafter(str(?ActualBiomass), "#") AS ?ActualBiomassValue)
  (strafter(str(?WaterType), "#") AS ?WaterTypeName)
  (strafter(str(?Country), "#") AS ?CountryName)
  (strafter(str(?LocalAuthority), "#") AS ?LocalAuthorityName)
  (strafter(str(?Date), "#") AS ?DateValue)
 WHERE {
  ?Sample rdf:type LHO:SalmonSample .
  OPTIONAL { ?Sample LHO:CS61 ?MortalitiesKilograms . }
  OPTIONAL { ?Sample LHO:CS56 ?ActualBiomass . }
  OPTIONAL { ?Sample LHO:CS60 ?WaterType . }
  OPTIONAL { ?Sample LHO:hasCountry ?Country . }
  OPTIONAL { ?Sample LHO:CS57 ?LocalAuthority . }
  OPTIONAL { ?Sample LHO:hasDate ?Date . }}
```

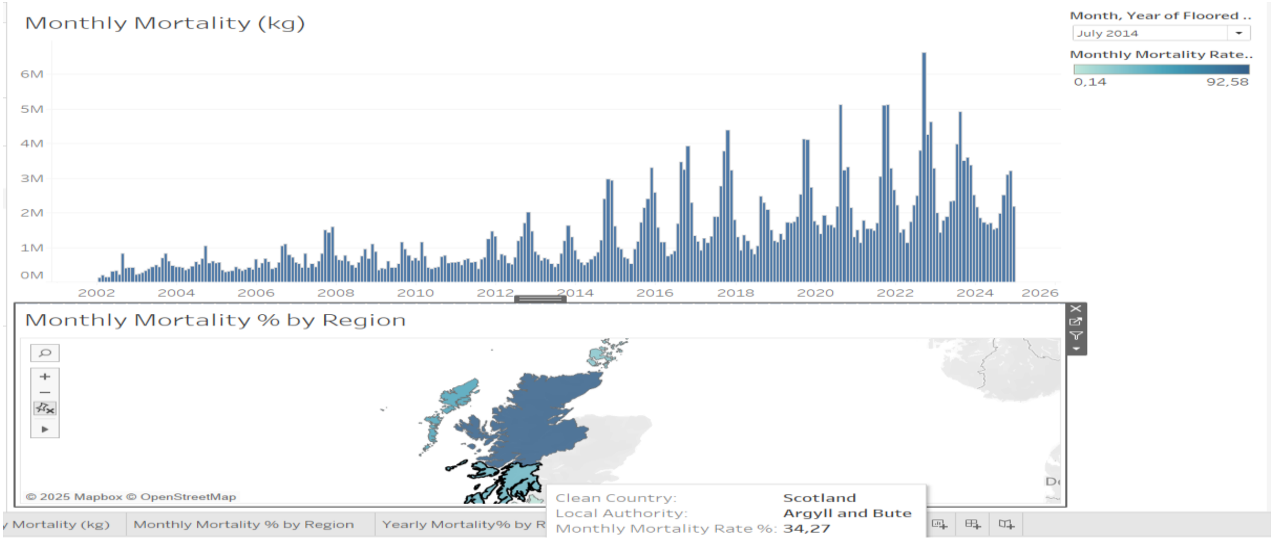


	SampleName	MortalitiesKilogramsValue	ActualBiomass(Tones)	LocalAuthorityName	CountryName	WaterType	Date
0	Lab1SalmonSample_0	0.0	0	Highland	Scotland	Seawater	2024-12-01
1	Lab1SalmonSample_1	1708.1	149	ArgyllandBute	Scotland	Seawater	2024-12-01
2	Lab1SalmonSample_2	21968.81	1840	Highland	Scotland	Seawater	2024-12-01
3	Lab1SalmonSample_3	820.0	575	ArgyllandBute	Scotland	Seawater	2024-12-01
4	Lab1SalmonSample_4	0.0	0	ShetlandIslands	Scotland	Seawater	2024-12-01
5	Lab1SalmonSample_5	0.0	0	Highland	Scotland	Seawater	2024-12-01
6	Lab1SalmonSample_6	5846.0	1815	Highland	Scotland	Seawater	2024-12-01
7	Lab1SalmonSample_7	10611.0	798	EileanSiar	Scotland	Seawater	2024-12-01
8	Lab1SalmonSample_8	0.0	0	EileanSiar	Scotland	Seawater	2024-12-01
9	Lab1SalmonSample_9	0.0	0	EileanSiar	Scotland	Seawater	2024-12-01

Figure 18: Query result for salmon samples (mortality, biomass, location, water type, and date).

4.6.5 Visualisation and analysis

For the Salmon use case, we adopted the same visualization approach used in the cattle case, using Tableau. The SPARQL query results are processed and then integrated into Tableau for visualization. Salmon Barometer, a visualization tool that mimics the functionality of the Norwegian salmon loss dashboard. The Salmon Barometer calculates and displays monthly and yearly mortality rates, providing both temporal trends and geographic distribution through interactive maps. This tool enables stakeholders to monitor salmon health performance across regions and over time, supporting data-driven decision-making, as shown in Figure 14. The visualization is planned to be published on the DECIDE case study website under the Case Studies > Salmon Barometer tab in the near future: https://decide-project-eu.github.io/case-studies-website/case-studies/salmon-barometer.html. The page is currently under development, but it looks as shown in the following Figure 19.



Month, Year of Floored ..: July 2014

Clean Country: Scotland
Local Authority: Argyll and Bute
Monthly Mortality Rate %: 34,27

Figure 19: Salmon barometer, focusing on monthly mortality rate.

5 Discussion and Conclusion

This report explains how we used the ODKFADM framework to improve animal disease surveillance. In the beginning, we reviewed existing ontologies and found that most are not active or available, which makes them hard to reuse. Only a few, like AHOL, ANDO, AGROVOC, and DECIDE, offer useful terms for disease monitoring, but even these do not fully cover important species like cattle, pigs, poultry, and salmon. To fix this, we developed the Livestock Health Ontology (LHO), which is focused on these animals. It builds on the DECIDE ontology and reuses selected terms from AGROVOC and ANDO to keep it focused and effective. Important disease concepts like *Mycoplasma bovis*, *Pasteurella multocida*, *Swine Influenza*, and *Porcine Reproductive and Respiratory Syndrome* are reused from AGROVOC, while terms like *Mycoplasma hyopneumoniae*, *Infectious Bronchitis Virus*, and *Bovine Coronavirus* are taken from the Animal Disease Ontology. LHO also reuses core disease classifications such as *Infectious Disease* and *Viral Disease* from the AHOL ontology, ensuring alignment with existing standards. Specific viral diseases like *Swine Influenza* and *Bovine Respiratory Syncytial Virus (BRSV)* are modelled as instances or subclasses, maintaining semantic clarity and consistency across the ontology. For consistency in geographic data, country and province classes use official URIs from the NCIT thesaurus. We shared LHO on public platforms like GitHub, BioPortal and AgroPortal and conducted a training session during the **ECPLF 2024 conference** to raise awareness and engage the community. In our method, we collected and cleaned diagnostic data, turned it into RDF, linked it to LHO, ran SPARQL queries, and showed the results in Tableau. We tested this full process in case studies for cattle, pigs, poultry, and salmon. These examples showed how the framework can help monitor diseases, support decisions, and improve farm health. The SPARQL results confirmed that our system works for real use.

Looking ahead, we aim to evaluate and compare different data access models developed during the project. These include:

- **Direct Data Sharing** (e.g., manual email exchange, Python processing, Tableau visualization),
- **Centralized Cloud-Based Pipelines**, and
- **Federated Access via Solid Pods**, enabling decentralized, on-demand queries using the Federated Query API and Comunica.

Each model presents unique benefits and challenges, ranging from data privacy and technical complexity to access control and scalability. As part of Deliverable 1.3, we are also conducting scalability testing of the federated endpoint, assessing performance under varying dataset sizes to evaluate its suitability for real-time, multi-source surveillance in distributed environments.

Overall, this report demonstrates how ontology-driven knowledge integration, combined with interactive visualization and modular data access models, can significantly advance the field of digital animal health surveillance across Europe.

References

Alreshidi, E. (2020). SAAONT: Ontological knowledge-based development to support intelligent decision-making systems for Saudi Arabian agriculture. *Int. J. Adv. Appl. Sci*, *7*, 49–59.

Babcock, S., Beverley, J., Cowell, L. G., & Smith, B. (2021). The infectious disease ontology in the age of COVID-19. *Journal of Biomedical Semantics*, *12*, 1–20.

Chen, X., Jia, S., & Xiang, Y. (2020). A review: Knowledge reasoning over knowledge graph. *Expert Systems with Applications*, *141*, 112948.

Cowell, L. G., & Smith, B. (2010). Infectious disease ontology. *Infectious Disease Informatics*, 373–395.

Doan, S., Ngo, Q.-H., Kawazoe, A., & Collier, N. (2008). Building and using geospatial ontology in the BioCaster surveillance system. *Nature Precedings*, 1–1.

Dórea, F. C., Vial, F., Hammar, K., Lindberg, A., Lambrix, P., Blomqvist, E., & Revie, C. W. (2019). Drivers for the development of an Animal Health Surveillance Ontology (AHSO). *Preventive Veterinary Medicine*, *166*, 39–48.

Elena, & Saba. (2022). *DECIDE ONTOLOGY*. https://bioportal.bioontology.org/ontologies/DECIDE/?p=summary

Fahad, M., Javid, T., Beenish, H., Siddiqui, A. A., & Ahmed, G. (2021). Extending ONTAgri with service-oriented architecture towards precision farming application. *Sustainability*, *13*(17), 9801.

Fuentes, V., Martin, T., Valtchev, P., Diallo, A. B., Lacroix, R., & Leduc, M. (2021). *Toward A Dairy Ontology to Support Precision Farming.* 24–35.

Fuentes, V., Martin, T., Valtchev, P., Diallo, A., Lacroix, R., & Leduc, M. (2022). DCPO: The dairy cattle performance ontology, a tool for domain modelling and data analytics. *Of the Series: ICAR Technical Series*, 57.

Gangadharan, V., & Gupta, D. (2020). Recognizing Named Entities in Agriculture Documents using LDA based Topic Modelling Techniques. *Procedia Computer Science*, *171*, 1337-1345.

Golik, W., Dameron, O., Bugeon, J., Fatet, A., Hue, I., Hurtaud, C., Reichstadt, M., Salaün, M.-C., Vernet, J., & Joret, L. (2012). *ATOL: the multi-species livestock trait ontology*. 289–300.

Gómez-Pérez, A., & Corcho, O. (2002). Ontology languages for the Semantic Web. *IEEE Intelligent Systems*, *17*(1), 54-60.

Heward-Mills, D. (2017). Visualizing SPARQL Results in Tableau Public. *OpenLink Virtuoso Weblog*. https://medium.com/virtuoso-blog/spasql-tableau-public-54f31a9767a8.

Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., De Melo, G., Gutierrez, C., Kirrane, S., Gayo, J. E. L., Navigli, R., Neumaier, S., Ngomo, A., Polleres, A., Rashid, S., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., & Zimmermann, A. (2021). Knowledge graphs. *ACM Computing Surveys*, *54*(4), 1–37.

Hu, G., Do, D. N., Gray, J., & Miar, Y. (2020). Selection for favorable health traits: A potential approach to cope with diseases in farm animals. *Animals*, *10*(9), 1717.

Kollia, I., Glimm, B., & Horrocks, I. (2011). SPARQL Query Answering over OWL Ontologies. In: Antoniou, G., et al. The Semantic Web: Research and Applications. ESWC 2011. *Lecture Notes in Computer Science*, vol *6643*. Springer, Berlin, Heidelberg.

Malhotra, A., Younesi, E., Gündel, M., Müller, B., Heneka, M. T., & Hofmann-Apitius, M. (2014). ADO: A disease ontology representing the domain knowledge specific to Alzheimer's disease. *Alzheimer's & Dementia*, *10*(2), 238–246.

Martini, D., Schmitz, M., & Mietzsch, E. (2013). *agroRDF as a Semantic Overlay to agroXML: a General Model for Enhancing Interoperability in Agrifood Data Standards*. CIGR Conference on Sustainable Agriculture through ICT Innovation.

Meyer, A., Faverjon, C., Hostens, M., Stegeman, A., & Cameron, A. (2021). Systematic review of the status of veterinary epidemiological research in two species regarding the FAIR guiding principles. *BMC Veterinary Research*, *17*(1), 270.

Monteiro, A., Santos, S., & Gonçalves, P. (2021). Precision agriculture for crop and livestock farming—Brief review. *Animals*, *11*(8), 2345.

Murali, E., & Anouncia, S. M. (2022). An Ontology-based Knowledge Mining Model for Effective Exploitation of Agro Information. *IETE Journal of Research*, 1–18.

Newton, J. E., Nettle, R., & Pryce, J. E. (2020). Farming smarter with big data: Insights from the case of Australia's national dairy herd milk recording scheme. *Agricultural Systems*, *181*, 102811.

Niloofar, P., Lazarova-Molnar, S., Francis, D. P., Vulpe, A., Suciu, G., & Balanescu, M. (2020). *Modeling and simulation for decision support in precision livestock farming*. 2601–2612.

Peyre, M., Hoinville, L., Njoroge, J., Cameron, A., Traon, D., Goutard, F., Calba, C., Grosbois, V., Delabouglise, A., & Varant, V. (2019). The RISKSUR EVA tool (Survtool): A tool for the integrated evaluation of animal health surveillance systems. *Preventive Veterinary Medicine*, *173*, 104777.

Rajbhandari, S., & Keizer, J. (2012). The AGROVOC Concept Scheme - A Walkthrough, *Journal of Integrative Agriculture*, 11(5), 694-699.

Ren, G., Lin, T., Ying, Y., Chowdhary, G., & Ting, K. (2020). Agricultural robotics research applicable to poultry production: A review. *Computers and Electronics in Agriculture*, *169*, 105216.

Rosa, G. J. (2021). Grand challenge in precision livestock farming. *Frontiers in Animal Science*, *2*, 650324.

Salaun, M.-C., Yon, J., Le Bail, P.-Y., & Reichstadt, M. (2019). *Animal health ontology for livestock*.

Schriml, L. M., Arze, C., Nadendla, S., Chang, Y.-W. W., Mazaitis, M., Felix, V., Feng, G., & Kibbe, W. A. (2012). Disease Ontology: A backbone for disease semantic integration. *Nucleic Acids Research*, *40*(D1), D940–D946. https://doi.org/10.1093/nar/gkr972

Schriml, L. M., Munro, J. B., Schor, M., Olley, D., McCracken, C., Felix, V., Baron, J. A., Jackson, R., Bello, S. M., & Bearer, C. (2022). The human disease ontology 2022 update. *Nucleic Acids Research*, *50*(D1), D1255–D1261.

Shankhdhar, G. K., Sharma, R., & Darbari, M. (2021). SAGRO-lite: A light weight agent based semantic model for the internet of things for smart agriculture in developing countries. *Semantic IoT: Theory and Applications: Interoperability, Provenance and Beyond*, 265–302.

Sophie. (2018). *Animal Disease ontology(ANDO)*. INRAE. https://agroportal.lirmm.fr/ontologies/ANDO

Tomic, D., Drenjanac, D., Hoermann, S., & Auer, W. (2015). Experiences with creating a precision dairy farming ontology (DFO) and a knowledge graph for the data integration platform in agriOpenLink. *Journal of Agricultural Informatics*, *6*(4).

Tomic, S. D. K., Fensel, A., Aschauer, C., Schulmeister, K. G., Riegler, T., Handler, F., Otte, M., & Auer, W. (2013). *agriopenlink: Towards adaptive agricultural processes enabled by open interfaces, linked data and services*. 408–413.

Topalis, P., Mitraka, E., Dritsou, V., Dialynas, E., & Louis, C. (2013). IDOMAL: The malaria ontology revisited. *Journal of Biomedical Semantics*, *4*(1), 16. https://doi.org/10.1186/2041-1480-4-16

Valentin, S., Arsevska, E., Rabatel, J., Falala, S., Mercier, A., Lancelot, R., & Roche, M. (2021). PADI-web 3.0: A new framework for extracting and disseminating fine-grained information from the news for animal disease surveillance. *One Health*, *13*, 100357.

van Schaik, G., Hostens, M., Faverjon, C., Jensen, D. B., Kristensen, A. R., Ezanno, P., Frössling, J., Dórea, F., Jensen, B.-B., & Carmo, L. P. (2023). The DECIDE project: From surveillance data to decision-support for farmers and veterinarians. *Open Research Europe*, *3*, 82.

Vasconcelos Gioia, G., Lamielle, G., Aguanno, R., ElMasry, I., Mouillé, B., De Battisti, C., Angot, A., Ewann, F., Sivignon, A., & Donachie, D. (2021). Informing resilience building: FAO’s Surveillance Evaluation Tool (SET) Biothreat Detection Module will help assess national capacities to detect agro-terrorism and agro-crime. *One Health Outlook*, *3*, 1–13.