

Faculty of Pure and Applied Sciences
Department of Computing and Informatics

AN ONTOLOGY-DRIVEN FRAMEWORK FOR ANIMAL TRACEABILITY

by

TSHEPISO LARONA MOKGETSE
Student ID number: 21100029
BSc (Computer Systems Engineering) (BAC), MSc (Computer Applications) (GTU)

A Thesis Submitted to the Faculty of Pure and Applied Sciences in Partial Fulfilment of the Requirements for the Award of the Degree of Doctor of Philosophy in Computer Science of BIUST

Supervisor(s): Dr, Hlomani Hlomani and Dr, Tshiamo Sigwele
Department of Computing and Informatics
Faculty of Pure and Applied Sciences, BIUST
E-mail Address: hlomanihb@biust.ac.bw, sigwelet@biust.ac.bw

July, 2025

Table of Contents

I

DECLARATION REGARDING THE WORK AND COPYRIGHT VI

CERTIFICATION VII

EVIDENCE OF ORIGINALITY VIII

ACKNOWLEDGEMENTS IX

DEDICATION X

PUBLICATION CONTRIBUTIONS XI

ABSTRACT XII

LIST OF TABLES XIII

LIST OF FIGURES XIV

LIST OF EQUATIONS XV

ACRONYMS XVI

1 CHAPTER 1 INTRODUCTION 1

1.1 Introduction to Animal Traceability 1
1.2 Challenges of Animal Traceability 1
1.3 The case of Botswana 3
1.3.1 Challenges Facing Botswana's Animal Traceability Systems 3
1.4 An Introduction to Ontologies 5
1.4.1 Benefits of Ontologies 6
1.5 Problem Statement 7
1.6 Objectives 8
1.6.1 Main Objective 8
1.6.2 Specific Objective 8
1.7 Research Questions 9
1.7.1 Main Research Question 9
1.7.2 Specific Research Questions 9
1.8 Significance of the Study 9
1.9 Contributions to this study 10
1.9.1 Main Contributions 10

1.10 THESIS STRUCTURE 12

2 CHAPTER 2 LITERATURE REVIEW 13

2.1 INTRODUCTION 13
2.2 ANIMAL TRACEABILITY 13
2.2.1 IDENTIFIED USES AND GAPS OF ANIMAL TRACEABILITY TECHNOLOGIES 13
2.2.2 ANIMAL TRACEABILITY FRAMEWORKS 18
2.2.3 ANIMAL TRACEABILITY IN BOTSWANA 20
2.3 ONTOLOGIES 22
2.3.1 ONTOLOGY DEVELOPMENT METHODOLOGIES 23
2.3.2 ONTOLOGY DEVELOPMENT TOOLS 26
2.3.3 ONTOLOGY LANGUAGES 30
2.3.4 TRACEABILITY ONTOLOGY FRAMEWORKS 32
2.4 CHAPTER SUMMARY 39

3 CHAPTER 3 RESEARCH DESIGN AND DATA ANALYSIS 40

3.1 INTRODUCTION 40
3.2 RESEARCH DESIGN AND METHODOLOGY OF THE STUDY 40
3.3 RESEARCH APPROACH 42
3.3.1 MIXED RESEARCH 42
3.4 RESEARCH METHODS 43
3.4.1 QUANTITATIVE RESEARCH 43
3.4.2 QUALITATIVE RESEARCH 43
3.5 DATA COLLECTION 44
3.5.1 QUALITATIVE 44
3.5.2 QUANTITATIVE 45
3.5.3 STUDY AREA 46
3.5.4 DATA SOURCES 46
POPULATION AND SAMPLING 47
3.6 VALIDITY AND RELIABILITY 49
3.6.1 PILOT STUDY 50
3.6.2 ETHICAL CONSIDERATIONS 50
3.7 DATA ANALYSIS 51
3.7.1 QUANTITATIVE 51
3.7.2 QUALITATIVE 51
3.8 TRIANGULATION 52
3.9 DATA ANALYSIS FOR THIS STUDY 53
3.9.1 QUALITATIVE DATA ANALYSIS 54
3.9.2 QUANTITATIVE DATA ANALYSIS 62
3.10 CHAPTER SUMMARY 62

4 CHAPTER 4: FINDINGS AND DISCUSSION 64

4.1 INTRODUCTION 64

4.2 QUANTITATIVE RESEARCH RESULTS – DESCRIPTIVE STATISTICS 64
4.2.1 RESPONSE RATE 64
4.2.2 RESPONDENTS DEMOGRAPHIC 65
4.2.3 CENTRAL TENDENCY OF ANIMAL TRACEABILITY AND TRACKING EFFECTS 72
4.3 QUALITATIVE RESEARCH RESULTS 73
4.3.1 THEMES AND CONCEPTS 73
4.4 CHAPTER SUMMARY 84

5 CHAPTER 5 PROPOSED ONTOLOGY-DRIVEN ANIMAL TRACEABILITY FRAMEWORK & ONTOLOGY DEVELOPMENT 86

5.1 INTRODUCTION 86
5.2 PROPOSED ONTOLOGY-DRIVEN ANIMAL TRACEABILITY FRAMEWORK (ODF-AT) 87
5.2.1 THE INPUT LAYER OF THE ODF-AT 87
5.2.2 THE SEMANTIC CORE LAYER OF THE ODF-AT 88
5.2.3 THE KNOWLEDGE MANAGEMENT LAYER OF THE ODF-AT 88
5.2.4 THE APPLICATION LAYER OF THE ODF-AT 89
5.3 ONTOLOGY DEVELOPMENT APPROACH 90
5.4 ANIMAL TRACEABILITY ONTOLOGY (ATO) DEVELOPMENT 90
5.4.1 ANIMAL TRACEABILITY CORE FEATURES 91
5.4.2 ONTOLOGY DEVELOPMENT - METHONTOLOGY 92
5.5 CHAPTER SUMMARY 125

6 CHAPTER 6 ONTOLOGY EVALUATION 126

6.1 INTRODUCTION 126
6.2 STRUCTURAL EVALUATION 127
6.2.1 ONTOMETRICS EVALUATION TOOL 127
6.3 USABILITY PROFILING EVALUATION 133
6.4 FUNCTIONAL EVALUATION 135
6.4.1 PELLET REASONER 136
6.4.2 KNOWLEDGE METRICS 136
6.4.3 AVERAGE POPULATION (AP) 136
6.4.4 CLASS RICHNESS (CR) METRIC 137
6.4.5 CLASS METRIC 137
6.4.6 EVALUATION USING APACHE JENA FUSEKI 138
6.5 ONTOLOGY PITFALL SCANNER! (OOPS!) EVALUATION 141
6.6 CHAPTER CONCLUSION 144

7 CONCLUSION, LIMITATIONS AND RECOMMENDATION 146

7.1 INTRODUCTION 146
7.2 SUMMARY 146
7.3 LIMITATIONS 148
7.4 RECOMMENDATIONS FOR FURTHER RESEARCH 149

8 REFERENCES 150

DECLARATION REGARDING THE WORK AND COPYRIGHT

Candidate: TSHEPISO LARONA MOKGETSE_____

Student ID: 21100029_____

Thesis Titled: AN ONTOLOGY-DRIVEN FRAMEWORK FOR ANIMAL TRACEABILITY_____

I, **TSHEPISO LARONA MOKGETSE**, certify that the Thesis is all my own original work and that I have not obtained a degree in this University or elsewhere on the basis of any of this work.

This dissertation/thesis is copyright material protected under the Berne Convention, the Copyright and Neighbouring Rights Act, Act. No. 8 of 2000 and other international and national enactments, in that behalf, on intellectual property. It must not be reproduced by any means, in full or in part, except for short extracts in fair dealing; for researcher private study, critical scholarly review or discourse with an acknowledgement, without the written permission of the office of the Postgraduate School, on behalf of both the author and the BIUST.

Signed: ---------- Date: ------20 June 2025----------

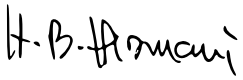
Primary Supervisor: DR HLOMANI HLOMANI_____

I, the Candidate's **Primary Supervisor**, hereby confirm that I have inspected the above-titled thesis and, to the best of my knowledge, it is based on the original work of the candidate.

Signed: ---------- Date: ---16 September 2025----------

CERTIFICATION

The undersigned certifies that they have read and hereby recommends for acceptance by the Faculty of Pure and Applied Sciences a dissertation titled, AN ONTOLOGY-DRIVEN FRAMEWORK FOR ANIMAL TRACEABILITY, in fulfillment of the requirements for the Degree of Doctor of Philosophy in Computer Science of BIUST.



.....

DR HLOMANI HLOMANI

(Primary supervisor)

EVIDENCE OF ORIGINALITY

The work in this dissertation is original and most of the content from the chapters have been published and accepted in high quality journals. This shows that other professionals in the ontology and animal traceability field have examined and accepted my work after running plaigiarism checks, methods and results analysis on my articles published showing that my work is original and there is no plaigiarism as evident from the publications. A list of the articles is provided on Page x Publication Contributions in this thesis.

BOTSWANA INTERNATIONAL UNIVERSITY OF SCIENCE & TECHNOLOGY

FORM PGS - 330

POSTGRADUATE SCHOOL

Antiplagiarism check

MEMORANDUM

Postgraduate Candidate:__Tshepiso Larona Mokgetse_____ Student ID: _21100029_____

Degree level: PhD [ ], or Master [ ] *(tick the appropriate)*

Thesis Titled: **__AN ONTOLOGY-DRIVEN FRAMEWORK FOR ANIMAL TRACEABILITY____**

The Antiplagiarism Check had been completed and the results is detailed below.

Software and version used: **_TURNITIN FEEDBACK STUDIO__**

Outcome: **8% SIMILARITY INDEX SCORE_**

Officer who completed the check (please write in caps or type): **_KELEBONYE BAGAI**

Signed:  Date: 19/06/2025

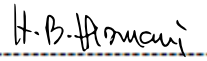
COMMENTS BY THE PRIMARY SUPERVISOR (*if requested*):

_____

_____

_____

Primary supervisor (please write in caps or type) **_HLOMANI HLOMANI_____**

Signed:  Date: 16/09/2025

ACKNOWLEDGEMENTS

First and foremost, I would like to thank the Almighty God for His Grace and keeping me strong and being the calm in the storm throughout the journey of this great achievement. I convey the highest gratitude to my esteemed Supervisors Dr Hlomani Hlomani and Dr Tshiamo Sigwele for their invaluable advice, immense knowledge, and abundant experience that encouraged me during my academic research time and life. My extreme appreciation to my family – my mother P. Mokgetse who is my pillar and best friend, my father W. Mokgetse who is my heart, my mother A. Mosweu who is my strength and peace, my grandmother R. Engleton who was my motivator, and friends for their treasured support and encouragement throughout my studies. Finally, I am thankful to myself for working hard and sacrificing to fulfill part of the vision and purpose that Jesus Christ planned for my life.

DEDICATION

I dedicate this work to God for the completion of the research despite the obstacles. I am grateful to my family for their unrelenting support that fueled my focus and discipline to succeed in my academics. My supervisors for their guidance. Special appreciation to my mother who has been right by my side throughout this journey.

"For I know the plans I have for you, declares the Lord, plans to prosper you and not to harm you, plans to give you hope and a future".

PUBLICATION CONTRIBUTIONS

Thre results and contributions of this research was shared with the body of knowledge through a number of research publications as detailed below:

1. **Tshepiso Larona Mokgetse**, Hlomani Hlomani, Tshiamo Sigwele, Irina Zlotnikova, An Ontology-driven Framework for Animal Traceability in Botswana. Journal: Agris online – Papers in economics an Informatics. 2025, Volume 17, Issue 1, pp. 95-112. DOI 10.7160/aol.2025.170108
2. **Tshepiso Larona Mokgetse**, Hlomani Hlomani, Tshiamo Sigwele, Irina Zlotnikova, Leveraging Ontology-Based Knowledge Management for Improved Data Quality and Enhanced Trade Compliance in Livestock Traceability Systems in Botswana, Interdisciplinary Journal of Information, Knowledge, and Management 2024, Volume 18, pp. 035. DOI: https://doi.org/10.28945/5409
3. **Tshepiso Larona Mokgetse**, Hlomani Hlomani, Sridaran Rajagopal, Mitigating the Technological Challenges of Regenerative Farming with an integrated Framework merging IoT, Ontologies, and GIS, 3rd International Conference on Advancements in Smart Computing and Information Security (ASCIS - 2024). A Springer Nature Computer Science book series: Communications in Computer and Information Science, 2025, Volume 2428, pp. 95-106. DOI: 10.1007/978-3-031-86302-8_6
4. **Tshepiso Larona Mokgetse**, Hlomani Hlomani, Sridaran Rajagopal, A FAIRification Framework for the Findability, Accessibility, Interoperability and Reusability of Cyber Security Ontologies Using FAIR Data Principles, International Journal of Advanced Networking and Applications 2023, Volume 15, Issue 4, pp. 185–195. DOI: http://dx.doi.org/10.35444/IJANA.2024.15433
5. **Tshepiso Larona Mokgetse**, Hlomani Hlomani, Tshiamo Sigwele, Innovative Emerging Ontology-driven Frameworks: A Systematic Literature Review, Indonesian Journal of Electrical Engineering and Informatics (IJEEI) 2023, Volume 11, Issue 4, pp. 1064 – 1081. DOI: http://dx.doi.org/10.52549/.v11i4.5106
6. **Tshepiso Larona Mokgetse**, Hlomani Hlomani, Tshiamo Sigwele, A Systematic Review of Animal Identification and Traceability: Benefits and Challenges. Journal: Scientific African. 2025. Publishing: Accepted with Revisions.

ABSTRACT

Traceability systems are crucial for regulatory compliance and consumer trust in food supply integrity and animal protection. Animal traceability systems ensure food safety and disease outbreak control. However, the current systems lack interoperability and are badly integrated across multiple stakeholders. The non-interoperable animal traceability systems pose challenges for farmers and the government in tracking stray animals, diseases, feeding, vaccinations, and livestock export regulations. The main cattle identification and traceability system being used in Botswana is the non-interoperable Botswana Livestock Identification and Traceability System. The six stakeholders concentrate on cow but lack a single system for interoperability, integration, standardization, and domain understanding among them. This research proposes the development of an interoperable and ontology-driven animal traceability framework to solve the gap of poor interoperability in current frameworks for animal traceability, with Botswana as a case study. Limitations of current traceability frameworks were identified, and the components for the design and development of the proposed interoperability framework were generated using the mixed method design methodology - qualitative and quantitative research. Involving robust data analysis. Quantitative identified system gaps and patterns in traceability practices though descriptive statistics. Qualitative research carried out a thematic analysis using NVivo tool on stakeholder interviews and documents, for in-depth understanding of socio-technical aspects of traceability systems. Together the designs reinforced the validity and reliability of the findings. The analysis identified cattle, farmers, system, and records as the key concepts. 39% of Botswana participants find BAITS ineffective, making it the main challenge of animal traceability.

An interoperable ontology-driven animal traceability framework is developed. In addition, a new and novel animal traceability ontology was developed using the METHONTOLOGY ontology development methodology, covering knowledge acquisition, specification, conceptualization, formalization, and evaluation. The process ensures a transparent and systematic ontology development of the animal traceability domain to model the key traceability components. The ontology was evaluated for quality based on consistency, clarity, adaptability, and coverage. The Ontometric evaluation tool was used to evaluate revealing a high level of inheritance richness 0.692308, relationship richness 0.678571, and average depth 0.9, indicating detailed knowledge ontology, strong class connections, and seamless data integration and exchange. The evaluation of reference and candidate ontologies revealed 66.67% of animal traceability ontology is more interoperable than animal health ontology at 33.33%. Altogether, results indicate that the framework achieved interoperability and standardization, producing seamless integrated systems. Overall, this study contributes both as a template that is methodological for agricultural research in semantic systems. Also, providing a foundation for adaptable solutions to support worldwide supply chain compliance and strategies in animal health. The ontology-driven framework is designed as a guiding principle for system developers, a blueprint to implement semantically correct integrated traceability. The ontology is developed to contribute semantic interoperability and can be integrated with data infrastructures already existing.

LIST OF TABLES

Table 1 Animal Traceability Stakeholders and Their Roles. 5
Table 2 Animal traceability technologies.....15
Table 3 Methodologies compared beside the ontology development life cycle [45] 24
Table 4 Comparison between the ontology development methodologies [104].....24
Table 5 Ontology methodologies compared [99]. 25
Table 6 General description of the developers, release dates, current versions, and pricing policy of the ontology tools [109], [110], [111].....27
Table 7 Comparison of ontology tools interoperability translations to/from languages.....29
Table 8 Usability features of the ontology development tools [48].....29
Table 9 Comparison of the Ontology languages [106]. 31
Table 10 Traceability ontology frameworks 34
Table 11Task list for the interview process for the research.....45
Table 12 Participants selected for study.....49
Table 13 Pilot study feedback examples.....50
Table 14 Sample attributes and values of case Person.....58
Table 15 Cattle count, theft, and found distribution 69
Table 16 Central Tendency statistics for animal traceability. 72
Table 17 Glossary of Terms for Animal Traceability Ontology.....97
Table 18 Glossary of terms for the class Cattle and class Farmer. 97
Table 19 Glossary of terms for class Farm Employee and class Tracking. 98
Table 20 Glossary of terms for the class Traceability and class Disease.....98
Table 21 Glossary of terms for the class Organizations and class Visitations. Glossary of terms for the class Organizations and class Visitations.....98
Table 22 Glossary of terms for the class Vaccination and class Farm. 99
Table 23 Glossary of terms for the class Financial and class Inventory. 100
Table 24 Glossary of terms for the class Market. 100
Table 25 Semantics of the Binary Object Properties 103
Table 26 Concept dictionary used to specify the properties, instances, and relations. 106
Table 27 Merger among the relations and attributes built in the concept dictionary. 107
Table 28 Description of the instance attributes from the concept dictionary.....108
Table 29 Description in detail of the class attributes. 108
Table 30 Axioms description of concepts definable in OWL. 110
Table 31 Description of instance Farmer. 110
Table 32 Description of instance Farm Employee. 110
Table 33 Description of instance Organization. 111
Table 34 Description of instance Traceability. 111
Table 35 Description of instance Vaccination. 111
Table 36 Description of instance Tacking. 112
Table 37 Description of instance Visitations. 112
Table 38 Description of instance Financials. 112
Table 39 Description of instance Inventory. 113
Table 40 Description of instance Disease. 113
Table 41 Description of instance Market. 113
Table 42 Description of instance Cattle. 114
Table 43 Description of instance Farm. 114
Table 44 Evalution Techniques and Metrics 127
Table 45 Comparison of the ATO and AHO 133

LIST OF FIGURES

Figure 1 Illustration of the modular design of the dairy ontology [118] 35
Figure 2 Pictorial of FoodOn Ontology [18] 35
Figure 3 Illustration of the taxonomy process of Pasta Filata [19] 36
Figure 4 Conceptual Framework for an Ontology-Based Traceability System for Halal Status of Flavour [120]. 37
Figure 5 Research Design of the Study Framework. 41
Figure 6 The interview protocol refinement method [153]. 45
Figure 7 Stages of planning a questionnaire [156]. 46
Figure 8 Concurrent triangulation design [138]. 52
Figure 9 Schematic diagram of the study data analysis process. 54
Figure 10 Import panel of NVivo software. 55
Figure 11 Sample documents imported in the Files subsection of the Data section. 56
Figure 12 Data Annotations 56
Figure 13 Sample Attributes 57
Figure 14 Sample presentation of the imported demographic of respondents. 58
Figure 15 Sample for coding 59
Figure 16 Sample for nodes 60
Figure 17 Explore panel with query options from NVivo software. 61
Figure 18 Response rate by district of respondents. 65
Figure 19 Age of the respondents. 66
Figure 20 Gender distribution of respondents. 67
Figure 21 Level of education representation of respondents. 68
Figure 22 Respondents source of income. 68
Figure 23 Cattle theft vs cattle found by district. 70
Figure 24 Cattle theft vs cattle theft in Botswana 71
Figure 25 The outlook of respondents to animal traceability challenges. 72
Figure 26 Word Cloud for animal traceability data. 73
Figure 27 Cluster groupings of animal traceability data. 75
Figure 28 Coding matrix of the five main concepts deducted from the text data. 76
Figure 29 Coding matrix illustration of interview results. 78
Figure 30 Main tree hierarchy map from animal traceability data. 79
Figure 31 Extract category 1 from the expansion of the “farm” component of the main tree hierarchy map Figure 30. 80
Figure 32 Extract 2 from the expansion of the “system” component of the main tree hierarchy map Figure 30. 81
Figure 33 Extract 3 from the expansion of the “cattle” component of the main tree hierarchy map Figure 30. 81
Figure 34 Project map of the four main components generated by NVivo software. 83
Figure 35 Project map of the four main components, including the coding relationships from each District. 84
Figure 36 The Ontology-driven Framework for Animal Traceability. 87
Figure 37 Ontology Development Activities as proposed by METHONTOLOGY [75]. 93
Figure 38 Knowledge acquisition and Specification activities of the animal traceability ontology. 95
Figure 39 The conceptualization steps for developing animal traceability ontology based on METHONTOLOGY. 96
Figure 40 Taxonomy for animal traceability. 101
Figure 41 Ad hoc binary relations built for the animal traceability ontology. 102
Figure 42 Protege Startup process 115
Figure 43 Formalization process for animal traceability ontology. 116

Figure 44 Class and Sub-class creation 117
Figure 45 Complete list of the animal traceability ontology classes, sub-classes, and sub-sub-classes. 117
Figure 46 Creating object properties for the animal traceability ontology. 117
Figure 47 depicts the list of object properties created in Protégé for the animal traceability ontology. 118
Figure 48 The domains and ranges for object property “employedBy”. 118
Figure 49 Creation of Data Properties 119
Figure 50 Created Data Properties of ATO using Protege 120
Figure 51 Adding features to data types 121
Figure 52 Creation of individual “foot and mouth” for class disease. 122
Figure 53 Individuals for animal traceability 122
Figure 54 Pictorial from Protégé OWLViz of the animal traceability ontology. 123
Figure 55 Pictorial from Protégé OntoGraf of the animal traceability ontology. 124
Figure 56 Base Metric 128
Figure 57 Schema Metrics 129
Figure 58 Graph Metrics 131
Figure 59 Class axioms Metric results from computing AnimalTraceabilityOntology.owl file 134
Figure 60 Property axioms 135
Figure 61 Knowledge metrics results from OntoMetrics 136
Figure 62 Class Metric 138
Figure 63 Running SPARQL query to search for farmer employee. 140
Figure 64 Output from SPARQL query to searching for farmer employee. 140
Figure 65 Running SPARQL query to search for name of farm the farmer owns. 140
Figure 66 Output from SPARQL query to searching for farm name. 141
Figure 67 Running SPARQL query to search for the farm location. 141
Figure 68 Output from SPARQL query to search for farm location 141
Figure 69 OOPS! Pitfall indicator 142
Figure 70 ATO results from OOPS! pitfall scanner 142
Figure 71 ATO pitfall detected P13: Inverse relationships not explicitly declared 142
Figure 72 ATO pitfall detected P22: Using different naming conventions 143
Figure 73 ATO pitfall detected P24: Using recursive definition 143
Figure 74 ATO pitfall detected P30 Equivalent classes not explicitly declared 143
Figure 75 ATO pitfall detected P41: No license declared 143
Figure 76 ATO results presenting no pitfalls detected in the ontology 144

LIST OF EQUATIONS

Equation 1 Cluster analysis algorithm 74
Equation 2 Coding matrix algorithm 76
Equation 3 Hierarchy tree map algorithm 78
Equation 4 Inheritance Richness (IR) 130
Equation 5 Relationship richness (RR) 130
Equation 6 Class/Relation ratio (CRR) 131
Equation 7 Average Population (AP) 137
Equation 8 Class Richness (CR) metric 137

ACRONYMS

BMC – Botswana Meat Comissioner
CAQDAS - Computer Assisted Qualitative Data Analysis
DCOP - Dairy Cattle Performance Ontology
DONT - Dedicated Domain Ontology
EU – European Union
FAIR – Findable Accessible Inteoperable Reusable
FMD – Foot and Mouth Disease
FTTO - Food Track & Trace Ontology
GDP – Gross Domestic Product
GPS - Global Positioning System
GS1 – Global Standards
GSM - Global System for Mobile communication
GUI – Graphical User Interface
HACCP - Hazard Analysis Critical Control Point
HTML - HyperText Markup Language
ICT - Information Communications Technology
ID – Indetification Device
JSON - JavaScript Object Notation
LITS – Livestock Identification and Trace-back System
MOA – Ministry of Agriculture
OIE - World Organization for Animal Health
OWL – Web Ontology Language
PDA - Personal Digital Assistant
PLF - Precision Livestock Farming
RDF - Resource Description Framework
RFID – Radio Frequency Identification Device
SPARQL - Simple Protocol for RDF Query Language
SPSS - Statistical Package for Social Sciences
SQL - Structural Query Language
SQWRL - Semantic Query-Enhanced Web Rule Language
TOVO - Toronto Virtual Enterprise
TSV - Tab-Separated Values
UML - Unified Modeling Language
WHO – World Health Organization
WTO - World Trade Organization
WSN - Wireless sensor networks
XML - Extensible Markup Language

1 CHAPTER 1 INTRODUCTION

The first chapter of this research introduces the main concept of animal traceability and ontologies. The academic context highlights and discusses the work and research that other researchers have carried out on ontologies and animal traceability, and further articulates the existing gaps within the literature regarding animal traceability systems and frameworks internationally. The chapter provides the reader with the context of this research by giving an insight into the challenges that informed the study. In alliance with the literature review and research methodology, the information conveyed in this section is presented to show the contributing research that sparked the initiation of this study.

1.1 Introduction to Animal Traceability

The EU regulation from the research presented by Zhao et al. [1] traceability is the ability to track and monitor food, feed, food-producing animals, or substances meant to be, or anticipated to be, included in feed throughout the phases of production, processing, and distribution. Traceability is the capacity to preserve animal or product identity throughout the whole production process [2]. A clearer and more detailed definition by Caporale et al. [3] Traceability is documenting the important elements of an animal, that is, processes, movements, and controls, that are required to describe the history of the animal's life.

Animal traceability systems are a means of using devices to track the life of livestock from birth until they are slaughtered. Detailing their movement, vaccination, diseases, and feeding. The tremendous growth in the demand for animal products globally has ignited the need for the implementation of robust animal traceability systems. The traceability systems are not only imperative to meet the regulatory requirements but necessary for building trust with consumers in the food supply integrity and in animal protection. Preservation of animals, in this case cattle, has provoked a lot of attention from individuals, external organizations, and governments. Traceability requires a focus on the animal's health and the hygiene of the food for consumers. That is why it was important to create systems that would manage the traceability of livestock. The modern system for the control of food safety, with the integration of food hygiene and the health of animals, is the foundation of traceability [4].

1.2 Challenges of Animal Traceability

The minimum for a traceability system is listed as having identification elements, information flow, and databases [5]. As a precursor to studies like this one, the World Organization for Animal Health (OIE) has been hosting international seminars since 1998 to address the problems of animal traceability for global impact [6]. Therefore, it is not a trivial problem. Their focus has been on the devices being used to monitor the animals and less on the management of the data collected. In addition, there also exists the Agri-Traceability Québec (ATQ), whose mandate is to group, in a single place, data related to the traceability of agricultural sectors when entered into the system through the development of multispecies databases [7]. Also, Zhao et al. [1] talks of another system that only focuses on the device used for animal traceability, with an ineffective information management system that relies on how the animal will be slaughtered to get information.

There are a few other animal traceability systems that deal with animal food safety, diseases, and export regulations and a few frameworks that talk about animal food safety. For example, the general framework for animal food safety traceability using Global Standards (GS1) and Radio Frequency Identification Device (RFID) by Cao [8] is a framework in animal traceability that deals with the tracing of animal health for the safety of the animal products that are going to be sold. There is an absence of

interoperability in all the systems mentioned, even though they operate in an environment with multiple processes and organizations.
Cattle diseases have played a major role in the development of animal traceability systems. Diseases like Foot and Mouth Disease (FMD), zoonotic disease, and bovine spongiform encephalopathy disease have caused disruptions and conflicts not only in the survival of farmers but the economy of countries and relations between countries. FMD and bovine spongiform encephalopathy are said to be widely spread diseases in Asian, African, Middle Eastern, and European countries, all linked to infected animals being traded [9]. FMD outbreaks require significant resources for control, impacting traceability, and wildlife conflicts in FMD-endemic areas worsen livestock management [10]. Environmental factors like high temperatures and inefficiencies in traditional farming, such as low off-take rates and high mortality, also hinder the effectiveness of the traceability system [10]. Overall, these issues underscore the need for improved infrastructure and technology solutions to support traceability [11].

The serious impact of cattle disease makes it difficult for animal products to enter markets easily for the safety of compromising public health. World Health Organization (WHO) made compulsory agreements on Sanitary and Phytosanitary Measures, which made trade barriers that were stricter for the traceability of animals [12]. The European Union (EU) also created regulations [13] under articles 112 to 115 of Regulation (EU) 2016/429, which presented tighter regulations for the import/export of cattle requiring effective animal traceability systems. The introduction of compulsory animal traceability systems required farmers and countries to implement animal traceability systems to continue managing their livestock and complying with the new regulations.

In other African countries, such as Kenya, Ethiopia, and Tanzania, animal traceability systems face challenges due to inadequate infrastructure and high costs associated with technologies like RFID, making adoption difficult for many farmers [14]. There is also a lack of regulation and oversight to enforce traceability practices, and limited stakeholder coordination further impedes progress [14]. In Kenya, the absence of proper animal identification methods affects disease surveillance and control, particularly for diseases like African swine fever and cysticercosis [15]. In Tanzania, a pilot study using the Livestock Identification and Traceability System (LITS) showed potential for tracking cattle, but accurate data collection, lack of incentives for adoption, and limited stakeholder coordination remain issues [16]. The unreliable Internet connectivity in rural areas complicates real-time data collection and monitoring, hindering the effectiveness of traceability efforts [17].

The essence of these problems is therefore that despite the introduction of several animal traceability systems and platforms, there are still few to no efficient solutions both to the usage of the systems and the animal data. In general, the methodologies used for animal traceability are not effective and are not gaining the results they should be. Both Dooley et al. [18], and Magliulo et al. [19] say the complexity of the traceability systems that exist has been a drawback; they are limited in input data, meaning only critical events are given attention. The identifiable common challenges of the systems include interoperability, integration, regulatory compliance, data that is complete and accurate, and a stable technology infrastructure. The main challenge is interoperability.

Power et al. [20] define interoperability as a means for two or more systems that are independent to perform the exchange of information, and interaction in a meaningful way. Moreover, use the exchanged information to achieve their objectives. The EU concurs that interoperability with technology can help organizations communicate towards reaching mutually beneficial goals [20]. This will involve an exchange of data through the systems to share information and knowledge. According to Koncová and Kremeňová [21] the interconnection of interoperable systems is important because it provides efficient communication, data protection, productivity increased, consistent data, timely data, costs reduced, error rate reduced, and time for obtaining data reduced.

1.3 The case of Botswana

The main animal traceability system being used in Botswana is called the Botswana Animal Identification and Traceability System (BAITS). However, at the inception of animal traceability systems in Botswana, the first system introduced was called the Livestock Identification Trace-back System (LITS). According to Oladele and Jood [22] LITS was introduced in Botswana to carry out the registration and identification of individual cattle in 2001. The sole purpose of developing LITS was to fulfil the mandated EU requirements for exportation by monitoring diseases and cattle traceability from farm to fork. LITS was developed to carry out the function of animal traceability. Vallat [6] posits that the use of LITS around the world has increased, mostly for the provision of trade, animal production, and public health interventions. Lavoie and Forest [7] explain that these types of systems are useful in providing feedback.

Animal traceability is used in Botswana mainly for identifying and registering cattle, issuing movement permits, observing and controlling disease, monitoring the cattle market, and partnering with international organizations. The identification and registration of cattle is conducted through the Ministry of Agriculture (MOA) and BAITS to track the cattle throughout their lifetime. This process involves farmers buying ear tags with registration numbers from MOA or Veterinary and tagging each cow. The ear tags are registered to the central database that is managed by MOA.

The movement of cattle within Botswana and across the border requires documentation and an approved permit from the local veterinary services. This is done to make sure the regulatory requirements for compliance and traceability are met. The farmer has to write a movement permit application to the veterinary containing information about the cattle location, destination, reason for cattle movement, and the farmer's profile. The application is recorded and processed in BAITS by the veterinary, who then issues an approval or denial to the farmer.

BAITS is useful in observing and controlling cattle disease across Botswana. One of the main aspects of animal traceability is preventing outbreaks of diseases like FMD. Power et al. [20] highlight the significant economic and social impacts the FMD had on Botswana in the trade industry throughout the years. He further states that the recurrence caused some farmers to illegally traffic cattle from South Africa. The challenge is that the FMD virus stays alive even after the cattle have been slaughtered and processed. The country has implemented multiple measures for prevention and controlling the spread of disease. The first is mandatory reports of diseases in cattle, campaigns that support vaccination, and protocols on quarantine. The data from BAITS will facilitate identifying and tracing the diseased cattle movement, promote interventions for the diseased cattle, and mitigate to prevent the risk of disease outbreaks.

To maintain the integrity of the market, promote the trade of cattle and cattle products internationally, and assure confidence to the consumers, Botswana monitors the market. The markets are checked for compliance with the safety and quality standards. BAITS is used to trace the history, origin, health status, treatment, and export certification of cattle from farmers. The animal traceability system is also used to strengthen and align with the best practices and global standards. It is done to maintain good collaboration with its international partners. Botswana is in partnership with regional economic communities and the World Organization for Animal Health. The partnership magnifies the ability of the country to respond to transboundary diseased animals, trade facilitation, and safeguarding the health of the public. Altogether, the combination of registration, identification, movement monitoring, track and trace, disease control, the integrity of the market maintained, and international trade facilitation is critical to animal traceability in Botswana.

1.3.1 Challenges Facing Botswana's Animal Traceability Systems

Effective animal traceability systems are essential for managing livestock health, ensuring food safety, and facilitating international trade. At the basic level, such systems require components like animal

identification, information flow management, and centralized databases to store and manage data [5]. Botswana employs the Botswana Animal Identification and Traceability System (BAITS), designed to enhance livestock traceability from birth to slaughter. It mandates the use of electronic and analogue ear tags, ensuring accurate identification and traceability of cattle while requiring farmers to report movements, departures, and mortalities to a centralized database within specified timelines [23].

LITS was replaced by BAITS to address the challenges presented by LITS. The main outcome was making the system less centered on the government and more farmer-centric, but none of the technical challenges were addressed [24]. Derah and Mokopasetso [25] acknowledge the fact that Botswana gained from adopting a computerized system at its inception. Even though the system was not able to be fully effective in managing the data. A study by Ndubo and Moreki attested to the fact that LITS, which is now known as BAITS, still has implementation challenges that need attention [26]. Results show that currently, BAITS is not enabling the exportation of beef to the EU and other countries by Botswana, as was intended at its inception [26].

Despite these measures, Botswana faces numerous challenges in maintaining effective traceability systems. These include limited interoperability, poor data integration, and compliance with regulatory standards. The reliance on open grazing systems complicates accurate tracking of individual cattle and contributes to low productivity and elevated mortality rates in communal grazing areas [27]. Furthermore, foot-and-mouth disease (FMD) outbreaks disrupt cattle movement and trade, undermining the effectiveness of existing traceability frameworks. The annual agricultural survey revealed that cattle losses from theft and straying remained high, emphasizing gaps in traceability systems [28].

An identifiable challenge with BAITS is the absence of interoperability and standardization with the other stakeholders, as well as the lack of semantic knowledge and data management. Cattle traceability in Botswana has a range of diverse stakeholders, including veterinary services, Botswana Meat Commission (BMC), law enforcement, farmers, slaughterhouses, Matimela (stray animals) officers, and the Ministry of Agriculture. This means that achieving seamless data exchange and coordination is increasingly complex. However, only veterinary services and the Ministry of Agriculture have access to the system and its data. For example, the government of Botswana still spends handsomely on the collection and maintenance of “Matimela” – stray animals in many kraals around the country.

This is to say, despite each animal being tagged, branded, and its data being stored on these systems, tracing and linking the animal with its owner still evades the government. This is supported by research conducted by Mooketsi, who states that the District Council officer at Matimela expressed concern for the lack of access and should be allowed to fully employ the system, instead of using separate systems [29]. He went further to contact the police officers, who also expressed the same concern for the lack of access [29]. Research findings of this study (Chapter 4) discovered that even BMC uses a separate system called an integrated system. The remarkable engagement that Botswana has with trading cattle cross-border also increases the urgency for animal traceability systems that are interoperable. Furthermore, the data from animal traceability is compromised by the absence of standardized interoperable protocols. Absenteeism is distressing to the privacy and security of data. The are five main stakeholders that are involved in animal traceability, the stakeholders along with their roles are listed in Table 1.

Stakeholder	Role
Farmer	Animal traceability helps a farmer monitor the movement, health, and origin of their livestock, enabling better disease control, record-keeping, and market access.
Veterinarian	Animal traceability enables a veterinarian to quickly access an animal's health history, movement records, and origin, facilitating accurate diagnosis, treatment, and disease outbreak management.
Ministry of Agriculture	Animal traceability allows the Ministry of Agriculture to monitor livestock movements, ensure food safety, control disease outbreaks, and support policy-making through reliable and timely data.
Local Administrative Authority (Matimela)	Animal traceability helps the Local Administrative Authority identify the owners of stray cattle, manage their return or relocation efficiently, and reduce conflicts and public safety risks.
Botswana Meat Commision	Animal traceability enables the Botswana Meat Commission to verify the origin, health status, and movement history of livestock, ensuring compliance with export standards and maintaining international market confidence.

Table 1 Animal Traceability Stakeholders and Their Roles.

The Performance of Veterinary Services (PVS) evaluation report by the World Organisation for Animal Health highlighted critical challenges in animal traceability that remain relevant more than 10 years later, such as inadequate infrastructure and resources, limited veterinary personnel, and high costs of physical barriers for disease management [30]. Addressing these challenges requires investments in modern traceability systems, enhanced disease control, and stakeholder coordination to improve the sector's competitiveness and market access.

To effectively address these challenges, Botswana must focus on implementing an interoperable animal traceability system and initiating interoperability standards. Ensuring public health, safety and trade facilitation at the same time, advancing economic growth and sustainable agricultural practices. The advancement in technology has emerged as a favourable technology for solving interoperability and standardization challenges in animal traceability. Ontologies are arguably the best way to manage data, in terms of information integration, interoperability, scalability, and knowledge management [31]. This emphasises that ontologies play a key role in the organization and interpretation of data that is complex across domains. They contribute a formal and shared conceptualization of a domain that allows a shared understanding to facilitate data integration, reuse, and exchange. Furthermore, ontologies provide good capabilities for data management in the context where scalability, semantic clarity, and integration are the main focus. However, they are effective depending on the domain requirements, resources available, and data. Ontologies can offer a robust solution for a specific context and cannot be seen as an approach that fits everything.

1.4 An Introduction to Ontologies

Ontology is defined by Gruber [32] as a set of representative primitives with which to model a domain of knowledge. These primitives are classes, attributes, and relationships [33], [34], [35], [36]. Brusa et al. [37] also defines ontology as a dictionary of terms that are formulated into a canonical syntax with

mutually acknowledged definitions, with the design to yield a taxonomical framework for the representation of knowledge shared by multiple communities of information systems. For the development of ontology, there is a need for ontology class definitions, taxonomic hierarchy arrangement of the classes that is the subclass and superclass, slots definition, a description of values allowed in the slots then filling the slots with the various values [38]. Ontologies are designed for the development of multiple applications that enable the reuse of knowledge from the created domain representation [39].

Brinkmann and Kvale [40] also, define ontology as a dictionary of terms that are formulated into a canonical syntax with mutually acknowledged definitions, with the design to yield a taxonomical framework for the representation of knowledge shared by multiple communities of information systems. For the development of ontology, there is a need for ontology class definitions, taxonomic hierarchy arrangement of the classes, which includes subclasses and superclasses. Attributes that are associated with a concept are traditionally called slots in a frame-based ontology, according to Nguyen [35] however, in the modern web-based ontology, the term property is used to define the attributes (data properties and object properties). Different types of development tools have been created to use in ontology development. Among them are Protégé, OntoEdit, Ontolingua, and Duet [35]. Nguyen [35] classifies ontologies based on the purpose they fulfil, which is in the assistance of human agents' communication, in the improvement of process and quality in engineering software systems and achieving interoperability as challenging as it may be.

1.4.1 Benefits of Ontologies

A description of the word ontology is a data schema that is a conceptual representation of specific domain entities with their relationships in terms of classes and relations [39]. Sequeda and Miranker agree that ontologies work as conceptual models with a description of the interested domain [41]. Ontologies use logic-based language for machine-readable representation, allowing formal definition of the semantics of ontology concepts [39]. The stored information in the ontology is easily interpreted and reasoned by the computer. Ontologies' main advantage is the provision of interoperability; they are the best to use to solve the lack of interoperability in animal traceability. Furthermore, the ability to provide semantic interoperability assists in multiple systems understanding information in a meaningful and accurate way even though the structure, vocabulary, and format in different. Berg explains that the focus of semantic interoperability is to ensure data is utilized and interpreted correctly between the systems, by using vocabularies that are standardized, accurate interpretation by data models, integrations, and data utilization between the various applications and domains [42]. Which will all be possible through the use of ontologies. Ontologies will promote efficiency, innovation, and flexible systems that allow the integration of new technology into existing systems.

The power of ontologies over other traditional data structures (e.g. databases) is inference; the ability to reason about data. Ontologies follow the open world assumption - the lack of knowledge does not automatically mean it is false. Therefore facts not explicity stated can still be true. While databases on the other hand follow the closed-world assumption – anything not explicitly stated to be true is false.

This research proposes an interoperable Ontology-Driven Framework for Animal Traceability (ODF-AT) as a guide to the development of effective and practical animal traceability platforms to solve the gaps of poor interoperability and standardization of vocabularies in current frameworks for animal traceability, with Botswana as a case study. Ontologies are said to be the best way to manage data, in terms of information integration and knowledge management [43]. They address both human and machine challenges since, by definition, they are machine-readable formal representations of a domain of discourse [39]. The most identifiable impact of such a framework is the interoperability of animal traceability stakeholders with an effective ontology knowledge management system. Also, the enabling of value derivation from animal traceability data will facilitate the reuse of the data and create new business opportunities. The opening up of such data would, for example, help drive a budding

innovative “apps economy” and make a significant contribution towards driving Botswana closer to its vision of a diversified economy – one that is knowledge-driven as opposed to one that is dependent on natural resources.

The technological infrastructure plays a crucial part in the success and proficiency of animal traceability systems. To achieve high-performing systems that yield desired results and gain high profits for farmers and countries, the first step is to evaluate and take action on the infrastructure. Research supports this by stating that a technology strategy of implementing good technology infrastructure brings significant and positive effects to the performance of the company, achieving highly valuable products [44]. This research develops a novel ontology from scratch for animal traceability. The development of the ontology for animal traceability was constructed using the ontology methodology called METHONTOLOGY. The choice for METHONTOLOGY was based on the comparison of ontology methodologies carried out [45]. The research showed that all development activities are supported by METHONTOLOGY and only partially supported by one, whereas there is less support for the ontology development activities by all the other methodologies.

According to Gómez-Pérez et al. and Khondoker, the development of ontologies is a time-consuming and complex task, and even more so when the developer implements it from scratch without a tool to support it [46], [47]. Ontology editor Protégé is described as a tool that permits users to define instances, concepts, and relations [34]. Alatrish says that Protégé is a user-friendly ontology development tool with advanced features for collaboration and cooperation [48]. It offers a customizable interface layout and visualization, allowing for viewing, zooming, and pruning of graphical taxonomy, unlike tools like Swoop, which lack a help function. Swoop, TopBraid Composer, OntoTrack, and protégé are among the six tools used in a study by Khondoker to develop ontologies, with Protégé being the most domain-independent and dominant tool [47]. This research chose the ontology development tool Protégé 5.1 for building the ontology using the ontology language Web Ontology Language (OWL). Heflin and Pan state that OWL is created for the support of the semantic web, and it provides exchanging and encoding ontologies as an international standard [49].

An Animal Traceability ontology was evaluated using two evaluation techniques. The first evaluation technique is called the OntoMetric evaluation tool. Lantow states that OntoMetric is used to evaluate the quality of the ontology through automatic metric calculations [50]. The OntoMetric tool provided multiple metrics and analysis techniques for evaluating different aspects of the Animal Traceability ontology. The aspects of the ontology that were computed include the ontology structure, coverage, and coherence. The OntoMetric tool is important in helping ontology users and developers understand the ontology's strengths and weaknesses and have informed ideas when making decisions on the improvement, use, and reuse. The second evaluation technique was conducted using Apache Jena Fuseki by running SPARQL Protocol and RDF Query Language (SPARQL) queries on the Animal Traceability ontology. It is important to run SPARQL queries for ontology evaluation because they assist in verifying the performance, correctness, and consistency of the Animal Traceability ontology. The verification process will ensure that the ontology meets the expected requirements to be used effectively in applications. Apache Jena Fuseki is used for its provision of a platform that is comprehensive for querying, managing, and ontology evaluation. Furthermore, it facilitates optimization and assessment for performance, scalability, and data quality of the ontology.

1.5 Problem Statement

This research aimed to address the existing gap in interoperable, knowledge-based frameworks for developing animal traceability systems and related platforms. The case of Botswana is used to provide an empirically validated framework based on stakeholder needs. The lack of interoperability has undermined the effectiveness of promising initiatives like LITS and BAITS. While these systems have succeeded in gathering animal traceability data, they did not offer a standardized and interoperable

framework that fully integrates all relevant stakeholders. Currently, these systems mainly serve only two categories of stakeholders (government and farmers), despite the involvement of at least five different categories of stakeholders, as demonstrated by our research, in the animal traceability process. There is a pressing need for a truly interoperable system that brings all stakeholders together within a unified environment, allowing them to access critical information more efficiently, without the constant need for veterinarian intervention for every cattle-related matter. Such a system should support reliable cattle identification through advanced tracking technologies that automatically feed data into the platform. It should also provide user-friendly interfaces in both Setswana and English, reducing reliance on technical terminology. Moreover, a centralized database would enhance the generation and updating of reports, significantly reducing processing times. By promoting knowledge management and information integration, this interoperable system would increase the value of collected data, leading to more efficient and effective animal traceability.

1.6 Objectives

To address the problem statement as stated above, this research has set out to achieve the following main and specific objectives:

1.6.1 Main Objective

The main objective of the research is to develop and evaluate an ontology-driven framework for animal traceability, contributing a novel approach to interoperable animal traceability systems design.

1.6.2 Specific Objective

1. To identify and analyze the limitations and challenges of current animal traceability frameworks and systems.

Outcome: Limitations of current animal traceability frameworks and systems.

2. To identify components for the design and development of the proposed interoperable framework for animal traceability.

Outcome: Framework components and ontology concepts through analysis of the data collected.

3. To design and develop an ontology-driven framework for animal traceability that enhances interoperability, integration, and knowledge management among stakeholders in Botswana's livestock sector.

Outcome: An ontology-driven framework as a guiding principle for the development of animal traceability systems.

4. To identify existing ontologies that can be reused, alternatively, in the absence of existing ontologies, identify an ontology methodology to develop a new ontology for animal traceability.

Outcome: Animal traceability ontology artefact for knowledge representation to facilitate interoperability and reusability effectively.

5. To evaluate the feasibility and effectiveness of the animal traceability ontology in addressing challenges related to disease control, food safety, and economic outcomes.

Outcome: Evaluation report on animal traceability ontology

1.7 Research Questions

1.7.1 Main Research Question

How can an interoperable ontology-driven framework be developed to enhance animal traceability across diverse systems and stakeholders?

1.7.2 Specific Research Questions

1. What are the perceived limitations and challenges of the current animal traceability systems according to the main stakeholders?
2. What are the key components of an interoperable and knowledge-driven framework for animal traceability?
3. How can stakeholder knowledge be effectively captured and modelled within an ontology without compromising technical interoperability, to develop an animal traceability framework?
4. Which ontology methodologies can be employed to align and integrate heterogeneous systems into a single interoperable animal traceability ontology?
5. What is the impact of the proposed ontology-driven framework on stakeholders' interoperability, knowledge management, and integration?

1.8 Significance of the Study

Botswana's livestock sector struggles with implementing effective animal traceability systems due to limited interoperability, inadequate data integration, reliance on open grazing systems, and systemic challenges like high mortality, theft, and disease outbreaks, which collectively undermine its competitiveness and compliance with international trade standards. In response to these critical issues, this research developed an ODF-AT specifically tailored to the needs of Botswana's livestock industry. By addressing the limitations of existing systems, the ODF-AT enhances interoperability, standardizes vocabularies, and improves data integration, offering a robust framework to tackle the challenges posed by open grazing systems and systemic inefficiencies. This ontology-driven approach supports better interoperability and knowledge management among stakeholders, strengthens disease management efforts, and ensures compliance with stringent international trade standards, positioning Botswana’s livestock sector for improved competitiveness and sustainable growth.

This research will provide a framework with identifiable components that can be further translated into programmable components of an animal traceability system in a given context, which can be extended to other developing countries. It is structured into four layers: an input layer for data collection, a semantic core layer for query processing, a knowledge management layer serving as a data hub, and an application layer for user interaction. This framework addresses gaps in scalability, data privacy, and standardization, aiming to improve disease control, food safety, and economic outcomes by providing a comprehensive animal traceability system.

It is important to keep finding ways to automate the manual work that is being done in our everyday lives. With technologies like ontologies, we can take simple information and make it more manageable to make the flow of work easier. Ontologies contribute to data modelling through structured knowledge representation, ensuring clarity and consistency across the different data sources, systems can meaningfully understand and exchange data, and a common framework is provided that maps data formats that are disparate into a coherent whole, allowing cross-platform querying. The contribution of ontologies to application management includes changes being made without major reprogramming to applications, rule-based reasoning permits applications to infer knowledge that is new automatically, standardized queries are used for data interaction, making the management and maintenance easier, and

lastly data handling rules in ontologies are modelled and enforced clearly across the applications. This is what we expect from the ODF-AT. Integration of the stakeholders and an interoperable system using OWL from the framework will ensure effective and dependable animal traceability systems.

The research will close the gap of the lack of a framework for the design and development of an animal traceability system by using ontologies to make sure that the data used in the systems meets ontology requirements of interoperability, integration, and data reuse for producing good-quality and fit data. Data reuse will ensure the longevity of the data and vocabulary used in the framework for developers and researchers to understand and continue the research.

The framework uses specific ontologies following certain common patterns, and most importantly, machine and human-readable language. This is so that the application is reusable in the world of research and development. The adoption of the framework by researchers to further expand and innovate in their area of interest will significantly add value to the body of knowledge. The developed animal traceability ontology in Chapter 5 of this research has been published in Figshare.com for developers and researchers to freely download and reuse.

The development of the framework is expected to also draw from growing disciplines such as open data and the semantic web, particularly from ontologies. The research will also culminate in the development of an animal traceability ontology as a means of evaluating the framework. We expect a tremendous impact from the framework and subsequent systems developed following the framework. This will be a catalyst for innovation and the creation of new technological products in the animal traceability industry. While this will also contribute towards Botswana getting closer to the vision of a knowledge-based economy, it will also facilitate Botswana's conformance to market (EU) requirements.

The most identifiable impact of the proposed framework is the enabling of value derivation from animal traceability data, which will facilitate the reuse of the data and create new business opportunities. The opening up of such data would, for example, help drive a budding innovative "apps economy" and make a significant contribution towards driving Botswana closer to its vision of a diversified economy – one that is knowledge-driven as opposed to one that is dependent on natural resources.

1.9 Contributions to this study

The main contribution of this study is the proposed ontology-driven framework for animal traceability and the developed animal traceability ontology. The animal traceability ontology is published online on Figshare. The ontology evaluation report based on two types of evaluation techniques is presented in Chapter 6 of the thesis. The multiple contributions of this research are further discussed in this section.

1.9.1 Main Contributions

The ODF-AT is presented as a guiding principle for developing interoperable knowledge-driven animal traceability systems. The framework presents a structured method for developing domain-specific animal traceability platforms. A novel and open-source animal traceability ontology has been developed for animal traceability researchers and developers. The ontology is available for free download to be reused for further research on Figshare. (https://figshare.com/articles/dataset/Animal_Traceability_Ontology_ATO_for_Botswana/26826700?file=48765313). A step-by-step illustration of the development of animal traceability systems providing main concepts, data, relationships, classes, sub-classes, and axioms.

1.9.1.1 Metrics-based Contributions

The contributions were analyzed and presented based on the contribution metrics, which state that a thesis/dissertation should answer the following questions to validate the contribution to research. What? When? How? and Why? [51], [52].

What? - According to Whetten [52], the “What” question in a study addresses both what is new and why it matters. This study proposed an ODF-AT and developed a new animal traceability ontology for the development of animal traceability platforms (systems, applications, websites). The recent conferences that addressed Animal welfare were in 2016 [53]. Despite all the efforts that have been made by different organizations there are still major animal traceability challenges that have not been successfully addressed. The main focus of the organizations is on devices to monitor the animals and less on the management and value derivation of the animal data collected. ODF-AT introduces a new aspect of managing and deriving value from knowledge-based information, focusing on the data about the animals. The biggest value-added contribution is interoperability and integration, which will be initiated by the implementation of ontologies. This will take the stakeholders from having heterogeneous systems to an integrated and interoperable system that deals with “cattle”, changing and improving the current organizational operations.

When? – The When? is based on finding the balance between parsimony and comprehensiveness [51]. That means that the study should find the simplest form of explanation for the evidence/solution without losing its effect or phenomena. Knowing when to add concepts, details, and variables that are understandable by people who are not specialists in the field of computer science. This study contributed to the body of knowledge as a whole and not just the computer science field. Firstly, it brought a solution to a real-life practical problem in the society, in the agricultural field in Botswana, particularly animal traceability. The data collected and the results presented are beneficial to understanding cattle theft, diseases, feeding, vaccination, computer literacy, farm employment, network connectivity in the farm areas, and the differences in the types of livestock farming (communal and commercial) challenges. The technicality of the ODF-AT proposed is broken down into details that researchers and developers can understand. This study went further by developing the animal traceability ontology with classes, data properties, objects, and individuals. The ontology depicts the contents, flow, rules, and variables of how animal traceability systems should look.

How? – addresses How robust is it? How far do variables affect each other? The animal traceability framework needs to guide in development of robust systems [54]. Denkena et al. [55] define robustness as a feature of a system that will make it possible to maintain the state planned even when the system's behavior, components, and environment change. Zhang [56] agrees that robustness is a form of quality assurance for software that is analyzed rigorously and continuously designed. The evaluation of Chapter 6 of this study contributes highly to the robustness and variable linkages of the animal traceability ontology. The metrics calculations and discussed results illustrate to developers and researchers the strengths and improvements that are necessary when developing animal traceability systems. According to Zhou [51] for a significant contribution, adopting a diagram that is designed with arrows and boxes showing relationships between the variables is necessary. The proposed ontology-driven framework for animal traceability in this study is presented in the form of a diagram in Chapter 5. The diagram is a form of boxes and arrows that shows a relationship among the variables. The diagram is a guiding principle for developers, showing the process that has been established in theory. Answering both the why? and how? factors of theory.

Why? – Contribution is achieved if the researcher can answer Why so? and Why now? [52]. The topic went through rigorous research through literature review and data collection to make sure the gap existed and it is still relevant. Chapter 2 of this study discusses what was found in the literature about animal traceability, frameworks, and ontologies, presenting and analyzing the identified gaps. This study was conducted in the case of Botswana, which meant that the location for data collection was Botswana. Chapter 3 presented the analysis of the data collected, and Chapter 4 discussed the findings from the data. These chapters present arguments showing the gap and supporting the need for the proposed framework for the five major animal traceability stakeholders in Botswana. The integration of the stakeholders and an interoperable system is the major contribution that will bring overdue, effective change. The topic is very contemporary because the current animal traceability system,

BAITS, has recently been in the news constantly for the failures that have slowed down monetary production for farmers and the country. The Matimela issue is still ongoing for the government, where farmers are losing cattle every day, but the cattle found by the council still get sold as they cannot find the owners. Police and BMC still have to go to the veterinary to get information on the cattle of farmers. All these stakeholders are working on cattle and need similar information, but there is currently no system in place that integrates all of them to work and communicate as one. The proposed ODF-AT can be used by system developers to easily bridge this gap.

1.10 Thesis Structure

The rest of the thesis is organized as follows:

- **Chapter 2** presents the literature review. This presents a synthesis of current literature on the areas of animal traceability technologies and ontologies, intending to perform a gap analysis and inform the design of the framework.
- **Chapter 3** presents a description of the research methods and data analysis approaches followed in collecting and analyzing the data that is needed for the design of the proposed framework and ontology.
- **Chapter 4** provides an analysis of the data collected in Chapter 3. The concepts from this analysis will form a basis for the design of the proposed framework and ontology.
- **Chapter 5** provides a detailed description of the proposed framework. The chapter also describes the design of the animal traceability ontology.
- **Chapter 6** provides a detailed evaluation of animal traceability ontology.
- **Chapter 7** concludes the thesis and provides recommendations and limitations of the work.

2 CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

The purpose of this chapter is to synthesize the current literature in animal traceability technologies and ontologies, intending to perform a gap analysis and also inform the design of the framework. It is important to investigate the research that has been done in the three main areas that are involved in designing the framework, which are animal traceability, ontology, and existing frameworks in animal traceability and ontologies. Much research has been done on ontology and animal traceability individually, but not together to produce an effective framework. Merging ontology, animal traceability, and frameworks will produce advanced knowledge representation, interoperability, and information integration with quality reusable data. This section will cover what has been done in these areas and the relations they have with each other that make them good to merge into one framework to produce good-quality systems. The purpose is to perform a gap analysis to justify this research. The purpose is also to review methods, techniques, metrics and measures that are relevant to this work to justify the selection of those appropriate for this research.

2.2 Animal Traceability

There are various types of systems and tools encompassed as animal traceability technologies that are used for monitoring and tracing animals during their lives. The technologies contribute to the safety of food, management of the supply chain, and control of disease, enabled by the technology's ability to identify, manage, and monitor the individual animals. Animal traceability technology examples include a Global Positioning System (GPS), ear tags, smart sensors, Radio Frequency Identification (RFID), blockchain, microchips, DNA barcoding, and biometric recognition.

RFID is the most commonly used animal traceability technology [57]. RFID, among the other technologies, stands out because it is reliable, compliant with industry standards, continuously improving technologically, can easily integrate with other systems, and presents opportunities for innovation.

RFID is a widely used technology system for animal traceability, which has emerged, shows great potential. It is a technology for identification that is contactless, with two main components being the interrogator and transponder located on the object [57]. It is said to be an automatic technology system that enables machines the power to identify and sense objects with the use of radio waves, and then record that metadata [58]. The objects here, being cattle, are attached with tags called RFID tags on a part of their body, usually the ear. The RFID tag acts as a reading device that reads, identifies, monitors, and tracks the object. The data is automatically read and updated.

2.2.1 Identified uses and gaps of Animal Traceability Technologies

Table 2 presents a summary of the technologies identified in research and the gaps identified. The paper by Resti et al. [17] discusses various technologies used in dairy recording tools, including mobile phones, applications, desktop applications, web applications, networks, and the Internet of Things (IoT). Mobile phone applications are identified as the most widely used due to their accessibility and ease of use, especially in regions with limited infrastructure. IoT has the potential to automate data collection and provide real-time information, enhancing farm management efficiency. The integration of IoT with other technologies, such as networks and mobile apps, is mentioned as a way to improve the user-friendliness and effectiveness of dairy recording tools. However, challenges related to Internet connectivity and the need for continuous data input are acknowledged as critical factors influencing the successful implementation of these technologies in developing countries.

No.	Technology	Main Components	Gaps Identified
1	RFID-based methods integrated with infrared sensors and wireless sensor networks. [59]	For livestock management, animal tracking, traceability, and animal behavior monitoring.	Does not propose any new animal traceability technologies but instead reviews the existing methods and their applications in different animal tracking scenarios
2	Technologies used in dairy recording tools include mobile phone applications, desktop applications, web applications, networks, and IoT. [17]	Integration of IoT with other technologies, such as networks and mobile apps, is mentioned as a way to improve the user-friendliness and effectiveness of dairy recording tools.	Challenges related to Internet connectivity and the need for continuous data input are acknowledged as critical factors influencing the successful implementation of these technologies in developing countries
3	Potential application of blockchain technology in animal product traceability systems. [60]	Provides a decentralized, secure, and transparent method of recording and sharing data across the entire production chain, from farm to consumer.	Does not propose any new blockchain technologies; rather, it compiles existing knowledge and examples from the literature to illustrate the benefits and challenges associated with adopting blockchain for animal traceability
4	Animal Traceability System for Botswana. [61]	Developed to replace the BAITS system in Botswana. The main aim was to provide objective performance information for the beef industry for the improvement of economic and genetic beef production efficiency. 3 specific objectives. 1. On-the-farm and on-the-station evaluations 2. Quality of the carcass evaluation 3. The facilitation of the implementation of a new act.	Absence of interoperability. Absence of any standardization. No suitable structures to handle and weigh cattle at the farm. Lack of record-keeping. Lack of animal identification by farmers. Farmers did not comprehend the benefits of record keeping. Lack of data collection at the farm, and no manpower. No database was developed to manage the data. Farmer's participation declined during schema trials. The switch from BAITS to LITS only accommodated one stakeholder. Still have identification problems with ear tags.
5	Cloud-based livestock monitoring system integrated with Blockchain technology and RFID sensors. [62]	Early detection and monitoring for abnormal behavior and infections	Technology integration challenges. Lack of information dissemination and data fusion.
6	Internet of Things. [63]	Real-time observation of an individual animal's production, health, environmental effects, and welfare. Detection of small significant change patterns of behaviour in cattle.	Cannot be used in farms with no network or internet connectivity. The data recorded focuses only on the health of the animal.
7	Pasture-based livestock system. [64]	Monitoring the behaviour of cattle and pasture management.	The low battery lifespan of the device, it would not be able to provide long-term functionality. The management of data in the open fields is not effective. Farms with no

			signal reception for the internet cannot work with the technology.
8	Proposed ontology tool. [19]	Integration of heterogeneous databases. Facilitation of interoperability	Unreliable because not been developed and tested yet.
9	Proposed Machine learning tool. [65]	Trace elements for animal fingerprints and classification models obtained as tools for the establishment of reliable traceability systems.	It is a concept that has not been implemented to test its functionality and effectiveness.
10	Proposed face recognition tool. [66]	Curb disease outbreaks. Improve false insurance claims in agriculture. Uses image analysis technology	It is a proposed idea that has not been developed to test the accuracy and reliability.

Table 2 Animal traceability technologies.

Pereira et al. [59] identified various RFID-based methods used for livestock management, animal tracking, traceability, and animal behavior monitoring. Passive UHF RFID tags are the most common, due to their cost-effectiveness and ease of use in tracking cattle and other mammals. RFID is often integrated with other technologies like GPS and cameras to enhance tracking accuracy and provide real-time location data, which is particularly valuable for monitoring animal movement and behavior in large or remote areas. Other integrations include infrared sensors and wireless sensor networks, which help gather additional health and environmental data, facilitating comprehensive monitoring systems. The paper does not propose any new animal traceability technologies but instead reviews the existing methods and their applications in different animal tracking scenarios.

A review paper by Zanetoni et al. [60] investigates the potential application of blockchain technology in animal product traceability systems, highlighting its emerging role in ensuring food safety and quality. Zanetoni et al. [60] emphasize that blockchain can provide a decentralized, secure, and transparent method of recording and sharing data across the entire production chain, from farm to consumer. By doing so, blockchain addresses key concerns such as traceability, transparency, and consumer trust, which are increasingly important in food supply chains. The review identifies various use cases where blockchain has been implemented or proposed, such as in the tracking of beef, milk, eggs, fish, and pork, showing its versatility across different types of animal products. However, this review paper does not propose any new blockchain technologies; rather, it compiles existing knowledge and examples from the literature to illustrate the benefits and challenges associated with adopting blockchain for animal traceability.

Modisa [61] discovered that the traceability system in Botswana was not effective and not beneficial to all the stakeholders involved. The traceability system was initiated by the government of Botswana in 2001 as the LITS system, which was only done to facilitate the new EU import regulations. LITS had too many challenges being fully funded by the government, other stakeholders having a limited role, the bolus for identification not being visible, the updating of the central database being prolonged, and bolus recycling causing inaccurate data [61]. The government of Botswana then revised LITS with BAITS. BAITS is supposed to be more farmer-centric, with better monitoring of animal and public health, costs not fully dependent on the government, replaces bolus with ear tags, farmers buy ear tags, and also bear the costs [61]. A proposal by [61] is for a national beef performance testing and recording schema as a better option for animal traceability in Botswana. The national schema proposed is based on three objectives: On-the-farm and on-the-station evaluations, Quality of the carcass evaluation, and the facilitation of the implementation of a new act. Trials of the schema were carried out, but were not successful. There were no suitable structures to handle and weigh cattle at the farm. Farmers had no record-keeping methods, so no history to follow. Lack of animal identification by farmers, most of

whom did not have ear tags. Farmers did not comprehend the benefits of record keeping. The scheme had lack of data collection at the farm because there was no manpower. No database has been developed to manage the data he collected. Farmer's participation declined during schema trials [61]. The schema ended at the research phase.

The prevalent danger from livestock infectious diseases could cripple economies, not to mention the uncertainty of food supply for the world [62], [67]. An approach that has shown promising results is cattle traceability, being able to monitor and track cattle from birth to fork will help in identifying diseases early and acting at early stages. Yang et al. [62] say that the existing systems have so far not been able to fully address the challenges, although RFID is widely used and promises good results. The systems give information about cattle that is basic and are vulnerable to hackers. The recommendation [62] is to use a cloud-based livestock monitoring system with a combination of Blockchain technology and RFID sensors, for humidity, movement, and carbon dioxide. The system will be able to detect and monitor cattle for abnormal behaviour that is infection-related early. Blockchain and cloud computing provide communication that is secured through a cryptographic blockchain system with the cloud server.

Yang et al. [62] believes the solution is Blockchain, but [63] Precision Livestock Farming (PLF) found that adopting not just one technology but multiple, like the Internet of Things, was a better approach for animal traceability. PLF is explicated as the unceasing real-time observation of an individual animal's production, health, environmental effects, and welfare [63]. The PLF has, over the years, advanced its technologies borrowing from the fields like Internet of Things, internet access, information and communication, and wireless communication. It supports farmers with non-intrusive and objective, endless data collection [63]. The Authors further state that it provides the ability to detect small significant change patterns of behaviour in cattle, which assists the decision management of farmers, reduces in workload, and less theft of cattle.

A review of PLF focused on pasture-based livestock systems in monitoring the behaviour of cattle and pasture management [64]. The review showed great strides in using PLF for pasture management, which agrees with the review from [63], [68]. Like the other traceability tools, it is not without its challenges. The first challenge is the battery lifespan of the device, which would not be able to provide long-term functionality. The management of data in the open fields is not effective. Lastly, farms with no signal reception for the internet present a problem with some of the technology used [63].

The traceability system's key factors, as delineated, include the ability of information sharing and integration in the supply chain [19]. Magliulo et al. [19] agrees that Bailey et al. [64] traceability systems have great benefits to farmers and consumers. For example, the ease and cheaper crisis management for businesses, but there are challenges present. Just like Cao et al. [8], and Magliulo et al. [19] found that the current systems have challenges to overcome. Magliulo et al. [19] proposed an ontology tool to overcome the integration challenges, dissemination of information, and data fusion.

An ontology tool is presented by Magliulo et al. [19] as what He believes is the best solution for the traceability challenges. Ontologies are a favourable choice because of their effective techniques in information modelling and management [19]. The traits of information modelling and management of ontologies were applied to the dairy industry. The ontologies are independent of the data model's lower level, which provides for the integration of heterogeneous databases. This also allowed the facilitation of interoperability between the disparate systems. Furthermore, providing services that are knowledge-based [19]. However, the ontology tool is a proposal, and no implementation has been carried out.

Machine learning is one of the technologies found could be feasible in the success of animal traceability [65]. Fernandes et al. [65] proposed the use of trace elements for animal fingerprints, and classification models obtained as tools for the establishment of reliable traceability systems. The classification models will be modelled from the different machine-learning approaches. Animal traceability is a domain that

still has so many challenges that many researchers are adopting what seems like any feasible approach. Xu et al. [66] and Kumar et al. [69] have proposed the use of animal face recognition for the first step of traceability (animal identification). They believe individual animal identification is a major factor in traceability to curb disease outbreaks and reduce false insurance claims in agriculture. Image analysis technology is the approach proposed by Xu et al. [66] to carry out facial expression and identification of cattle. His paper evaluated the possibility of what they call "a cutting-edge object detection algorithm" [66]. Kumar et al. [69] present a multi-resolution algorithm for cattle face recognition. The algorithm will use Local Binary Patterns and Speeded Up Robust Features provided by the Gaussian pyramid levels to perform the facial recognition. From the tests they performed, the algorithm showed an accuracy of identification of 92.5% of cattle faces database from the 1200 images of cattle that were collected [69]. Both approaches are promising great results in solving cattle identification challenges, but the true results will show when they are put to the test in real-world scenarios.

Juan and González-Vaqué [70] suggest the use of blockchain technology for enhancing animal traceability within the food supply chain by providing a secure and transparent way to record and verify each step in the production and distribution processes. Blockchain's decentralized ledger system ensures that data related to animal origins, health records, and movement is tamper-proof and easily accessible to all stakeholders, thus improving food safety and quality control. By enabling real-time tracking of animals from farm to fork, blockchain helps meet regulatory requirements and consumer demands for transparency in the supply chain. The technology's capacity to create a permanent and verifiable record of transactions makes it a reliable tool for preventing fraud and ensuring compliance with food safety standards. Overall, blockchain offers an efficient, scalable, and trustworthy solution for enhancing traceability and accountability in the animal agriculture sector.

García-Infante et al. [71] investigated the effectiveness of machine learning algorithms as tools for meat traceability, specifically to classify Spanish Mediterranean lamb carcasses based on production systems. The study evaluated six different algorithms, namely, artificial neural networks (ANNs), decision trees, k-nearest neighbors (KNN), naive Bayes, multinomial logistic regression, and support vector machine (SVM), using datasets that include organoleptic, sensory, and nutritional traits of lamb meat. The ANN and SVM algorithms demonstrated the highest accuracy in categorizing lamb production systems, with the ANN achieving accuracy scores of up to 0.88. In contrast, the KNN algorithm showed the lowest performance, with accuracy scores of 0.54 or lower. The study highlights the potential of machine learning tools to enhance the traceability and classification of meat products, suggesting that their use can improve transparency and consumer confidence in the meat industry. Using machine learning algorithms to enhance meat traceability aligns with our research's goal of improving animal traceability systems by demonstrating the importance of advanced technological tools for ensuring food safety, quality, and consumer trust.

Gbashi and Njobeh [72] explored the potential application of artificial intelligence (AI) and machine learning (ML) for enhancing food integrity and, specifically, animal traceability. The authors examine how AI and ML can automate and improve the accuracy of tracking and monitoring livestock movements, health status, and environmental conditions. The review highlights various AI and ML techniques, such as predictive analytics, image recognition, and anomaly detection, that could optimize data collection and decision-making processes in animal traceability. The paper also discusses the integration of these technologies with existing systems to enhance data management, provide real-time insights, and improve food safety and disease outbreak response. do not develop any new technologies for animal traceability based on AI and ML; instead, they focus on reviewing existing research.

Navia et al. [73] reviewed various sensor technologies used for monitoring the vital signs of livestock, emphasizing applications that enhance animal health and welfare. The study identified long-range communication technologies such as LoRaWAN as the most common for transmitting data from sensors monitoring parameters like movement, geo-location, body temperature, and heart rate. Short-range

technologies like Bluetooth and WiFi are also used, often in combination with long-range systems for better coverage. The paper suggests opportunities for developing more comprehensive sensor solutions that enhance traceability by providing continuous location data and robust communication capabilities. It also emphasizes the need for integrating energy-harvesting technologies to improve the autonomy and sustainability of these traceability systems.

León et al. [74] evaluated the effectiveness of handheld Near-InfraRed Spectroscopy (NIRS) as a non-destructive tool for traceability and authentication of the ageing process in Angus beef steaks. Using partial least squares-discriminant analysis, the study demonstrates high accuracy in distinguishing between aged and non-aged beef and predicting refrigeration storage times, with over 90% accuracy in external validation. This research suggests that NIRS technology could be valuable in digital transformation strategies for meat supply chain traceability, ensuring product authenticity and enhancing consumer trust. The paper highlights the importance of advanced, non-destructive technologies for ensuring authenticity and enhancing traceability, which aligns with our goal of developing interoperable and standardized traceability systems for the livestock sector.

2.2.2 Animal Traceability Frameworks

Füzesi et al. [75] proposed a framework for animal traceability. The study introduces the concept of a Digital Business Ecosystem (DBE) as a potential solution to improve traceability and information exchange in the meat industry. The DBE framework is designed to create an Internet-based environment where businesses can efficiently interact, share data, and implement traceability methods. This approach emphasizes using modern technologies like barcodes, RFID, and XML-based information integration to ensure continuous and reliable traceability across different stages of the meat production and supply chain. The DBE framework aims to address interoperability challenges and promote standardized practices, which are essential for effective animal traceability.

The framework for animal traceability is structured around key components, including the identification of traceable resource units (TRUs), a centralized database to store and manage data related to these TRUs, and an information flow mechanism that facilitates the collection, processing, and inquiry of traceability information [5]. The framework aims to ensure effective tracking of animals from their origin to the final product, enhancing transparency and compliance with health and safety regulations.

Marinello et al. [76] presented a framework for a traceability system specifically designed for the animal product supply chain using blockchain technology. This framework aims to enhance the transparency, security, and efficiency of tracking animal products from their origin to the end consumer. By utilizing blockchain's decentralized and immutable nature, the proposed framework ensures that data related to animal health, feed, and movement is securely recorded and easily accessible by all stakeholders in the supply chain, including farmers, regulators, and consumers. This approach is intended to improve food safety, comply with regulatory requirements, and increase consumer trust by providing verifiable and tamper-proof information.

Addo-Tenkorang et al. [58] proposed a conceptual framework for advanced animal traceability using Internet of Things (IoT) technologies to enhance supply chain management. The framework integrates RFID technology, solar-powered electronic tags, GPRS, GSM, and Google Earth to track animal health, birth records, ownership history, and location. By employing a cloud-based architecture, the framework aims to provide a low-cost solution for farmers, addressing vulnerabilities such as stock theft, straying animals, and inadequate health management. This IoT-driven approach promises to improve visibility and data transmission across the supply chain, contributing to better management practices in animal husbandry. The proposed system highlights the potential for digitization to add value to agricultural businesses, especially in free-ranging environments in Africa.

Chen et al. [77] presented a mobile solution for pork traceability using 2D barcode technology, focusing on improving food safety and quality in China's pork supply chain. The system enables the collection, transformation, and delivery of traceability information from breeding to retail, integrating seamlessly with government supervision systems. It offers a cost-effective alternative to RFID by utilizing batch-based traceability rather than individual animal tagging, making it accessible for smaller-scale operations common in China. The implementation in Jilin Province demonstrated enhanced transparency and credibility of pork traceability, promoting consumer confidence and regulatory compliance. This approach is relevant to our research on animal traceability technologies as it illustrates the successful application of mobile and barcode technologies to create an effective and scalable traceability system.

A framework for animal identification and traceability using AI-based muzzle recognition technology utilizes deep learning algorithms and convolutional neural networks to analyze unique muzzle patterns of bovine animals, providing an accurate and cost-effective alternative to traditional methods like ear tags [78]. Key components of the framework include muzzle detection, face alignment, feature extraction, and pattern recognition, which collectively enable the precise identification of animals. The system integrates a mobile application interface for capturing and uploading images, along with a MySQL database to store image links and identification numbers, ensuring efficient data management and verification. By leveraging open-source technologies and ensuring cross-browser compatibility, the proposed framework offers a scalable and field-friendly solution for enhancing animal traceability in various applications such as farm management, disease control, and livestock insurance.

A "butcher-to-farm" traceability framework designed for smallholder pig systems in Western Kenya aims to enhance disease control, food safety, and production efficiency [15]. This framework integrates the use of coded plastic ear tags to uniquely identify pigs, with records of pig movements, health events, and sales maintained by farmers, traders, and slaughterhouse operators. The system integrates data capture and database management technologies to ensure accurate tracking from the farm through to the slaughterhouse and butcher shops. It emphasizes the importance of involving local veterinary authorities and stakeholders to implement and maintain the system effectively. The traceability framework is intended to facilitate early detection and containment of diseases like African swine fever and cysticercosis, thereby improving public health and potentially opening new market opportunities for smallholder farmers.

A conceptual framework for animal traceability, suggested by Mwanga et al. [11] as a result of a literature review, it integrates various components, including electronic record-keeping, RFID technology, and mobile devices, to track and manage livestock data effectively. It involves a centralized system that allows data collection at the farm level, which is then transmitted to local and national databases for comprehensive tracking. RFID tags and sensors are used to monitor animal movements, health, and other critical information, enhancing traceability from farm to fork. The framework also incorporates mobile applications for farmers to input and access data easily, facilitating real-time decision-making and compliance with international traceability standards. This integrated approach aims to improve disease control, food safety, and marketability of livestock products by ensuring accurate and reliable animal identification and record-keeping throughout the supply chain.

Tripoli and Schmidhuber [79] introduced a framework for optimizing animal traceability using digital technologies like IoT, blockchain, AI, and big data analytics to enhance efficiency, accuracy, and security. The framework includes electronic identification (EID) methods and digital ledger technologies (DLTs) to securely store and share data across the supply chain. AI and predictive analytics are used for monitoring animal health, detecting diseases, and ensuring compliance with food safety standards. The framework also emphasizes the need for supportive policies, infrastructure, and capacity building to facilitate the adoption of these technologies. Together, these components aim to improve animal health management, ensure food safety, simplify trade, and raise consumer awareness.

Arvana et al. [80] proposed a blockchain-based framework for animal traceability in the agri-food sector, focusing on the production of Portuguese hams. The framework utilizes a multi-layer architecture that includes a business layer to define the supply chain processes, an integration layer for data collection and sharing, a cloud computing layer incorporating blockchain technology for decentralized and immutable data storage, and an application layer for user interaction. Smart contracts are employed to automate the recording and validation of transactions, ensuring data reliability and transparency throughout the entire product lifecycle. The system enhances traceability by capturing detailed information at each stage, from the animal's origin to the final consumer, making the process more transparent and trustworthy. This approach addresses the limitations of traditional centralized traceability systems, reducing the risk of data tampering and increasing consumer confidence in the quality and safety of meat products.

A framework for integrating Digital Twins (DT) of livestock production into a unified digital platform for agricultural management is proposed by Medennikov [81]. This framework emphasizes using digital twins to model and monitor various aspects of livestock farming, such as feeding, milking, and breeding, ensuring precise traceability and management. The DT framework includes cloud-based sub-platforms for collecting and storing primary data, technological databases for managing farming operations, and a knowledge base for algorithmic management and decision-making. By integrating these components, the framework aims to create a holistic system that connects livestock production with other agricultural sectors like crop production and processing. The unified approach seeks to improve sustainability, enhance breeding practices, and meet global standards for food safety and environmental protection

A framework for integrating blockchain technology into animal healthcare by Walunj and Gourkar [82] utilizes a permissioned blockchain platform, specifically Hyperledger Fabric, to manage various aspects of animal health data securely and transparently. The proposed framework involves a system design that includes multiple nodes representing authorized entities, such as veterinary clinics, adoption centers, and emergency services, with each node authenticated through a Membership Service Provider (MSP). Smart contracts are utilized within this framework to enforce business rules and access control policies, ensuring that only authorized personnel can update or view specific types of data. This approach aims to enhance privacy, trust, and the traceability of animal health records, contributing to better management of animal healthcare.

Witt et al. [83] proposed a framework to validate the use of slaughterhouse indicators as proxies for on-farm animal welfare and health assessments in pigs. It integrates data collected from 12 on-farm and seven slaughterhouse indicators, combining them into three health indices: limb health, other organ health, and respiratory health. The study examines the agreement between these indices across different production stages, such as farrowing, rearing, and fattening, to determine the feasibility of using slaughterhouse findings as retrospective indicators of on-farm conditions. The research shows that while some slaughterhouse indicators can partially substitute for on-farm assessments, complete replacement is not feasible, and agreement varies with the timing of assessments. Although this framework does not currently incorporate ontologies, an ontology-driven approach, as proposed in our current study, could significantly enhance data integration, consistency, and semantic interoperability, leading to more robust and scalable animal welfare monitoring systems.

2.2.3 Animal Traceability in Botswana

2.2.3.1 Livestock Production in Botswana

Agriculture is a vital component of livelihood for many people in Botswana, even though it contributes only around 2% to the country's GDP [84]. The sector is primarily dominated by livestock production, especially cattle, which constitutes roughly 80% of the agricultural GDP [84]. Botswana’s beef heritage is deeply intertwined with its socio-economic and cultural history, highlighting the importance of cattle as both a livelihood and a symbol of identity [85]. The country's free-range, grass-fed beef is known for

its high quality and unique flavor [86]. Botswana cattle production systems employ technologies like electronic rumen boluses for traceability to ensure compliance with stringent international market standards.

The Ministry of Agricultural Development and Food Security, commonly known as the Ministry of Agriculture, is tasked with enhancing food security and fostering agricultural development in Botswana [87]. The Ministry's efforts span several departments and parastatals, aiming to diversify the agricultural sector, reduce reliance on imports, enhance value addition, and promote sustainable practices. Specifically, the Department of Animal Production (DAP) seeks to enhance the competitiveness of Botswana's livestock sector by reducing reliance on imports, transferring skills and technologies, providing expert advice, and supporting both subsistence and commercial farmers in improving productivity and sustainability. A parastatal organization, the Botswana Meat Commission (BMC), established in 1965, aims to sustainably produce premium meat for the global market [86].

2.2.3.2 Development Challenges Facing Botswana

The sharp decline in Botswana's cattle population from 2010 to 2019 is attributed to poor farm management, low birth rates, high mortality, and theft [88]. Climate change-induced droughts, like the severe 2015/2016 drought, significantly worsened these challenges, impacting both livestock numbers and water resources. Additionally, the number of cattle holdings dropped drastically, with over 62% of farmers leaving the sector, reflecting systemic issues in the sustainability of livestock farming in Botswana. These challenges highlight the critical need for improved management practices, the adoption of climate-resilient strategies, and enhanced traceability systems to stabilize and grow the livestock industry.

The 2023 report by FAO includes statistical data demonstrating that livestock production in Botswana significantly contributes to hidden costs in agrifood systems, particularly through environmental impacts such as greenhouse gas emissions and land degradation, reflected in the high climate costs [89]. Limited access to markets and technological adoption also exacerbate inefficiencies, as does the lack of secure land tenure in communal grazing areas, discouraging investment in improved livestock management practices [90].

2.2.3.3 Barriers to Effective Animal Traceability in Botswana

Animal traceability is being done in the form of devices like RFID, bolus, electronic ID, and ear tags. The devices have different types of sensors and chips to gather data that is required for animal identification and traceability. The information is later sent and stored in basic databases. In the case of Botswana, [4], [91] discovered that in 2001, Botswana implemented a system at the Department of Veterinary Services funded by the government called the LITS, which carried out animal traceability. After experiencing problems with the collection of data, Botswana further developed the BAITS used for animal record keeping and data collection, to coincide with LITS.

The central database maintenance that Botswana used proved to be extremely poor, which further led to database updating delays and, henceforth, a generation of reports [92]. Boy [93] concurs that LITS then BAITS manages data collected from animals by sending it all to a single computerized system that houses the database for the data collected. Furthermore, just storing the data in a database has shown to be a major problem, and even the maintenance of the central database is extremely poor [91], [93]. The government of Botswana then employed local private companies to manage the LITS central database, but their services failed due to not being able to handle the capacity of the complex database which further led to database data lagging, loss of information, and henceforth the generation of reports and the system is not available online, so the farmers do not have access to the information [91].

According to Monyatsiwa et al. [94] LITS was introduced in Botswana as a cattle computer-based individual animal identification system. Furthermore, a ceramic reticular bolus containing an RFID

microchip was used as the cattle identifier. The data is sent to a data repository called the LITS database that is synchronized across the country in field stations, slaughterhouses, and border points. Monyatsiwa et al. [94] describes the two main sub-systems in the LITS systems, the first being the core system, which includes a central database, applications, remote distributed data, requirements for training and handling, including specialized equipment. The other sub-system is the remote data capture system, which includes the radio frequency devices, software and database application, and the radio frequency ID chip that is placed in the animal.

Botswana mainly launched LITS because of the EU regulations, but they later discovered the vast advantages of animal traceability using RFID, even though they have not been able to implement it. Animal traceability also offers the advantage of curbing cattle rustling activities, which Botswana has yet to curb this challenge. The stealing of animals is known as cattle rustling [57]. The results (Chapter 4) of cattle rustling are tragic; livelihoods are affected, socioeconomic conditions are disrupted, and communities are left in conflict. However, veterinary and farmers are not the only stakeholders of cattle, there is BMC and Matimela officers who have animal traceability operations. Each stakeholder has their separate animal traceability system, while veterinary and farmers use BAITS, BMC uses a system called the integration system, Matimela office currently has a system being developed for them by BITRI, and the Police contacts the veterinary to get information. All stakeholders have thing "cow" in common; the incorporation of Ontologies will provide an interoperable, integrated, and shareable system for all stakeholders. An interoperable system will reduce costs, time, lagging reports, response time by stakeholders, and cattle theft.

Despite the potential benefits of RFID in enhancing animal traceability and mitigating cattle rustling − a significant challenge yet to be fully addressed in Botswana − the integration of these technologies into effective management practices remains limited. The multiplicity of stakeholder-specific systems complicated the traceability process; veterinarians and farmers use BAITS, the Botswana Meat Commission (BMC) utilizes an integration system, and the Matimela[1]. The office is developing a new system, leading to fragmented and often incompatible data management efforts across the sector. An integrated and interoperable livestock traceability system could resolve these issues by consolidating stakeholder efforts into a unified framework, reducing costs, improving response times, and enhancing the effectiveness of anti-theft measures.

2.3 Ontologies

Ontologies are classified by Verdonck et al. based on the purpose they fulfil, which is in the assistance of human agents' communication, in the improvement of process and quality in engineering software systems, and in achieving interoperability [95]. Several basic questions arise when building a new ontology: What type of methodologies and methods can be used to build an ontology, a new ontology from scratch or reusing ontologies? What type of tools are to be used for the ontology development process? What type of language can be used for ontology implementation? [96]. The characteristics of ontologies provide advanced methods for computer science to search, manage, and enrich knowledge. This subsection discusses and justifies the choice of methodology (METHONTOLOGY), ontology development tool (Protege), and use of the modern semantic web standards (OWL 2 and RDF). OWL 2 is an ontology language used for the semantic web, used with information written in RDF. It provides data values, classes, individuals, and properties stored as semantic web documents (https://www.w3.org/TR/owl2-rdf-based-semantics/).

The existing traceability ontology frameworks are presented, along with the main ontology components and gaps identified. The research gaps of the identified traceability ontologies are summarized in detail. This research chose to develop the animal traceability ontology from scratch after thoroughly reviewing

[1] Matimela - the management of stray cattle in Setswana

the existing ontologies. This section justifies the choice of developing the animal traceability from scratch rather than reusing or extending an already existing ontology. The process of developing an ontology from scratch and the choices made are explained

2.3.1 Ontology Development Methodologies

Over the years, an evolution of ontology methodologies has been developed and improved throughout the growth of ontologies. Corcho et al. [96] relays the timeline of different ontology methodologies starting with the general first steps of Cyc creation by Lenat and Guha in 1990, followed by the enterprise ontology and Toronto Virtual Enterprise (TOVO) in 1995, based on experience gathering. Next, Bernaras made a presentation during the 12th European Conference for Artificial Intelligence in 1996, showcasing the electrical networks ontology for the Espirit KACTUS project. In the same year METHONTOLOGY was developed, followed by SENSUS ontology in 1997, presented for building ontologies, after a couple of years named after the project appeared the methodology On-To Knowledge [96].

In 1994, the TOVO was developed, involving the construction of logical knowledge models for the ontology domain [97]. Grüninger and Fox [98] explain that they developed TOVE intending to create an enterprise model for the next generation, with the ability to deduce answers to queries that inquired about shallow knowledge of the domain. Contreras et al. [99] discovered that the TOVE ontology was able to represent a common-sense enterprise model. Uschold and King in 1995 proposed a methodology for ontology development called the skeletal methodology [100]. As stated by Uschold and King [100] the skeletal methodology included a few numbers of steps that they believed would be a requirement in a comprehensive methodology. Corcho et al. [96] lists the steps of the Uschold and King's method as 1. Identifying the ontology's purpose, 2. Building the ontology, 3. Evaluating the ontology, and 4. Documenting the ontology. In 1996, Uschold and Grüninger refined the skeletal methodology by including steps 1. Capture the ontology, and 2. Integration and coding of ontologies [99]. However, Contreras et al. [99] argues that the methodology did not have defined, practical guidelines that were provided for the evaluation of the ontology.

The Kactus project showcased an ontology development methodology called the Kactus methodology, named after the project [101]. The Kactus methodology was presented at the 12th European Conference for Artificial Intelligence (ECAI'96) held in 1996, with the key objective of evaluating knowledge reuse feasibility for systems that were complex [99]. During the same year, METHONTOLOGY was introduced [101] and refined and extended [101], [102] in the following years. According to Corcho et al. [96] METHONTOLOGY was developed at the Technical University of Madrid in the artificial intelligence lab for the provision of creating ontologies from scratch or ontology reuse. Contreras et al. [99] describe METHONTOLOGY as a methodology that defines in detail the ontology development process, with great interest being the conceptualization stage that has activities of building a glossary of terms, concept taxonomies, binary relations diagrams, concept dictionary, description of each binary relation, description of each instant attribute, and description of each constant. METHONTOLOGY has the objective of facilitating ontology construction at the knowledge level [97]. This research chose METHONTOLOGY ontology development methodology to develop the animal traceability ontology, following the phases 1. Knowledge acquisition, 2. Specification, 3. Conceptualization, 4. Formalization, 5. Implementation, 6. Evaluation, 7. Documentation. METHONTOLOGY provides the benefits of having a stable methodological rigour, including stakeholders, and adapting to real-world domains, such as animal traceability, with the focus on the Botswna scenario. The following tables and discussion justify the reason for choosing METHONTOLOGY over the other ontology development methodologies.

Activities	Uschold and King	TOVE	METHONTOLOGY	Brusa et al.	UPON
Specification	√	√	√	√	√
Conceptualization	p	p	√	√	√
Formalization	p	√	√	p	√
Implementation	√	√	√	√	√
Maintenance	x	x	p	p	x
Knowledge acquisition	√	p	√	√	√
Evaluation	x	x	√	√	√
Documentation	√	x	√	√	√

√: Supported; p: Partially supported; x: Unsupported

Table 3 Methodologies compared beside the ontology development life cycle [45]

Choosing which methodology to adopt for the ontology development process is important to the success of the ontology project [103]. A comparison of the different methodologies is presented [45] in Table 3, based on the ontology development life cycle. Alfaifi [104] acknowledge the importance of a comparison of the methodologies based on the ontology development life cycle. He made a similar comparison depicted in Table 4. It is apparent from Table 3 that the methodology METHONTOLOGY is more appropriate for the ontology development process. METHONTOLOGY supports all the activities and only partially supports one, whereas the other methodologies have less support for the ontology development activities. METHONTOLOGY is accepted as the common methodology for developing ontologies from various domains, even though UPON is the most recently released methodology [45]. Furthermore, most of the methodologies do not provide essential activities for ontology development, like the TOVE methodology, which does not have evaluation activities that are important for ensuring the quality of the ontology.

Methodology	Year	Ontology development process
TOVE (Toronto Virtual Enterprise)	1995	1) Identify motivating scenario, (2) define informal competency questions, (3) define the terminology of the ontology, (4) define formal competency questions, (5) specify the definitions and constraints on the terminology, (6) test the competency of the ontology
(Enterprise Ontology) skeletal methodology	1995	(1) Identify purpose, (2) build the ontology (ontology capture, ontology coding, and integrating existing ontologies), (3) evaluation, (4) documentation, (5) guidelines for each phase
METHONTOLOGY	1997	(1) Specification; (2) Knowledge acquisition; (3) Conceptualization, (4) Integration, (5) Implementation, (6) Evaluation, (7) Documentation
Simple Knowledge-Engineering Methodology	2000	(1) Identify the ontology's domain and scope, (2) think about repurposing current ontologies, (3) enumerate important terms in the ontology, (4) define the classes and the class hierarchy, (5) definition of properties of classes, (6) definition of restriction of the classes, (7) create instances
The On-To-Knowledge Methodology	2001	(1) Feasibility study, (2) Knowledge Kick-off, (3) Refinement, (4) Evaluation, (5) Maintenance

Table 4 Comparison between the ontology development methodologies [104].

Table 4 chronologically lists by published date the different types of common ontology methodologies from the oldest to the newest, inclusive of the name, author, and ontology development process. According to Alfaifi [104] the methodologies in Table 4 have conceptually similar requirements and steps despite the difference in the ordering of steps. Furthermore, He gives an example of the TOVE methodology and Simple Knowledge-Engineering methodology, which employ "identify motivating scenario" in their first step, whereas Enterprise Ontology employs "identify purpose". The methodology words used for the steps are semantically similar but differ linguistically [104]. The different methodology comparison in Table 4 by Alfaifi [104] is similar to the methodology comparison in Table 3 [45]. Both authors conducted a comparison of the methodologies based on the ontology development activities. However, Table 3 focuses mainly on the type of support that the methodologies have for the main steps of the activities, whereas Table 4 concentrates on each of the specific ontology development processes.

Methodology	User-centred design	Modular oriented	Domain oriented	Incremental and iterative	Quality oriented	Competency-based evaluation
Lenat and Guha (1989)	No	Yes	No	Yes	No	No
Grüninger and Fox (1995)	Yes	No	Yes	No	Yes	Yes
Uschold and King (1995) Uschold and Grüninger (1996)	Yes	Yes	Yes	No	Yes	Yes
Bernaras et.al (1996)	No	Yes	Yes	No	Yes	No
METHONTOLOGY	Yes	Yes	Yes	Yes	Yes	Yes
CommonKADS	No	Yes	Yes	Yes	No	No
Noy and McGuinness (2001)	Yes	No	Yes	Yes	No	Yes
NEON	Yes	Yes	Yes	Yes	No	Yes

Table 5 Ontology methodologies compared [99].

A comparison is provided of various ontology methodologies in Table 5 based on the user-centred design, modular-oriented, domain-oriented, incremental and iterative, quality-oriented, and competency-based evaluation [99]. Modular orientation is explained as a main factor in the achievement of maintenance and reutilization [99]. The grouping of domain-related terms into hierarchies and giving them class identities is known as domain-oriented. Contreras et al. [99] describe user centre as a methodology that examines unique requirements for users in defining the competency questions, while incremental and iterative approaches deal with a sequence of phases in the procedures. The last methodology criterion is quality-oriented, presented at the start of the design and used in the evaluation to check for success. As depicted in Table 5, the methodology that is most complete based on the set criteria is METHONTOLOGY. Al-Baltah et al. and Contreras et al. [45], [99] agree that METHONTOLOGY is the most suitable methodology to use for ontology development from their provided metrics. Even so, Contreras et al. [99] states that METHONOTOLOGY has the disadvantage of generating an effort overload of documentation in the design and development process.

Bautista-Zambrana [103] suggests adopting a methodology that shows defined, ordered steps to follow for developing an ontology. According to Al-Baltah [45] a good ontology methodology should describe essential activities, how each activity should be carried out, in which order each activity should flow, and the techniques required to implement the ontology. Periamalai et al. [105] states that METHONTOLOGY provides instructions that have the most details and association with the ontology editors. Moreover, Bautista-Zambrana [103] advises using METHONTOLOGY as a methodology for ontology development because it fulfils the most criteria than other methodologies. Periamalai et al. [105] agree that METHONTOLOGY is the best choice due to its clear step-by-step refined process and detailed instructions. This research chose the METHONTOLOGY methodology to develop the animal traceability ontology. Using the detailed phases 1. Knowledge acquisition, 2. Specification, 3. Conceptualization, 4. Formalization, 5. Implementation, 6. Evaluation, 7. Documentation (5.3.2 Ontology Development – METHONTOLOGY). METHONTOLOGY supports tools such as documentation templates, a checklist for evaluation, and ontology lifecycle management, which makes it suitable for ontology projects with the intention of being deployed in real-world systems. The animal traceability ontology intends to serve as an effective guiding principle for the development of animal traceability systems.

2.3.2 Ontology Development Tools

There are various ontology development tools, and the most relevant ontology development tools will be discussed. Ontology development tools are integrated suites used for building ontologies from scratch or merging existing ontologies [46]. Gomez-Perez [46] says it is important to have some kind of supporting tool when developing ontologies because it can be time-consuming and complex. In addition, the tools provide support for ontology export and imports, ontology documentation, ontology languages, management of ontology library, and ontology graphical edition.

Slimani [106] explain the DUET tool as software that makes use of a simple UML profile that contains stereotypes that are for ontology modelling and properties. Whereas the Ontolingua tool is one of the first ontology development tools developed in the 1990s at Stanford University [107]. Gokhale [107] states that Ontolingua assists users in the creation of new ontologies and modifying already existing ontologies collaboratively. OntoEdit is said to be an exceptional tool that integrates multiple aspects of ontology engineering because it is a modern method of ontology development [106]. Corcho [108] adds that OntoEdit provides a flexible and extensible environment with functionality for browsing and editing ontologies. Ontology development tool WebOnto was developed in the UK to support the creation and editing of ontologies, also providing collaborative browsing [107]. Protégé is an open-source, standalone application developed at Stanford University by Stanford Medical Informatics [108]. Furthermore, its main feature is being an ontology editor with multiple plugins. The OntoSauru is a tool described by Slimani [106] as a web browser that provides a graphical hyperlinked interface for LOOM knowledge bases. Slimani [106] describes the tools OILEd and WebODE. OILEd editor supports the development of OIL-based ontologies, whereas WebODE was constructed as an extensible, integrated, and scalable workbench for covering and giving support to activities associated with ontology development.

A broad overview of the available ontology editors that can be used to build ontologies is provided. A brief discussion of the tables is provided along with a comparison of the tools considering the differing properties of the editors. The tables that follow are categorized into: general description, interoperability, and usability, adopted from [48], [96].

	Developers	First release date	Current released version and date	Pricing policy
Protégé	SMI (Stanford University)	Nov 1999	5.6 Oct 2014	Free and Open source
Top Composer	Top Quadrant	Dec 2009	3.0 Apr 2012	Software licensing
Neon	NeOn Foundation	Dec 2009	2.5.2 Mar 2014	Free and Open source
Fluent Edition	Cognitum Service team	2012	2.0 Dec 2014	Commercial (Software licensing)
VocBench	ART Research Group (University of Rome Tor Vergata)	2008	3.0 Sep 2017	Free

Table 6 General description of the developers, release dates, current versions, and pricing policy of the ontology tools [109], [110], [111].

The development of an ontology is a base for knowledge-based systems, requiring standard-compliant, robust, and user-friendly tools. The first ontology development tool presented in Table 6 is Protégé. Protege is developed by the engineers at Stanford University as an open-source editor. Since its release in 1999, it has been continually providing updates, and it is on version 5.6.5, released in 2024. The new version supports RDF, OWL 2, and the integration of reasoners [110]. What makes it more flexible than other ontology development tools is the provision of the cloud-based extension called WebProtege. WebProtege is freely accessible to the public and enables ontology editing that is collaborative [111]. The ontology tool Fluent Editor allows developers to model using controlled natural language (A. Seganti), which is helpful to non-technical users in developing an ontology. However, the tool was released in 2012, and the last available update was in 2014 (Table 6). No recent development activity has been seen, which is a concern in terms of its long-term sustainability. Fluent Editor, being commercial, is also not user-friendly to researchers and developers who have limited funds, unlike the Protege tool, which is free and open source. Even though the VocBench ontology tool is open source and provides multilingual vocabularies, it is tailored more to the management of thesaurus and vocabularies [109]. Instead of providing the complete OWL 2 ontology, reasoning and engineering. It is best suited for collaborative ontology management and data projects that are multilingual. VocBench was developed in 2008 by the ART Research Group at the University of Rome Tor Vergata [109] as seen in Table 6.

TopQuadrant developed the TopBraid Composer in 2009 as presented in Table 6. The tool provides enterprise-grade modelling of ontologies. It has the advantages of supporting SHACL, RDF, OWL 2, and SPARQL, that advanced. It is popularly used for semantic integration and data governance in industrial applications. However, like the Fluent Edition tool, it is commercial and requires a paid license, limiting its use to the public. A tool that came in 2009 with great promise was the NeOn Toolkit by the NeOn Foundation, with funding from the EU. It enables collaborative and modular ontology development, advanced features such as ontology reuse, and plugin-based extensibility. Despite the promising start, the NeOn Toolkit has since stalled development, and recent updates have not been seen

[111]. This makes it inappropriate for projects that need active support and standard compliance that is current.

The Protege ontology development tool is chosen by this research for the development of the animal traceability ontology. The Protege tool was chosen over the other tools because of the rich feature set it provides, the domain-specific, compatible semantic web standards, and active support. The research domain of animal traceability needs modelling of entities that is accurate, like cattle, farmers, animals, movements, farms, regulations, stakeholders, and health. These are all features that can be supported by SPARQL, OWL 2, and RDF in Protégé. Unlike NeOn Toolkit, which came out looking effective for modular ontology development, but now lacks maintenance that is recent for current OWL 2 standards. It cannot provide interoperability that is up-to-date as required by animal traceability. Fluent Editor was also not an option for this research since it had been last updated in 2014 and is commercial. Whereas, Protege is available for free and is maintained actively, as seen in Table 6, with the latest update in 2024. VocBench offers features that are convenient for taxonomy, but not enough for reasoning and formal semantic modelling in movement permit verification, monitoring of health, and tracking of animal movement. Similarly, TopBraid Composer enables features that are robust in enterprises, the inclination towards data integration that is required in large-scale and software licensing makes it unfeasible for the research-based animal traceability project. While Protege provides formal semantic modelling, open-source, and automatic reasoners like Pellet and Hermit [110], that ensures that the animal traceability ontology is consistent. These reasons justify the Protege tool as the right choice for developing, testing, and deploying the animal traceability ontology with its provision of a secure and extensible platform. It is the most suitable tool for the objective of this study to develop an interoperable, knowledge-driven ontology for animal traceability.

The information presented in Table 7 is based on the interoperability of the ontology development tools with other ontology tools and languages, as well as the translations from and to other ontology languages [48]. This information is important for ontology integration in applications. The ontology-driven framework proposed in this study, which includes a developed animal traceability ontology, intends to be adopted by developers and integrated into animal traceability applications. Therefore, the information in Table 7 is important to understand the interoperability of the ontology environments. Especially since few studies compare ontology language translators in the development tool. The functionality to export and import to ad-hoc XML is possible with new tools, but the possibility of exchanging ontologies among differing tools has not been explored. Kapoor and Sharma [112] acknowledge the research by Corcho et al. [96] about the lack of empirical studies on exchanging ontologies.

The term interoperability in ontologies means that semantic systems can interchange ontologies and also use them. This is made possible by the standards defined by W3C, like OWL 2 and RDF ontology language specification [113]. Table 7 presents a comparison of five ontology development tools (Apollo, Swoop, Protégé, TopBraid, OntoStudio) from Alatrish et al.'s [48] research. It shows that most of the tools support importing and exporting in multiple formats, such as OWL 2, RDF, and SKOS. TopoBraid and Protégé can read or write ontology serializations that facilitate ontology exchanges between them. The import of database tables, text files, and RDF files is supported by Protégé. Furthermore, Protégé supports the export of converted RDF graphs and RDF. Protégé can easily collaborate across the different ontology development environments. For example, an ontology designed using OWL 2 in Protégé can be imported into the VocBench tool for vocabulary editing that is multilingual, after being re-exported for validation and reasoning in Protégé.

	Apollo	Swoop	OntoEdit	TopBraid Composer	Protégé 2000
With other ontology tools	No	No	OntoAnnotate, OntoBroker, OntoMat, Semantic and Miner	Sesame, Jena, and AllegnoGraph	PROMPT, OKBC, JESS, FaCT, and Jena
Imports from Languages	Apollo Meta language	OWL, XML, RDF, text formats	XML, OWL, Excle, RDF, ULM 2.0 database schemas, Outlook E-mails	RDFa, WOL, XML, RDF, XHTML, UML, GRDDL, RDB with D2RQ, Microdata, Web Data Sites, SPIN, Spreadsheets, Oracle databse, text file, RDF file, News Feed, Email and Excel	XML, RDF(S), XML Schema
Exports to languages	OCML nad CLOS	RDF (S), OIL, DAML	OWL, RDF, RIF, SPARQL, F-Logic, Excel	HTML, UML, XSD, Excel, RDB, Oracle database, RDF file, XML File, Text file OKBC syntax	XML, RDF(S), XML Schema, FLogic, CLIPS, Java HTML

Table 7 Comparison of ontology tools interoperability translations to/from languages

	Graphical taxonomy	Graphical prunes (views)	Zooms	Collaborative working	Ontology libraries
DUET	Yes	Yes	No	No	No
OILEd	No	No	No	No	Yes
Onto Edit Professional	No	No	No	Yes	Yes
Ontolingua	Yes	No	No	Yes	Yes
OntoSaurus	No	No	No	Yes	No
Protégé 2000	Yes	Yes	Yes	No	Yes
WebODE	Yes	Yes	No	Yes	No
WebOnto	Yes	Yes	No	Yes	Yes

Table 8 Usability features of the ontology development tools [48].

As noted by Alatrish [48], usability in ontology development tools is closely associated with the availability of graphical editors, access to reusable ontology libraries, and support for collaborative work, as illustrated in Table 8. Both Alatrish [48] and the study presented by Kapoor and Shama [112] identifies Protégé as the most user-friendly tool, distinguished by its intuitive and easily navigable graphical interface. Protégé is further recognized for its advanced features that facilitate collaboration and cooperation in ontology construction [112], making it the most frequently recommended tool for ontology development [48]. Among its notable functionalities are the ability to customize the interface layout and enhance visualization, including features such as taxonomy viewing, zooming, and pruning. In contrast, alternative tools such as Swoop lack fundamental support mechanisms, including integrated help functions, which may limit their effectiveness and ease of use [48].

There may be an enormous number of ontology development tools that are available and free on the web, but multiple studies have singled out Protégé as the best ontology development tool to use when developing ontologies from scratch. Khondoker and Mueller [47] carried out a study involving participants to find out which of the six tools was the best. His study used Protégé, Swoop, TopBraid Composer, OntoTrack, Internet Business Logic, and IHMC Cmap Ontology Editor ontology development tools. The results of the study show that Protégé is the most domain-independent and dominant of all the tools, with 75%, and 55.5% found protégé to be the most interesting, easy, and good to develop ontologies [47]. Furthermore, the participants said protégé could be learned in one month, and this was proved by the results, where 11 participants out of 32 completed their task in protégé in 4 minutes. The multiple literature of ontology development tools found that all led to Protégé being the best tool to use. Thus, this study chose protégé as the ontology development tool for developing the ontologies for animal traceability.

2.3.3 Ontology Languages

Ontology languages are defined as formal languages that are used for the construction of ontologies in computer science [114]. Moreover, the languages permit knowledge encoding of specific domains with the inclusion of reasoning rules for processing knowledge. There are two categories of ontologies: the traditional ontology languages and the web-based ontology languages [115]. Su and Ilebrekke [115] list the first category being KIF, OCML, CycL, Ontolingua, and F-logic. The second category belongs to traditional and web-based, known as RDF and OWL. Kalibatiene and Vasilecas [114] explains that part of the language is implemented in the ontology development tools. Slimani [106] has created a summary of languages by presenting a comparison of the languages as shown in Table 9. The thorough comparison with details of how the listed languages manage to meet the presented criteria in Table 9, General issues, Attributes, Facets, Taxonomies, Relations, Axioms, Instances and others. The symbol √ means the language corresponding satisfies the requirement, and the symbol X means the language corresponding does not satisfy the requirement [106]. The results from Table 9 show that OWL 2 ontology language satisfies all the requirements except for one, being named axioms.

	LOOM	OCML	FLogic	SHOE	OML	XOL	RDF(s)	OIL	DAML + OIL	OWL	OWL2
General issues											
Concept partitions	√	X	X	X	√	X	X	√	√	X	√
documentation	√	√	X	√	√	√	√	√	√	√	√
Attributes											
Instance attributes	√	√	√	√	√	√	√	√	√	√	√
Class attributes	√	√	√	X	√	√	X	√	√	X	√
Local scope	√	√	√	√	√	√	√	√	√	√	√
Global scope	√	X	X	X	√	√	√	√	√	√	√
Facets											
Default slot value	√	√	√	X	X	√	X	X	X		
Type constraints	√	√	√	√	√	√	√	√	√	√	√
Cardinality constraints	√	√	X	X	X	X	X	√	√	√	√
Slot documentation	√	√	X	√	√	√	√	√	√	√	√
Taxonomies											
Subclass of	√	√	√	√	√	√	√	√	√	√	√
Disjoint Decomposition	√	√	X	X	X	X	X	√	√	X	√
Relations											
n-ary relations	√	√	X	√	√	X	X	X	X	X	√
Operational definition	√	√	√	X	X	X	X	X	X	√	√
Axioms											
1st order logic	√	√	√	X	√	X	X	X	X	√	√
2nd order logic	X	X	X	X	X	X	X	X	X	√	√
Named axioms	X	√	X	X	X	X	X	X	X	X	X
Embedded axioms	√	√	X	X	√	X	X	X	X	√	√
Instances											
Concept instances	√	√	√	√	√	√	√	√	√	√	√
Facts	X	X	X	X	√	X	X	X	X	√	√
Other											
Rules	√	√	X	√	√	X	X	X	X	√	√

Table 9 Comparison of the Ontology languages [106].

The semantic language called Web Ontology Language 2 (OWL) is designed for the representation of complex and rich knowledge about things and relations among things [116]. OWL was recommended in February 2004 when it was made a World Wide Web Consortium recommendation [49], [106]. Thus, making it compatible with eXtensible Markup Language while extending the Resource Description Framework (RDF) [49]. The OWL documents are usually published online to be referred to [116]. Hitzler et al. [116] explain that OWL 2 is useful in expressing ontologies with a complex history in computer science, and is useful in creating computational artifacts. Maniraj and Sivakumar [117] mentions that OWL ontology interprets data in the form of individuals and property assertions that relate the individuals to one another. Maniraj and Sivakumar [117] further describes OWL as being both a syntax for the exchange and description of ontologies and being able to formally define semantics that give meaning to them. OWL 2 is recommended for use in ontology because of proven usability and its ability to present the ontology in a format that is comprehensive [110]. With these multiple benefits of OWL 2 and RDF, this research chose to use OWL 2 and RDF ontology languages for the development of the animal traceability ontology.

2.3.4 Traceability Ontology Frameworks

No.	Framework	Main Components	Gaps Identified
1	Dairy Cattle Performance Ontology (DCPO) Framework. [118]	It is a unified framework that brings together the complex existing structure of a heterogeneous kind of dairy data. This will ensure extensibility to data that is external. The framework offers a vocabulary that is common for all stakeholders involved. There is a provision of knowledge management tools that are automated. The main focus of the ontology framework is modularity.	The framework lacks explainability support. There is no data-centered ontology for knowledge discovery. Multiple ontology models have not been integrated to make one main model. Lack of interoperability.
2	FoodOn Ontology Framework. [18]	Framework built to focus on accessible and easy international farm-to-fork food ontology. It describes common cultural food from around the world that is accurate and consistent. It addresses the problem of data harmonizations that spans food security, quality, consumer health, distribution, and production domains. The framework should produce systems that are interoperable with software-driven intelligence. FoodOn provides the provision of growing production of food knowledge base for reasoners.	The framework is too broad as its focus is on different types of cultural food around the world. It is broken into multiple ontologies, and there is no integration. Still has to refine the food product knowledge base for reasoners.
3	Food Track & Trace Ontology Framework (FTTO). [119]	FTTO is an ontology framework developed to manage food traceability. Food tracking along the entire supply chain – from farmer to consumer. Provides standardization along with conceptualization of food. Should be connected to the Information System for Global Track and Trace. Aims to create a food traceability domain inclusive of all food concepts ordered in a single hierarchy. FTTO has four main modules food, actors, products, and processes, part of the chain of supply.	Does not have measures in place to ensure correct information extraction from farmers. Focusing on all food products will make a system that is complicated. It will not be easy to use and navigate with over a thousand foods to consider. Mainly focusing on the regulations of food quality being checked by stores and not consumers. It should involve consumers fully. Benefits are one part, which is companies and not the producers (farmers). A framework developed to work with an already existing system.
4	ValueGo traceability ontologies. [19]	ValueGo is a Business Process Modelling Ontology. Proposed to model the processes of a business at a semantic level.	The business model so far is only focused on filata, one pasta out of the various kinds of pasta available. It is not efficient nor effective for a whole dairy supply chain.

		Integrates knowledge of the context of the organization, the workflow activities, and the semantic web services. It is developed to specifically manage dairy. Tracing the business models and operations of fresh cheese pasta filata. ValueGo should be able to detect and then record the quality-related activities during the lifecycle of the products in a database. The consumer will be able to trace all product information.	
5	Ontology-Based Traceability System for Halal Status of Flavour: A Conceptual Framework. [120]	Conceptual framework traces the Halal status of flavours for Halal for faster certification. Used to determine the source of raw material and the status of Halal. The ontology is used to link multiple food sources for flavour information representation. The ontology framework for halal traceability will be validated for tracing by using the flavour Halal status tracing system. The ontology framework will integrate references that are related to Halal flavouring.	Ontology serves as a minor component of the framework. There is no interoperability involved, which is the main feature of ontologies. No description of the knowledge base and reasoners provided. Does not take full advantage of the opportunity's ontologies offer. A conceptual framework is developed to aid in the design of another traceability framework. A framework developed to work with an already existing system.
6	TraceFood Framework. [121]	Food supply chain. Sector-specific ontologies, data element relationships. Standardises traceability data, enhances interoperability	General framework, not specifically tailored for animal traceability
7	MEat Supply Chain Ontology (MESCO). [122]	Meat supply chain. Activities, actors, food products, lots, processes, service products. Enhances traceability and safety in the meat supply chain.	Specific to meat, not all-inclusive for livestock management
8	prFood. [123]	Food provenance and contamination risk. Food history, contamination risk, and provenance. Integrates with existing ontologies, focuses on contamination risk	Primarily safety-focused, with less emphasis on comprehensive traceability
9	Food Supply Chain (FSC) Ontology. [124]	Food supply chain. Supply chain stages, relationships, knowledge representation with OWL, and reasoning capabilities. Applicable to all food types, enhances semantic mediation	Broad focus, not specific to livestock traceability
10	Outbound Logistics Process-Internet of	Food supply chain. IOT4Log for Internet of Things concepts, OLP-Onto for outbound logistics	Focused on IoT integration, with less emphasis on semantic detail

	Things (OLP-IOT). [125]	processes and resources, and a Space-Time module for spatiotemporal relationships. Support real-time monitoring, improve safety and transparency	
11	UHT Milk Production Ontology. [126]	UHT milk production. Raw materials, processing, transportation. Integrates blockchain for enhanced transparency	Limited to UHT milk production, not applicable to all livestock systems
12	Ontology-Based Traceability System (Onto-Fog/Onto-Cloud). [127]	Interconnected supply chains. IoT, cloud, fog computing. Enhances visibility and coordination, supports tracking.	General application, not specific to food or livestock traceability
13	Animal Traceability Ontology (ATO)	Current study. Livestock traceability systems in Botswana. Classes (e.g., cattle, farmer, disease), hierarchical taxonomy, object and data properties, formal axioms, and instance's for animal traceability. Enhances interoperability, supports data integration, includes robust disease management, is scalable for small and large farms, and aligns with Botswana's specific livestock practices.	Generalizability to other geographical contexts might be limited (not yet investigated), currently focused on cattle only

Table 10 Traceability ontology frameworks

The literature presented ontology-driven frameworks designed specifically for different aspects of traceability. Table 10 lists the ontology-driven frameworks' traceability.

The Dairy Cattle Performance Ontology (DCPO) Framework

Fuentes et al. [118] discovered that most of the data regarding animal health issues, nutrition, genetics, and animal pedigree are heterogeneous and complex. The organization of datasets that are complexly structured will require considerable data modelling. Furthermore, the whole farming process needs to be computerised. This inspired the domain modelling of the farming dairy process called the dedicated domain ontology (DONT) [118], which is the DCPO as seen in Table 10. DONT is beneficial in its ability to provide rich data schema, a domain expert centralized repository, and vocabulary that is standardized for stakeholder collaboration. The ontology traceability project was carried out in Canadian dairy farms. The main component is modularity. Fuentes et al. [118] states that the ontology is divided into multiple aspects representing the dairy process. The aspects are core, testing, production, breeding and genetic evaluation, and quality control. Figure 1 is adopted from Fuentes et al. [118] depicting the modular design of the dairy ontology. Figure 1 shows the ontology entities, classes start with uppercase letters, and the lowercase letters are the properties.

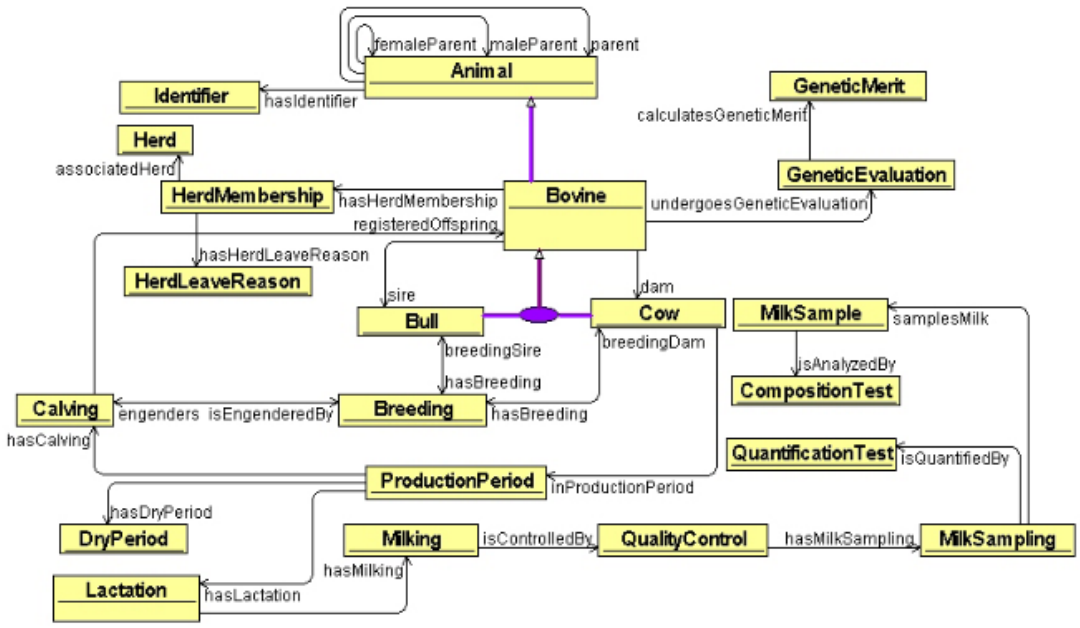


Figure 1 Illustration of the modular design of the dairy ontology [118]

FoodOn Ontology Framework

The FoodOn Ontology Framework was developed to provide food description vocabulary that can be used by consumers, industrial applications, and researchers, as seen in, which is the DCPO, as seen in Table 10. Dooley describes ontology as a resource that is made up of term hierarchy facets to describe the type of raw food ingredients, preservation and cooking, packaging terms, and multiple product type schemes that categorizes food products [18]. Figure 2 is an illustration of the FoodOn ontology framework [18].

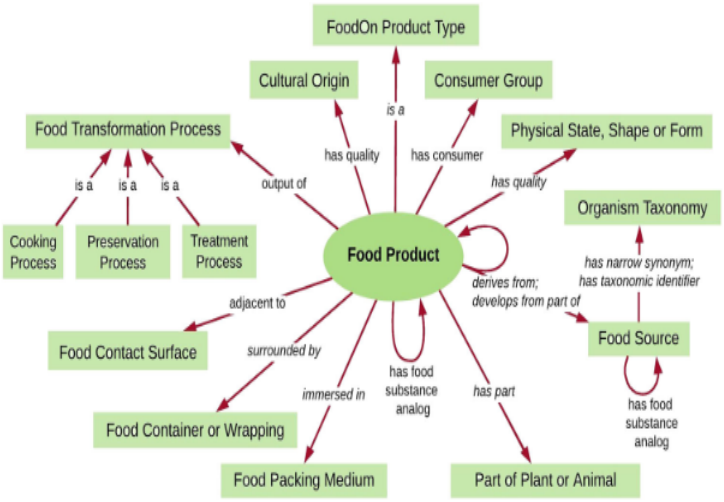


Figure 2 Pictorial of FoodOn Ontology [18]

Food Track & Trace Ontology (FTTO)

The FTTO is a framework that is proposed to be used in the future for all food traceability, as summarized in Table 10. The ontology is developed with a traceability semantic model for the reusability of existing information resources that are part of the tracing and promotion of efficient, accurate, and reliable information management systems [119]. Four modules make up the ontology framework: actor, product, process, and food [119]. Agents in the ontology have their knowledge base

containing information that could be expressed. Communication through the food supply chain by the agents happens through constructs from the FTTO ontology. The framework is designed to work with the Global food track & trace system as the base for FTTO, making it possible for agents to simultaneously communicate and cooperate while receiving and delivering messages. The main focus stresses Pizzuti et al. of FTTO is to act as a reference ontology model in the creation of food traceability ontologies in future research, to support safety control and quality [119].

ValueGo Traceability Ontologies

Beneficial attributes noted in using ICT-based solutions in the supply chain are the support of tracing and tracking products, capturing data, storing, recording, and disease tracking, with measuring variables in the environment [19]. The main benefit they seek is the capability to integrate and share the information in the food supply chain. Magliulo et al. Illustrates an ontology application used as a tool for modelling the business process and rules of the agricultural food chain [19]. The Bovlac project was used as the business case; it is a technological and scientific platform for tracing the production of fresh cheese. The description of people in each stage is done through use cases, including the activities and roles. Tracing is done through the use cases in their supply chain macro-phases. Magliulo et al. [19] state that the knowledge of the interested domain can identify any significant figures in the business use cases. ValueGo ontology was developed to define the business models, operations, and production of Pasta Filata, the fresh cheese, as illustrated in Figure 3. Figure 3 is the taxonomy of the Pasta Filata processes that take place in the supply chain. The business processes that are specific provision of detecting and recording the quality-related activities of the product lifecycle in a database. Furthermore, allowing for easy traceability of product information by consumers when they purchase.

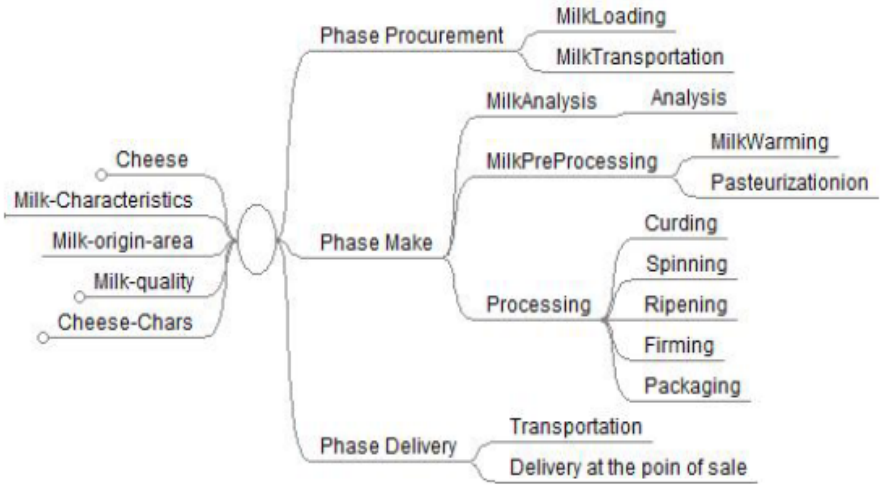


Figure 3 Illustration of the taxonomy process of Pasta Filata [19]

Ontology-Based Traceability System for Halal Status of Flavour: A Conceptual Framework

Hashim identified challenges that were contributors to delaying the halal certification application approval [120]. The challenges were identified through interviews with JAKIM authorities and conducting a review of related works, as summarized in Table 10. Hashim found that the main delay in issuing approval of halal certification was the difficulty in determining the halal status of flavours [120]. Moreover, the cause was the absence of a halal certificate for processing application certification documents by the auditors. The other identified issue was that tracing of halal status of flavouring was difficult because of the inconsistent terms being used by auditors and food producers. The research aims to solve the issues by integrating the constructs and factors as a framework. The concepts of halal certification process, reference sources, halal concept, and traceability using ontology technology will

be represented as relationships in the conceptual framework, as seen in Figure 4. Based on the conceptual framework, it will aid in the design of another framework to trace flavouring information to make the halal certification faster [120].

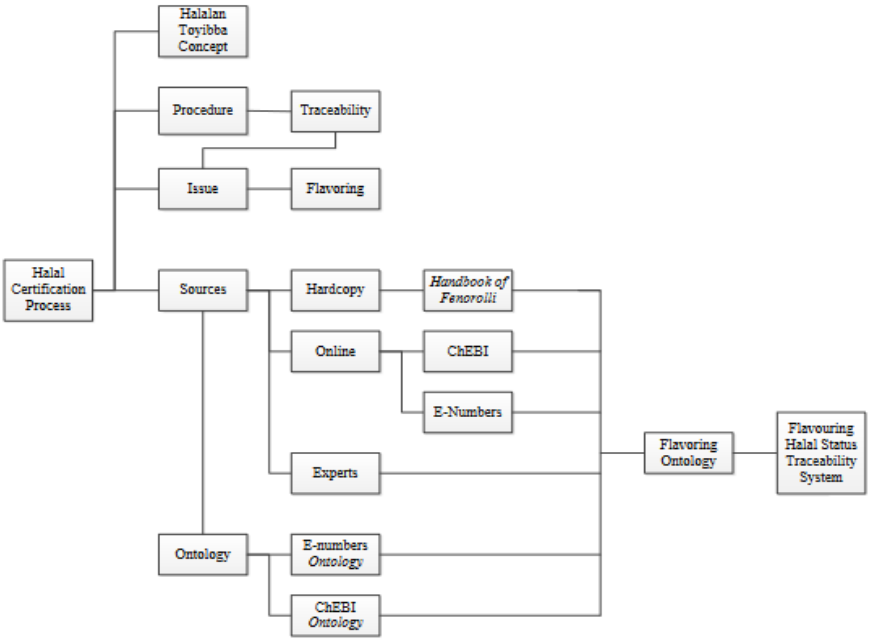


Figure 4 Conceptual Framework for an Ontology-Based Traceability System for Halal Status of Flavour [120].

2.3.4.1 Identified Research Gaps and Context-Specific Challenges

Scalability presents another significant challenge for ontologies like FoodOn [18], the DCPO [118], and UHT milk production ontology [126]. These ontologies often need substantial customization to fit various geographical and operational contexts, making their deployment complex and resource-intensive. Such limitations reduce their effectiveness across global supply chains that require flexibility and adaptability to diverse conditions and regulatory environments.

Accessibility for small-scale producers, especially in developing countries, is limited by the high technological requirements embedded in ontologies like FTTO [119], FoodOn [18], and the UHT milk production ontology [126]. These ontologies typically rely on advanced technological infrastructure, which can be a barrier to adoption among smaller stakeholders. Developing simplified versions or adaptations with lower technological demands could facilitate broader adoption, ensuring more inclusive participation in traceability initiatives.

While some ontologies, such as DCPO [118], address aspects of dairy health management, others like Bovlac/ValueGo [19], FTTO [96], and MESCO [99] lack comprehensive disease management functionalities. Effective animal traceability requires detailed modules for disease monitoring, including prevention, early detection, and response strategies. The absence of these critical features limits the ability of these ontologies to support comprehensive health monitoring and implement biosecurity measures effectively.

These identified gaps underscore the need for future advancements in animal traceability ontologies to focus on improving scalability, integrating accessible technologies for small-scale producers, and incorporating robust disease management systems. Addressing these areas will significantly improve the efficiency, inclusivity, and reliability of traceability systems.

In addition to the identified gaps in existing studies on traceability ontologies, our study aimed to address challenges specific to the Botswana context. Botswana' livestock sector faces intertwined

challenges of declining cattle populations due to poor management, climate-induced droughts, and systemic inefficiencies, compounded by inadequate traceability systems, limited market access, infrastructure deficits, low technology adoption, and technical literacy, all of which hinder food security, economic sustainability, and compliance with international trade standards. The animal traceability ontology addresses Botswana's identified challenges by standardizing data representation to improve management practices and decision-making, and response to systemic inefficiencies. It tackles the issue of inadequate traceability systems by enhancing interoperability and integrating disease monitoring, vaccination records, and movement tracking to manage livestock health and reduce mortality rates. Additionally, the animal traceability ontology supports market compliance by aligning traceability data with international trade standards, while its modular, cost-effective design promotes accessibility for small-scale farmers, addressing infrastructure deficits, low technology adoption, and limited technical literacy.

2.3.4.2 Limitations of Existing Ontologies and the Rationale for the Development of the Proposed Ontology

The reuse of existing ontologies is encouraged to reduce the effort, but in some cases, like this research, reusing ontologies was not conducted. The animal traceability ontology was developed from scratch. Some ontologies considered for reuse include Food Track & Trace Ontology Framework, Ontology-Based Traceability System, TraceFood Framework, and prFood Ontology. However, after thorough research, they could not be reused or extended for the animal traceability ontology. There were multiple reasons discovered. The existing ontologies did not cover domain-specific terminology and a specific scope. The conceptualisation developed by the ontologies was inconsistent for the development of the animal traceability ontology. The documentation of ontologies like the Ontology-Based Traceability System and the prFood Ontology was insufficient documentation to use. Using existing ontologies may also have restrictions in licensing and usage for distribution and publication. The animal traceability ontology is developed to be distributed and published to the research and development community. Specifically for Botswana to adopt and improve the current animal traceability system. The animal traceability ontology included the creation of specific axioms, but the existing ontologies could not support the custom rules.

The Ontology-Based Traceability System was specifically focused on modelling processes and relationships to determine the halal status of food ingredients. With concepts like halal certification procedures, traceability of chemical properties, and reference sources. They do not encompass the classes and properties for animal traceability, such as cattle movement and farm management practices; there was domain misalignment. The ontology structure of Food Track & Trace Ontology Framework, Ontology-Based Traceability System, TraceFood Framework, and prFood are designed for their domain, notably differing from the processes of animal traceability. It would have required substantial restructuring to accommodate the entities and relationships that are different. Attempting to reuse the existing ontologies could have caused semantic inconsistencies because the ontology constraints and definitions may conflict with the requirements to accurately model the animal traceability ontology. Furthermore, the animal traceability ontology does not reflect only the structure. But also, the purpose, semantic misalignment may occur when reusing a design for a purpose that is different.

The Animal Health Ontology was the closest potential ontology that could have been used to extend and develop the Animal Traceability Ontology. However, only two classes, three data properties, and five object properties could be adopted. The ontology had no individuals created. The creation of classes, data and object properties, and individual by class for the animal traceability ontology was substantial as it would have been the same as creating the ontology from scratch. As a research student who wants to venture into the ontology field, developing the ontology served as a learning tool to increase my understanding of the ontology design principles

2.4 Chapter Summary

From the literature review carried out, it is apparent that there are more animal traceability systems designed and developed than frameworks that guide the actual system development. It is also evident that the same systems face many issues and would benefit from a total overhaul following proper guidelines. This research proposal posits that to get a better chance of developing a good system, there needs to be guidelines that guide the process in the form of a framework. Furthermore, the few information management methods that were reviewed were mainly storing data on basic databases, overlooking some very pertinent data concepts like metadata and the semantics (meaning) of the data. Fortunately, there exist tools like ontologies, which were created to bring meaning and descriptions to data, an aspect that has not been applied to animal traceability, though suggested by some researchers. As important as managing the data is, making sure the data that is being managed is of good quality should be a priority. This research, therefore, envisions the development of a framework that guides the creation of Animal Traceability systems. This is consistent with suggestions by researchers in the shown literature who suggested that having more than one methodology may give a better opportunity of reaching success.

3 CHAPTER 3 RESEARCH DESIGN AND DATA ANALYSIS

3.1 Introduction

The purpose of this chapter is to describe the research methods and data analysis approaches followed in collecting and analyzing the data that is needed for the design of the proposed framework and ontology, which are defined in Chapter 5. Research is described as a systematic and methodical approach that must meet crucial standards and norms of reliability and validity [128]. Moreover, it must serve the purpose of solving a problem or adding a contribution that is innovative to the body of knowledge. To accomplish channel work in research, a vital role should be played by the research design [129]. Furthermore, the research design must be grounded on a methodology after the problem and content have been formulated. The selection of the research design and methodology should be based on the research problem and questions [130].

Methodology is defined as a method of solving a problem in research systematically by adopting certain steps logically [131]. The methodologies are described and analyzed, also shedding light on their limitations. A methodology is a mode that entails theoretical principles, with the provision of a framework for guiding how the research is conducted [130]. The purpose of this research is to explore and describe the user experiences and current system functionality of animal traceability in Botswana. To achieve this, an explanatory research design was chosen. This chapter discusses the research methodology, data collection process, data analysis, chosen population and sample, pilot study process, and the reliability and validity of the research design. Chapter 3 of this research addresses

RQ1. What are the perceived limitations and challenges of the current animal traceability systems in Botswana, according to the main stakeholders?

3.2 Research Design and Methodology of the Study

Researchers must think about what theoretical perspectives and knowledge they intend to provide. A researcher should reflect on the types of strategies intended to be used in the research, which should inform the chosen methods, and think about data collection and analysis procedures [132]. A blueprint for data collection and analysis constitutes a research design. This research has made decisions of the research design to keep the researcher mindful of bias they may bring to the study, how the decision affected the research objectives and questions, the benefit and use of the proposed ontology-driven framework for animal traceability and choosing effective tools for the collection of the research data. Research design should be built on a precise established statement of the problem, research objectives and questions, and a clear purpose of the research [128], [133]. The research design was chosen and built based on the research problem, “This research aimed to address the existing gap in interoperable, knowledge-based frameworks for developing animal traceability systems and related platforms in Botswana.”

Researchers should ask themselves questions about their theoretical perspective and the new knowledge being brought to their research area [134]. They must think about the strategies to be used in their study that will inform them of the methods chosen. The methods are based on data collection and data analysis strategies. The explanatory research design is chosen for this study, which focuses on the why questions. Explanatory research is defined as research with a problem which has not been thoroughly explored previously, requiring priorities that generate achievement, description, and produce a more effective research model [129]. It is chosen to focus on identifying what is happening, addressing challenges identified, gaining new insights, and better explaining and detailing the aspects of this research. It will assist in providing a lead to future researchers and studies in ontologies, animal traceability, and various fields. Explanatory research is used to explore the research topic at different depth levels.

Research designs are various methods of inquiry that are within qualitative, quantitative, and mixed approaches [134]. The research benefited from the use of mixed methods, employing both quantitative and qualitative research approaches to answer the research questions. For the implementation of evidence-based and innovative practices that could be complex, one methodology is not adequate [135].

The Collection of data from the stakeholders will provide first-hand information on the current challenges being faced in animal traceability. Furthermore, the information will form the solution to the concepts used to develop the proposed ontology-driven animal traceability framework and the interoperable animal traceability ontology. The data collection tools that were used were survey methods. The survey methods are face-to-face interviews and questionnaires. The data collected was analyzed using both quantitative and qualitative data analysis tools. Content and Thematic analysis were carried out using computer-assisted qualitative data analysis software called NVivo 12 for qualitative data analysis. For quantitative analysis, descriptive statistics analysis where conducted. The results from the analysis are presented in the findings and discussion in Chapter 4. Figure 5 illustrates the research data design of this study, depicting the stage-by-stage process that was used for better clarification and visualization.

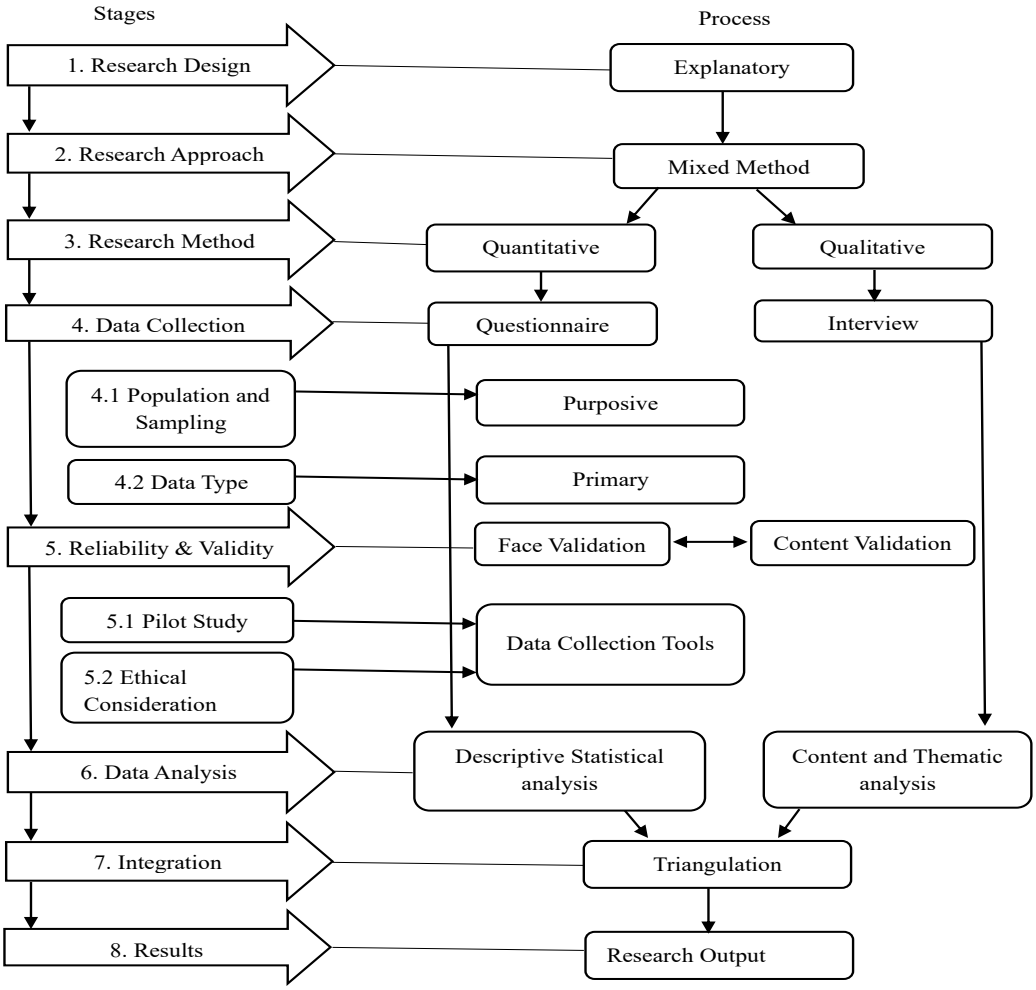


Figure 5 Research Design of the Study Framework.

3.3 Research Approach

The first stage was requirements elicitation. The research stakeholders were identified for requirement elicitation. The tools, interview questions and questionnaires, for data collection were then developed. Upon completion of the tools, next was the identification of participants for the pilot interviews and requests sent. A schedule was made for pilot interviews and conducted with the selected participants. The data collection tools were updated based on the pilot interview results. The final data tools were sent to the supervisor for approval.

The second phase began while waiting for approval of data collection tools. A schedule was made for the first interviews and site visits. The documents for the research permit application were filled out, processed, and sent for approval. The Ministry of Agricultural Development and Food Security (reference DVS 7/4/1 VII (90)) and the Ministry of Communications, Knowledge and Technology (reference MCKT/1/11/1 I (38)) approved the research permits to carry out this study. Upon the approval of the data collection schedule and research permit, requests for interviews and questionnaires were sent to stakeholders. Site visits to stakeholders commenced, and relevant data were collected. Sorting, synthesizing, and analyzing data began after the first site visit using data analysis tools. The cycle continued with the second and third site visits. Gaps were identified, and data clarity was probed for a second visit. The data was synthesized and deduced for user requirements and components for the development of the ontology-driven framework. Addressing the research questions of this study. Data analysis continued, and the data analysis chapter began after completion of data collection.

3.3.1 Mixed Research

The research notation system mixed method was created in 1991 by Morse [136]. The Morse system depicts whether a research project orientation is quantitative, qualitative, or both. It indicates if the methods are equal, one is less dominant, and the other is dominant. The notation system also describes whether the research project is conducted concurrently or sequentially. This study chose to employ the mixed research method, which uses both qualitative and quantitative methods. Radhakrishnan states that the mixed method may not be superior, but its advantages of the methods complementing one another, continuity, rectification of differences, and theoretical insights that are heightened speak volumes [137].

The research problems are better understood and expanded by the combination of the quantitative and qualitative methods. In mixed-method designs, the quantitative method is used to gain a breadth of understanding of the implementation success of the predictor. The qualitative method explores and gains a depth of understanding of the success and failure of the research area and identifies strategies to allow implementation. To achieve the required results for this study, a quantitative method was used to carry out questionnaires for data collection and descriptive statistics for data analysis. Qualitative method data collection was conducted by interviews and computer-assisted qualitative data analysis for data analysis.

A. Timing

There are different types of mixed method research designs classified into six: sequential explanatory, sequential exploratory, sequential transformative, concurrent triangulation, concurrent nested, and concurrent transformative [128], [138], [139]. Creswell and Guetterman [140] encourage concurrent design because it will reduce the number of site visits for the researcher, and data collection will be better managed at the same time. Research carried out by Hanson et al. [138] acknowledges Creswell and Guetterman [140], suggesting that the most effective method is concurrent triangulation. Almalki [132] concur that the most effective design is the concurrent triangulation after using it to carry out his research. The multiple studies that recommend and justify the use of concurrent triangulation validate the choice for this study.

The research design that was found suitable for this study was the concurrent triangulation design. Using the concurrent triangulation design requires consideration of timing. The timing choice depends on the researcher's intent of which method comes first. Concurrently, both quantitative and qualitative

methods are conducted at the same time. For the ontology-driven framework for animal traceability study, the concurrent triangulation timing was used for data collection. This was done to better manage the data collection activities at stakeholder sites and make sure the limited resources would be enough to complete the study. This meant that less finances were spent conducting the interviews at the same time in the nine districts of Botswana, also being conscious of the time.

3.4 Research Methods

3.4.1 Quantitative Research

Quantitative research is a method that involves the quantification and analysis of variables to produce results [141]. Statistical techniques are used for numerical data analysis to answer questions like when, what, who, how, and where. Quantitative research includes using scientific sampling with a developed questionnaire for measuring the characteristics of a population by statistical methods [141]. The type of data used in quantitative research is measurable figures, numbers, and percentages [142].

To explain and predict a phenomenon, answering questions from measurable variables is quantitative research involving numbers and anything measurable systematically [143]. It is research that is concerned with phenomena measurements based on quantity and carried out by statistical analysis [131]. Additionally, it is commonly relayed from primary data in the form of questionnaires. Gathering and analyzing numerical data is a form of quantitative research [144]. The non-experimental quantitative research has no experiments involved and uses a group design called a survey design. The survey type produces numeric descriptions of trends, attitudes, and opinions from a selected sample of a population [128]. Generalizing and summarizing data findings is easier when using quantitative research methods [145] .

This study employs a mixed research method, which includes quantitative research. The survey design from the non-experimental group was selected during the data collection phase, specifically questionnaires. The data collected was quantitatively analyzed using Microsoft Excel software. The software is used for descriptive statistics and pattern analysis. The initial stages of analysis are the descriptive analysis methods used for the summarization and description of data [146]. Furthermore, it is used for the establishment of a comprehensive guide to where, how, and what the impacts of the lack of frameworks in animal traceability are.

3.4.2 Qualitative Research

Qualitative research is a form of inquiry. Ospina et al. [147] define it as the value of a mode of inquiry that is systematically empirical. Empirical meaning experiences of the world inquiry, systematic by planning, and value of understanding the participants' experiences. Inductive mode is used to allow the data to be expressive and speak for itself [147]. Chinyere and Val [144] acknowledge the use of qualitative research, interpreting rich meaning to data. Furthermore, intricate information is uncovered, revealing new concepts in research. Qualitative research is utilized to identify the experiences, ideas, and opinions of the participants. The data is non-numerical, like audio, text, and video. The main qualitative research methods are action research, narrative research, grounded theory, phenomenology, and ethnography [144].

Qualitative research is defined as a form of phenomenon description and interpretation based on the participants' views to develop theories and concepts [143]. Hancock et al. [148] affirms that the development of new theories and concepts is made possible by focusing on describing and interpreting data. The data gathered and analyzed is usually from documents such as open-ended answers to questions. Mohajan [143] list that qualitative research uses observations, interviews, diaries, open-ended questionnaires, and journals. Busetto et al. [149] add field notes, interviews, focus groups, and audio recordings as qualitative methods for data collection. Kabir [150] is accord listing interviews, participation, and observation as the basic methods for qualitative data collection. For data analysis,

audio is transcribed to transcripts, and coding is performed using software for qualitative data management.

This study employed qualitative research as the other part of a mixed methodology to answer the research questions. The answer is determined by seeking an understanding of the questions from a sample selected from the local population. This was done using the predefined set of procedures to collect data, analyze the data, and produce results. The collected data results are useful beyond the boundaries of this study. The data collection techniques used for qualitative research were semi-structured interviews. After gathering the data, it was analyzed using software called NVivo 12 to discover themes and concepts. As stated by Busetto et al. [149], qualitative designs are employed to discover invisible and unpredictable patterns in data.

3.5 Data Collection

Data collection will be used to gather data to gain first-hand knowledge and original insight into the research problem. The data collection methods selected for mixed research are the survey methods. Specifically, a questionnaire for quantitative data collection and qualitative interviews. The main aim of using these instruments for data collection is to gain an understanding from the perspective of participants of their lived world experiences, communicated in specific details in their own words. According to Roberts [151] the reason data collection is done is to get a sense of the subjectivity of the participants.

3.5.1 Qualitative

A. Interview

The survey methods will be used to gather experiences, opinions, and characteristics by directly asking people questions. A conversation that is face-to-face with a participant is called an interview. Bhome et al. [152] explain that the interview process does not require recording answers by the interviewer only, but also observing the interviewee's non-verbal communication. The researcher observed the body language of the interviewees, for example, hand gestures and facial expressions. This study chose a semi-structured type of interview process. With semi-structured interviews, the interview was formal, with an interview guide that was created with a list of questions that addressed the research topic. The semi-structured interview allows the interviewer to follow a guide but also interview topical trajectories straying from the questions, which are appropriate for research [150].

Subsequent knowledge, including what happens during the interview, may change and alter the interpretation of research participants' experiences. Roberts [151] proposes that the interviewer and participant shall cooperate when answering research questions in an interview. Moreover, carrying out qualitative interviews will give privileged entry into the lived world experiences of research participants. It is a planned and focused conversation that is carried out to be able to see the world from their point of view, to ascertain what they have experienced, and then find out about life before science explanations [40]. The main aim is to obtain an understanding of the sense and experience from the participants, dictated and described in their own words, in specifics for the researcher to put aside what they know and note what is said [151].

Guidance for the interview tool was adopted from Castillo-Montoya [153] who created the Interview Protocol Refinement Method, which has four phases as seen in Figure 6. Roberts [151] is in accord and recommends this type of method for preparing the interview tool. A set of open-ended questions will be prepared and validated for the stakeholders' interviews, forming the interview tool. Peredaryenko and Krauss [154] advise that it is important for the researcher to adopt an open stance and be aware of their personal biases with each interview; they should have an open attitude, willing to discover new knowledge. The interview tool was first trialled through a pilot study with a selection of professionals. Feedback from the professionals applied in the document was relevant. Updates were made when it was given final approval by the study supervisor and used in the field to guide in interviewing respondents.

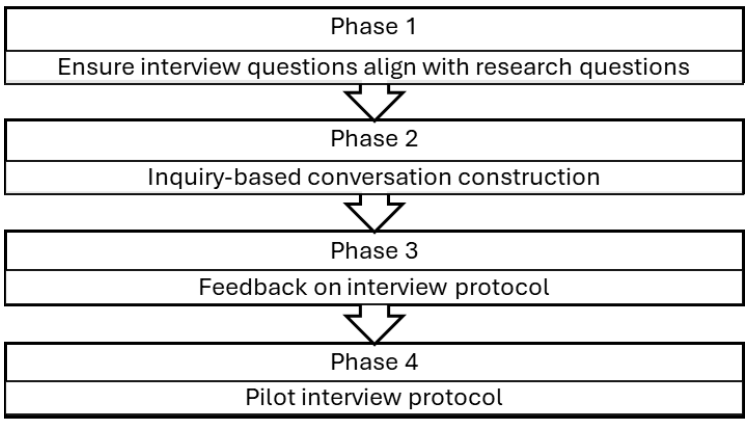


Figure 6 The interview protocol refinement method [153].

B. Procedures for conducting Interviews.

Table 11 is a schedule made for the interview process of the research. It depicts the tasks taken along with their timeline during the interview process.

1. Pilot interview
2. Choose participants
3. Schedule interviews
4. Conduct interviews
5. Triangulation
6. Reinterviewing
7. Analyzing data
8. Write up

Task ID	Task	Earliest start date (week)	Duration (weeks)	Parallel/sequential	Dependent upon Task ID
A	Develop an interview schedule	1	1	Sequential	-
B	Conduct pilot interviews	2	1	Sequential	A
C	Finalize interview schedule	3	1	Sequential	B
D	Select Participants	1	2	Parallel	-
E	Conduct interviews	4	12	Sequential	C & D
F	Triangulation	4	12	Parallel	E
G	Reinterviews	6	8	Parallel	F
H	Analyze Data	11	4	Sequential	F & G
I	Write up	13	8	Parallel	A & H

Table 11Task list for the interview process for the research.

3.5.2 Quantitative

A. Questionnaire

The questionnaire was first created by Sir Francis Galton [145]. It is used for various reasons, like gathering statistical data. The questionnaire is a document that has a set of questions along with answers that participants fill in, giving information required by the researcher [145]. It is one of the most used types of data collection methods. Kumar [155] describes a questionnaire as a method where participants read the questions for themselves and make an interpretation of their expectations, then choose an answer from the given options.

The questionnaire tool for the research was created following the guidance of Roopa and Rani [156]. Following his five stages of planning a questionnaire, as seen in Figure 7. The questionnaires were given to the employees who are part of the stakeholders' organization; some were interviewed, others were not. The plan was to give paper questionnaires to respondents who do not have the means of accessing the online questionnaire; the rest sent the online questionnaire. This plan was not fully successful. The online questionnaires sent were never returned, some of the paper-based questions were not filled out, and others were not returned. Questionnaires are used to reach a larger number of participants because interviewers are limited to the number of people that will be interviewed.

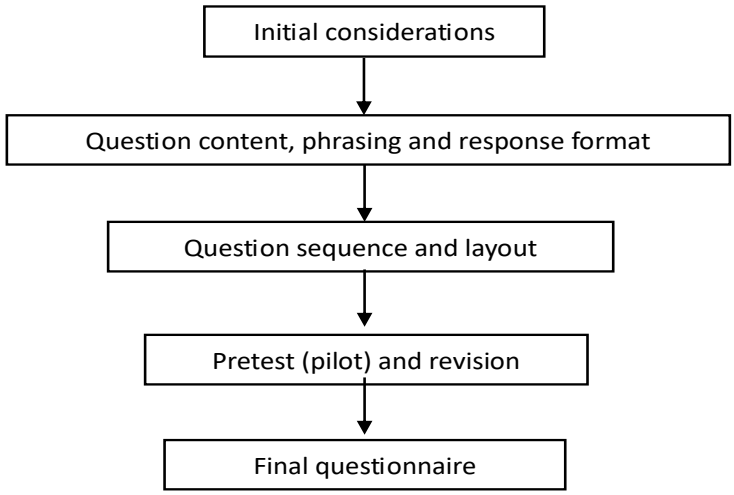


Figure 7 Stages of planning a questionnaire [156].

3.5.3 Study Area

The research study was conducted in Botswana, a landlocked country which is located in the center of Southern Africa, bounded by South Africa to the south and southeast, Namibia to the north and west, with Zimbabwe and Zambia to the northwest. An estimation of 965km from east to west and 965km from north to south, and a protruding sharp point at the eastern side [157]. Botswana is a country with a rich variety of wildlife and plenty of species like fish, reptiles, mammals, amphibians, and birds. It is traditionally viewed as a cattle country with the livestock population concentrated in the north around the perennial rivers and hill ranges at the eastern hard veld. Considering the plentiful water, pasture, and animal disease controls in place, the environment is healthy enough to raise high-quality and high-bulk indigenous beef cattle. Botswana is divided into nine districts, namely, Central, Kgatleng, Southern, South-East, Kgalagadi, Gantsi, North-west, North-east, and Kweneng. This study was carried out in each of the districts. The researcher conducted site visits to farms and organization offices for MOA, BMC, Veterinary, and Local Administrative Authority (Matimela).

3.5.4 Data Sources

There are two types of data sources: primary data and secondary data. The research used both data sources, the primary type was used during data collection, and secondary data throughout the study.

A. Primary Data Source

The primary data was acquired from the original information source in the field. The data sources were farmers, veterinary officials, BAITS Agents, BMC, MOA, and Matimela officials. Data was obtained from them through interviews and questionnaires. The advantage of direct contact with the events as they occurred in primary data made it more reliable.

B. Secondary Data Source

Record books, reports, templates, and management books are the secondary data sources that were obtained from the stakeholders of the research. The organizations provided reports and templates for animal registration, ear tags, cattle transfers, and permit applications. The farmers had record books and management books, but that was a small number of farmers. The majority of the farmers who took part

in the study did not keep records. Reputable journals, articles, proceedings, books, newspapers, and websites were used throughout the research. The farmers record books, reports, management books, veterinary templates, and policies and guideline documents were added along with the primary data in NVivo software for processing and analysis.

Tacit knowledge refers to the personal, experience-based understanding which cannot be easily articulated, transferred to others, or written down. Unlike knowledge that is explicit which can be codified and shared easily. Tacit knowledge includes wisdom, skills, intuition, and practical know-how that is gained by doing. Tacit knowledge was not included as a possible data source in the study, it was deemed out of scope of the current and can be considered as future work to extend work done in this thesis.

Population and Sampling

A. Population

The research population included participants/stakeholders who are involved and have a background in animal traceability and livestock in the Botswana area. Stakeholders are farmers, veterinarians, Matimela, MOA, and BMC. The current and past experience in using animal traceability techniques in Botswana validated the choice in the population selected. They were able to provide their opinions, experiences, and recommendations that provided answers to the research questions.

B. Sampling

Purposive sampling is explained as a sampling technique that selects specific types of respondents to include in the research to fulfil the expected variations relevant [149]. The respondents are based on theory, literature, and previous experience. Campbell et al. [158] argue that the use of purposive sampling is for the selection of participants that provide a high chance of yielding useful and relevant information, also choosing cases that save limited resources. The sampling method assures that relevant cases that could be included make the final research sample. Choosing a purposive and small sample is made with the focus of improving the depth of understanding in research based on the objectives and research questions [159].

Tongco [160] describe purposive sampling as a deliberate selection of a participant based on the participant's qualities, as well as not needing a definite number of participants or underlying theories. Furthermore, the researcher decides to seek respondents who are willing and can provide information based on their experience and knowledge. Etikan [159] is in accord with Tongco [160], the selection of participants is based on respondents' willingness to participate by communicating their opinions and experiences in a reflective, articulate, and expressive manner. Tongco [160] say samples chosen using purposive can be used both qualitative and quantitative. Purposive sampling was chosen for this study because the research topic and questions required participants with specific characteristics that will assist effectively with the relevant feedback. Not like studies that are random including a diverse selection of cultures, and backgrounds. Isaac [161] say if purposive sampling is conducted correctly, it will help filter out responses that are not relevant and not suitable for the study framework. He further says the sources being appropriate for the research framework will greatly lower the margin of error during the data collection phase.

The study used the purposive stratified sampling method for the selection of participants. The nine districts of Botswana make up the different locations used as the nine strata for the sample frame. A list of farmers was obtained from the Government Veterinary officials for representation of each district strata. The list was used to randomly select farmers from each district to participate in the study. In districts that did not provide a list of farmers, the researcher went to farm areas and randomly selected farmers to interview.

C. Sample size

The sampling choices were made based on research found on sample sizes for qualitative and quantitative research. According to Omona [162] data saturation should be considered when choosing

a sample size such that data should not be too small it cannot achieve saturation or too large that there is repetition of data. In a study on animal traceability by Plessis and Rand [31] used a sample of 30 participants for their data collection. A recommendation by Creswell [138], and Creswell [139] for research design that follows the mixed method design is to interview 30 based on the site visits. Furthermore, there is more involved in the selection of sample size than just the participant's number. Omona [162] say that it depends on how many times contact is made with each person, and how long contact is made. These were taken into consideration when conducting the data collection with the participants. The sample size chosen for this study was 60 participants, covering interviews for the qualitative methodology as recommended questionnaires for quantitative methodology [31], [133], [134].

After completing the data collection activities, the research ended up with 66 respondents who were willing to participate in the research. This was because some of the farms included both the farm owner and farm managers. The participants included farmers, farm managers, veterinarians, Matimela (stray cattle) officers, BMC, MOA, and all who were involved in animal traceability in Botswana. The participants were interviewed and filled in a two-part questionnaire for the study. The 66 participants provided adequate information to yield a sufficient depth of understanding of the research topic.

D. Data Saturation

Data saturation is the point of completeness in the process of research, relevant collected data, themes, and codes required for answering the interest of inquiry are found with no emergent new data [163] . A sampling approach that is developed towards saturation of data is possible for data collection completion before the data analysis phase. Ames et al. [164] define saturation as data concepts that are adequate and rich to satisfy the research questions. According to Etikan [159] the sample size of research is determined through data saturation and not analysis statistical power. When data that is new from cases does not contribute anything new or different to the creation of the emerging theory, the research has reached its data saturation point [142]. At that stage, the researcher should stop data collection to avoid data redundancy and use what they have for the next phase of research. Kumar [155] concur that the point where information is negligible or none is new, research assumes it has reached the data saturation point.

An experiment was conducted to determine data saturation, He found that 12 interviews were optimal depending on the type of research and demographic area [165]. Moreover, Guetterman, [165] states that Ph.D. students have a habit of using a sample size that is large unnecessarily, without considering the data saturation point. For theoretical data saturation using comparative methods Hancock et al. [148] suggests a sample size of 60 to avoid data saturation. This study employed a purposive sampling method with a sample size of 60 participants but ended up with 66 respondents. The sample size was chosen with the consideration of data saturation. Purposive sampling puts emphasis on data saturation, by obtaining deep understanding through the sampling frame till there is no substantive data that is new is obtained [159]. This study effectively met data saturation, ensuring no new data emerged with additional interviews, and ultimately involved a diverse group of stakeholders from the animal traceability sector in Botswana.

E. Selection of Participants

Table 12 depicts a list of the participants for data collection and their relevance to the study. Like the animal traceability study by Prinsloo and de Villiers [166] the impact on farmers by a traceability system is important to finding a solution to the problems of animal traceability. The best way is to get first-hand information from the key stakeholders. This information will play a major role in the development of the proposed ontology-driven framework.

Interviewees	Relevance to the research
Farm owner	Provide practices and requirements on animal traceability
Farm manager	Provide practices and requirements on animal traceability

Veterinary representative	Provide details on tasks performed for and with farmers for animal traceability
Local Administrative Authority representative	Provide details on Matimela and animal traceability
Ministry of Agriculture	Provide details on BAITS in Botswana
Botswana Meat Commission	Provide details on their system and cattle market in Botswana

Table 12 Participants selected for study.

3.6 Validity and Reliability

[128] argue that for anything to be called research, it should have methodical and systematic procedures that meet specific and relevant standards and norms for validity and reliability. Good research should have consideration of the two fundamental and important aspects in the evaluation of research tools, validity and reliability [143]. Sürücü and Maslakçi [167] concur with the fact that validity and reliability are indispensable instruments in research. Research conducted employing measuring instruments that do not go through validity and reliability will not produce results that are beneficial.

A. Validity

The validity of the research tool is concerned with what the instrument measures and how well it is doing that [143]. Surucu [167] agrees with Mohajan [143] that the concern of validity is if the instrument is measuring the quality and behavior as intended and how well the instrument is performing. Validity describes the extent to which data collected covers the investigation area [145]. Etikan [159] state that efficiency and validity are the most important part in purposive sampling. Observation strategies like questionnaires, interviews, and document analysis make validity stronger [168].

Validity means the measurement at which a chosen instrument is able to measure its intended purpose of measure. According to Amin, [169], validity is defined as the ability to deliver precise results and the degree of measurement by which an instrument measures what should be measured. The instrument was put through face and content validity. The interview instrument was created using the interview protocol refinement framework to ensure a reliable and valid set of questions that will provide answers to the research questions and guide the proposed ontology framework [153].

B. Reliability

Reliability is defined as the rate at which the measuring tools control random error, that is faith that is placed on the obtained data from the instruments used [143]. For testing and transparency of sampling and collection instruments Ames et al. [164] say it is important for the assurance of trustworthiness and reliability of findings. According to Taherdoost [170] the main focus of reliability is how far the phenomenon measurement provides results that are consistent and stable. Sürücü and Maslakçi [167] acknowledge that reliability indicates the stability of values measured from repeated measurements. Asenahabi [128] points out that the degree of reliable results is also determined by the participant's cooperation, trying to control the bias of the sample.

C. Validity and Reliability of the study

The questions for the questionnaire and Interview were given to three course colleagues for peer review, and three lecturers from the Department of Academic Literacy & Social Sciences in the university to check if the documents were written in simple language, easy to read, and appealing to the respondents. This was done for face validity and corrections were made as advised. For content validity, the documents were assessed by my supervisor, co-supervisor, and one lecturer with knowledge of ontologies and animal traceability. They were checking if the main aim and objectives of the research matched with the questions. Their advice was taken into consideration and corrections were made. Upon approval of all documents, pilot interviews were conducted with participants knowledgeable in the animal traceability field.

3.6.1 Pilot study

The pilot study of the research was conducted with 10 participants who have knowledge of research data collection, ontologies, and animal traceability. The sample was chosen as suggested by Kunselman [171] for a small study, 10-12 participants should be sufficient. It was undertaken using the interview and questionnaire tools developed and pre-tested for validity and reliability. The pre-testing was done by checking that each questions align with the study research questions. The reviewers were testing the tools for completeness, clarity, and appropriateness to the research topic. The first step to the pilot study after identifying the participants was sending participant request email's to the participants of the tools. The participants used review option on Microsoft word to write their comments. The participants sent their feedback and corrections were made to the tools according to the feedback. After correcting the documents the validity and reliability was conducted by carryng out interviews with the participants face-to-face. The participants were interviewed and asked to answer the questionnaire. When complete they gave more feedback to the tools and interview process.

The feedback received from the pilot participants were all minor changes, included suggestions to re-word some questions from the interview tool, combining BAITS and LITS questions because they fall under the same organization, and insert university headings in all the data collection tools. The suggestions for the questionnaire were to swap some questions for logical flow preservation, include a section for participants to enter their details, and grammatical corrections. The pilot survey provided a great contribution to the improvement of the main interview and questionnaire tools, which contributed to successful data collection in the actual collection activities.
Table 13 presents examples of the modifications that were made to the data collection tools based on the feedback from experts. The recommendation from the experts was used to modify the tools.

Data collection tool	Modification	Feedback
Interview Questions – BAITS	How do you cater for users that do not know how to use technology? "Especially in a country where most farmers are old?"	The last part of the questions is misleading and should be removed.
Interview Questions – Farmer	How do you record and manage information "about the livestock ?"	Rephrase part of the sentence to "about your livestock".
Demographic survey	What is your "ethnic" background?	Not "Racial"? Ethnic is too narrow. Eg. My ethnicity is not African, neither is it a Motswana.
Questionnaire	~~Have~~Has the development/ adoption of your system objectives been met since inception?	Replace ~~Have~~ with Has.

Table 13 Pilot study feedback examples

3.6.2 Ethical Considerations

Taherdoost [145] argue careful consideration of securing ethical concerns while all the research design and methodology choices are being made, like respondents' confidentiality. Ethical considerations were conducted in preparation for the data collection process. Confidentiality and anonymity were the main areas of note, where it is applied, and data presented by respondents is anonymized. An ethical clearance form was filled out and submitted along with the study's supporting documents to the university research ethics committee. The committee assessed the legal problems that could stem from the research studies and how they could be avoided. The legal challenges can be the country's laws on approaching citizens for interviews, respondents' mental health risks in participating, the organization's reputation, and effects on animals and the environment. The data collection documents: interview questions, questionnaire, demographic survey, consent form, and participation document were all submitted for ethical review by the committee. Taherdoost [145] insist that the ethical clearance process is conducted to make sure the study is designed based on the ethical rules required. Some of the rules are [145];

1. If the respondent's personal details are collected like names, contacts, and email the researcher should provide reasons and provide a consent form for them to sign.

2. It should be stated if participants will face difficulty during the data collection process, and how the research will assist.
3. A method of identification of the researchers should be provided to the participants to help them identify who they are.
4. Participants should be provided with a document stating the researcher's intention with the study and the expected outcome at the end of the study.
5. Participants should be clearly aware they are volunteering and can withdraw from participation at any point.

These rules were all covered in this study through the consent and participation documents submitted to the research ethics committee. The ethical approval for this study was obtained from the Directorate of Research and Development at Botswana International University of Science and Technology (reference HREC-011). The Ministry of Agricultural Development and Food Security (reference DVS 7/4/1 VII (90)) and the Ministry of Communications, Knowledge and Technology (reference MCKT/1/11/1 I (38)) approved the research permits to carry out this study. Participant consent was obtained, and data collection followed ethical guidelines, ensuring transparency and participant protection.

3.7 Data analysis

The next step after collecting data that is reliable and accurate successfully is to extract information that is useful, pertinent, and hidden in the data to interpret and manipulate [172]. Performing specific evaluations and calculations to extract necessary information from the raw data is called data analysis. Data analysis was performed in the study through several phases of both quantitative analysis and qualitative analysis. Creswell and Creswell [134] state that mixed methods data analysis should be related to the research strategy type. Furthermore, the analysis should occur within both the quantitative and qualitative research methods.

3.7.1 Quantitative

The data collected was useful in gathering the opinions, characteristics, and experiences of animal traceability from the participants. All the data was loaded into Microsoft Excel and NVivo 12 software prior to data analysis during and after data collection. The initial stages of analysis are the descriptive analysis methods used for the summarization and description of data [146]. It is used for the establishment of a comprehensive guide to where and what are the impacts and causes of animal traceability challenges. Data representation using graphics only illustrates tentative pictures of data and does not provide a specific distribution picture and prediction or estimate of value.

Data description in numerical shape employing central tendency is very important [172]. Marshall and Jonker [173] is in accord that for a better demonstration of data meaning based on numbers requires the use of statistics. Raw data like group observation characteristics can be descriptively illustrated by statistics. Descriptive statistics were used in this study for the summarization of data in a manner that is organized and provides a description of relationships among the variables in the sample. Microsoft Excel was used to compute descriptive statistics from the data collected. The central tendency feature of descriptive statistics was used specifically for the main primary measures mean, mode, and median. According to Kothari, [174] central tendency measures illustrate the main point that items have a clustering tendency. The value that gives a description of a whole set of data as a single measurement is known as the central tendency [175].

3.7.2 Qualitative

Qualitative data collected for this study was analyzed using a computer-assisted qualitative data analysis tool called NVivo 12 data analysis software. The features of NVivo include handling data by writing, coding, linking, demographics addition, pattern searching, and providing data reporting and exportation [176]. Dollah et al. [177] state that the software can analyze wide ranges of data types through progressively complicated searches with various modes of output like images, pdf files, web pages, audio, graphs, video, maps, surveys, charts, and textual. NVivo 12 is a qualitative analysis tool used to

enhance the researcher's academic work. For the assurance that the researcher is working methodically and rather thoroughly they are encouraged to employ the use of NVivo [178].

Using the grounded theory approach means that theory needs to be drawn from the data analysis. NVivo can push for a theory to be drawn from the researchers' data. Welsh [179] explain that the NVivo software has tools for memo writing which enable theory to be built for the data. He further states that this approach lets the data speak for itself unlike attempting the data from within, like using existing frameworks. NVivo was used to efficiently organize the data collected with its ability to code text onto a screen rather than manually cut and paste pieces of text into a single code. The analysis process is rigorous in permitting the researcher to accurately and quickly search specific types of data [180]. It provides valid search results by confirming all instances of the specific usage are found [179].

The first step in using NVivo was creating a project folder called Animal Traceability.nvp. The data files produced from data collection were then imported into the project folder. Codes are first created from the data contained in the files based on the relevance, similarity, and objective of the study. The codes go through a three-stage coding cycle to refine them by merging, renaming, and deleting. Once completed cases and case classifications were created to get more detailed and valid analysis results. The final stage of analysis was computing the cases created by exploring the NVivo pattern analysis tools. Based on word and data frequency and focusing on the themes and concepts that emerge. The results from the data analysis are summarized and presented in the form of tables and figures in the findings and discussion chapter of the study. The figures and tables are from both NVivo software and Microsoft Excel.

3.8 Triangulation

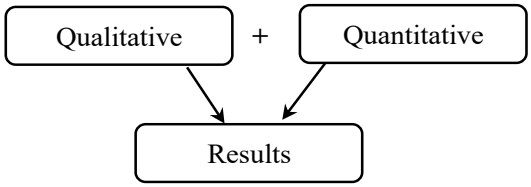


Figure 8 Concurrent triangulation design [138].

Concurrent triangulation is described as using both qualitative and quantitative methods for data collection and data analysis at the same time and the results are presented as one [138], as depicted in Figure 8. He further states that priority is given to both forms equally depending on the type of research carried out. The researcher collects data concurrently for both quantitative and qualitative approaches. The data is then compared for convergence determination. Between the methods the weight is usually but sometimes practically one method may be given higher priority [134]. The Analysis of the data is performed separately, and the integration of the data happens at the data interpretation phase. This research will follow the same manner using both qualitative and quantitative methods for data collection at the same time. The qualitative method will play a role in the requirement elicitation from stakeholders through interviews. The quantitative method will be used for the requirement's elicitation from stakeholders through questionnaires.

The mixed method research methodology is important to provide broad purposes of breadth and depth of corroboration and understanding for data collection, data analysis, and interpretation [136], [181]. Using the triangulation method will provide various viewpoints of the research from different angles. By using the different methods, it will give a better complete picture of the result. Caine [182] states that within the mixed method quantitative and qualitative research methods can fulfill purposes that are different but complementary. Asenahabi [128] agree of the triangulation design seeking to collect complimentary data for the research topic that will be integrated for interpretation and analysis. This method gives the study an opportunity for corroboration and convergence of the results derived through research methods that are different.

3.9 Data Analysis for this study

The data analysis section is based on exploring and discovering answers to the research questions posed by this study. Research questions one and two are the main questions of focus in this section. Ibrahim [172] states that summarization of data collected using closely linked processes and organizing the data in a manner that answers the questions, best describes data analysis. Bhatia [183], found that implementation of outstanding analysis techniques will substantially enhance making decisions, valuable insights are unearthed, and risks minimized. The analysis of the research took multiple complex steps to arrive at a conclusion that answered the research questions. As Ibrahim [172] said, it is very easy to organize data that is simple but appropriate processing is required for complex data.

The data analysis methodology adopted for this study is a mixed data analysis that caters to both qualitative and quantitative data. The interview transcripts and questionnaires from the collected data were used for data analysis. Computer Assisted Qualitative Data Analysis (CAQDAS) was used to conduct qualitative data analysis, specifically the NVivo 12 that provides a variety of beneficial analysis techniques. NVivo 12 was released in July 2020. Lumivero developed NVivo; they publicly avail the documentation and licensing information at https://lumivero.com/products/nvivo/. The university provides an educational license that allows the use of the analytical modules from NVivo throughout the project. NVivo 12 software is used for extensive pattern analysis by querying the data to deduce themes and concepts. NVivo is defined as a qualitative analysis tool used to enhance the researcher's academic work [177]. Author Hilal [178] concurs that for the assurance that the researcher is working methodically and rather thoroughly they are encouraged to employ the tools provided by NVivo.

The software selection was guided by open-science considerations, including potential impact on scientific reproducibility and research outputs accessibility. Piloting of alternative open-source CAQDAS such as Taguette, RQDA, and computer assisted text markup and analysis was conducted. NVivo 12 was employed because it is the only one that combines coding comparison, hierarchical cluster analysis, word frequency and a single environment with visualization functions. NVivo furthermore permits full exportation of coded data, project analysis logs and metadata in CSV/XML format (plain text) and ensures independent researchers can reproduce all the analytic steps with or without an NVivo licence. All the coded transcripts in XML format make them openly interchangeable and readable by free software. On the date of submission complete .nvp animal traceability project file, node-structure report (HTML), codebook (CSV), and query parameters (CSV) will be archived in the university repository and made available openly.

The content and thematic analysis are explored using NVivo software, found at https://lumivero.com/products/nvivo/. Hsieh and Shannon [184] define qualitative content analysis as a method in research used for the interpretation of data that is subjective by a methodical sorting procedure of coding for theme identification. Content analysis is employed for the observation of textual data to single out patterns and structures. NVivo provides multiple tools to perform content and theme analysis like coding, creating memos and annotations, coding stripes, and content linking [185]. The stages suggested by Maureen [185] for content analysis using NVivo are adopted, along with the guidelines by Saldana [186] for data coding. It allowed the researcher to be aware of the respondent's interview transcript's constant reflection, re-examinations, and confirmation of specific aspects. Microsoft Excel software is employed for the quantitative aspects of the data through calculations provided by the descriptive statistical analysis methods for accurate data results.

The statistical methods derived from descriptive statistical analysis provide ways of describing numerical data collected that is then assimilated [173]. This results in practical research findings that are evidence-based, therefore reducing the theory-practice gap. The numeric measurements provided by central tendency mean, median, and mode are useful in examining data in more detail and presenting accurate visual displays. The results are perfect for summarizing quantitative data, thus gaining beneficial insight, like deducing potential trends in animal traceability respondents' demographics.

Scherman [187] suggests a three-phase mixed data analysis, two phases for qualitative analysis and one phase for quantitative analysis. Their method proved that quantitative and qualitative analysis are both different and similar. Similar in that they are both concerned with answering research questions, data reduction, linking research literature, data analysis, and transparency. However, their differences are in the textual and numerical data presentations. Scherman et al. [187] used a CAQDAS (Atlatis.ti) for qualitative analysis and descriptive statistics for quantitative analysis. This research adopted parts of the mixed data analysis [187], for qualitative analysis this study used CAQDAS NVivo software instead. For quantitative analysis, this study followed Scherman's process for descriptive statistics data analysis.

3.9.1 Qualitative Data Analysis

This study found the NVivo toolkit data analysis process to be useful and relevant . The analysis process provided systematic phases to follow for data analysis that produced results in a logical manner. The qualitative analysis process will be followed throughout this chapter with an addition of features that are necessary for this study. The four main stages of analysis using NVivo as guided by O'Neills [185] are as follows:

1. Descriptive: It is importing data sources to NVivo software.
2. Topic: The process of data organization and data coding.
3. Analytic: performing data analysis and data querying.
4. Conclusion: Answers illustrated from the data.

The four stages are further broken down to form the sub-stages that were followed for the analysis as seen in Figure 9.

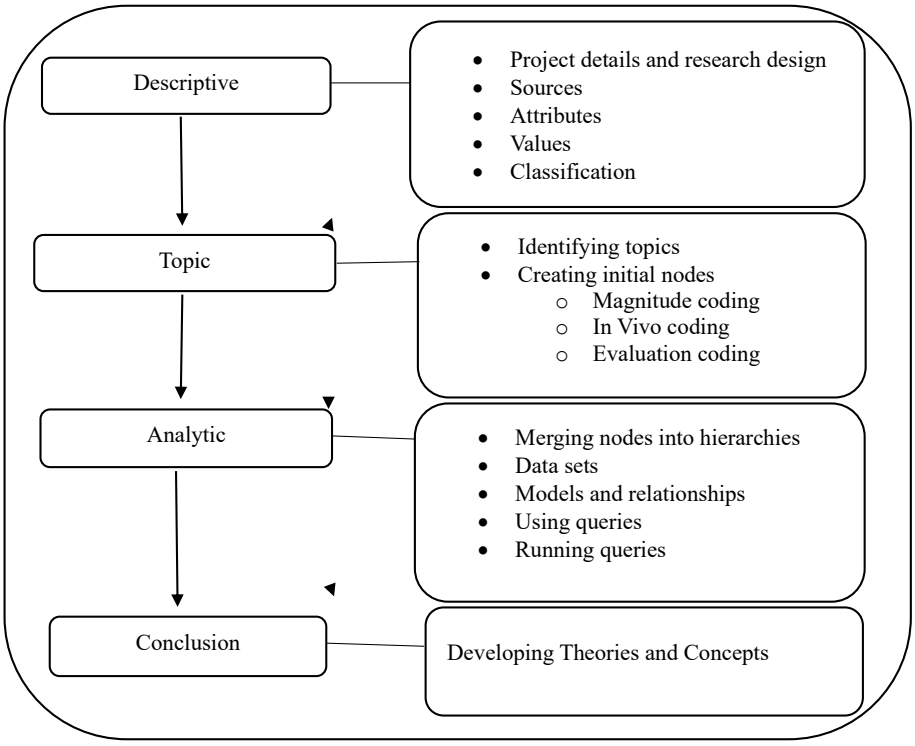


Figure 9 Schematic diagram of the study data analysis process.

3.9.1.1 Stage 1: Descriptive

This is the data analysis stage where project details and data are entered into NVivo to create a database for the research study. This study created a project called Animal Traceability.nvp in NVivo and started by importing data files into the project. The import panel used in NVivo is shown in Figure 10. The files included the interview schedule, interview recordings, interview transcripts, field notes, ethics approval, participant demographics, and a questionnaire. The documents are contained under the Data > Files in NVIVO.

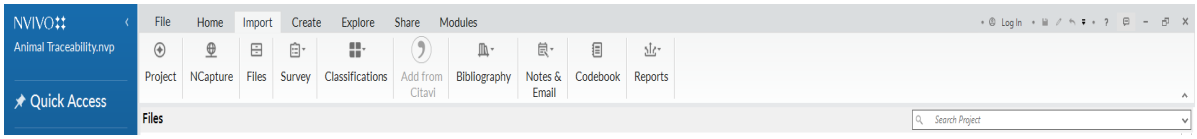


Figure 10 Import panel of NVivo software.

An initial open interview that is semi-structured and questionnaire guideline was created based on literature found on animal traceability. The data collection tools were validated by conducting a pilot study with professionals who understood the research topic. The pilot study assisted in modifying and making the data collection tools effective enough to derive feedback from respondents who answered the study research questions. Consent forms were given, explained, understood, and signed by each respondent who was interviewed and given a questionnaire to fill out. The respondent's consent allows this study to use the information they provided while keeping their identity anonymous. Field observations were also conducted during the interview sessions at the stakeholder's location, pictures of the farms, cattle, and ear tags were taken for the purpose of this research.

The text data files used for analysis included in-depth interviews that were conducted with farmers, veterinary officers, local administrative authority (Matimela officers), BAITS Agents, BMC, and MOA representatives. The main participants who are affected by animal traceability are the Botswana farmers and that is why they make up the majority of the study participants. The veterinary officers and the Ministry of Agriculture enforce and monitor the animal traceability procedures and guidelines. BMC checks if the farmers meet the given guidelines and are marketable to the European market and other markets (South Africa, Zimbabwe, Zambia, and Tanzania). BMC is the largest and highest paying beef market in Botswana.

The interview timeline for communal farmers took approximately 30-45 minutes including observation around the farm. The commercial farmer's time was longer at approximately 1-1:30 minutes of interview sessions because their form of farming was more advanced thus providing more information. The respondent's organization interview session took approximately 1-2 hours to get an insight into their current systems.

A. Data Selection

All the descriptive details of the animal traceability project were entered into NVivo under the Data section. The Data section includes subsections Files, File Classifications, and Externals. The researcher imported all data collected from the interview transcripts, demographics, and questionnaire under subsection Files.

The guidelines from O'Neill's [185] research used a different version of NVivo in which the data sets were imported to the section called Sources. The sources subsections were internals, memos, and externals which are different from the new version of NVivo 12. The primary material for research that is imported into NVivo in the internals consists of documents, pictures, PDFs, data sets, audio, and video. The external subsection represents material from research that is not possible to import to NVivo like books. Externals were not used in this study.

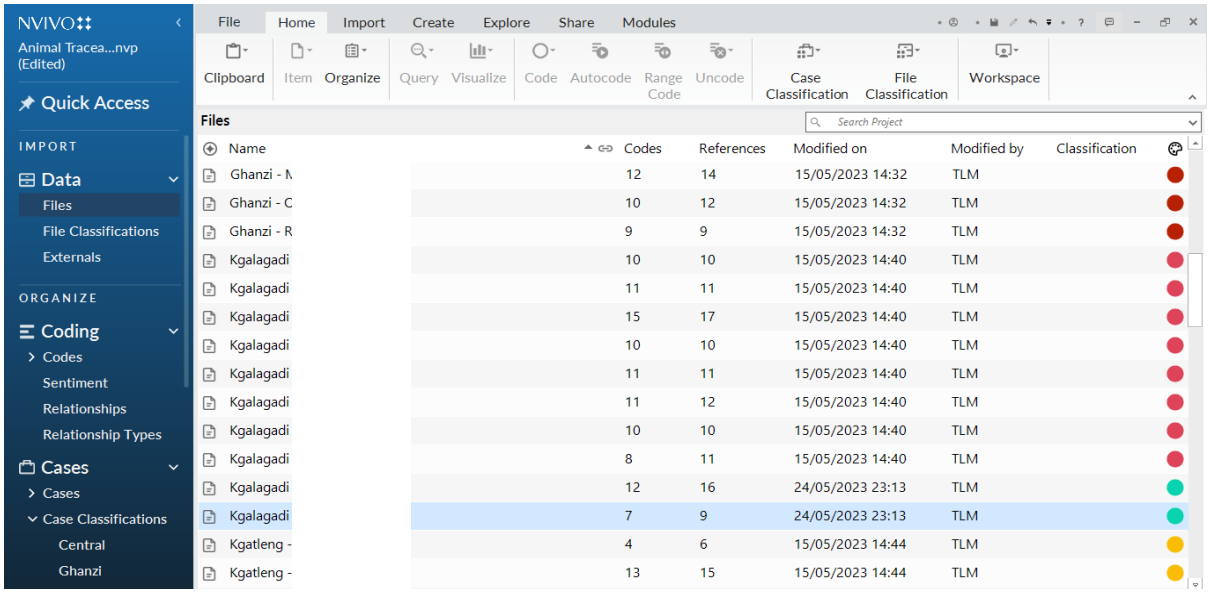
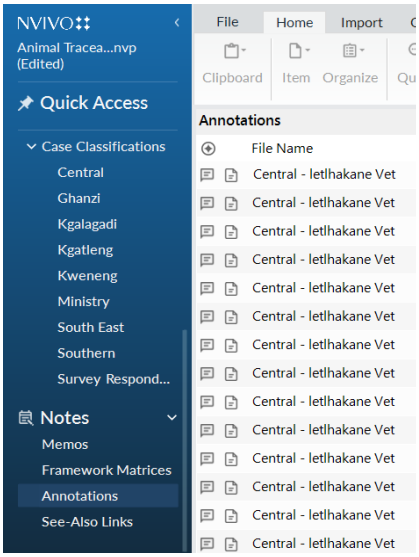


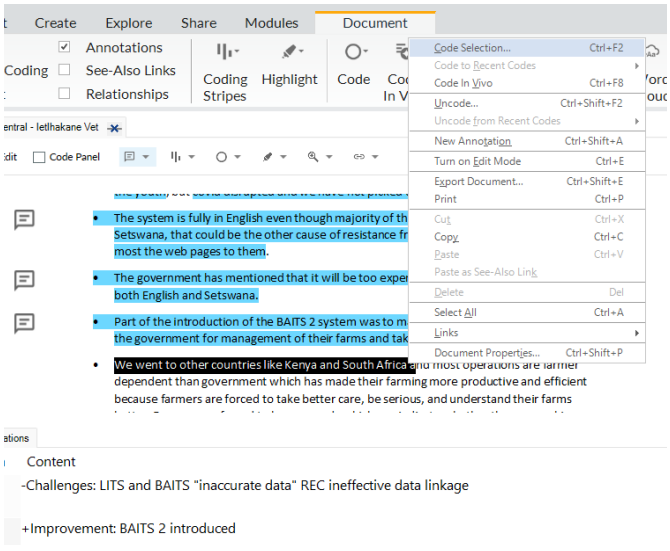
Figure 11 Sample documents imported in the Files subsection of the Data section.

Figure 11 depicts some of the files that were imported into NVivo under the Files subsection. The files are also categorized based on their 9 different districts and the respondent's name beside the district, because of anonymity the names are blocked out in Figure 11. The study had a total of 66 respondents taking part in this study. The researcher used color codes for information categorization marking as illustrated in Figure 11 to the right of the page. The color attributes corresponding to the characteristics were grouped according to the districts of each transcript. This is useful when conducting cross-case queries and coding matrix analysis of the data.

B. Annotations and Memos



(a) Sample of annotations of an interview transcript.



(b) Preview of new Annotations feature under Notes section.

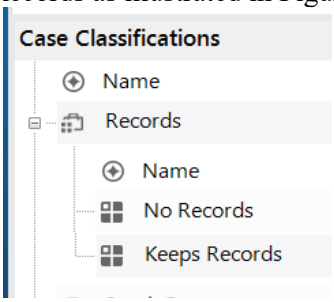
Figure 12 Data Annotations

Memos allow for the separation of analysis made and material being analyzed. They are a storage for reflections and ideas that are related to each interview document [177]. Maureen [185] describe memos in NVivo to be documents that give provision to write ideas and interpretations, understand that it is growing, and gain insights. The memos usually evolve and end up being part of the write-up of thesis findings and discussion chapter. Dollah [177] describe an annotation as being a connection to the comment written by the researcher creating a link to the precise data in a file. Figure 12.b shows the memo and annotations of the animal traceability project. Unlike the NVivo version used by Maureen [185] where the memo is a subsection of sources. NVivo 12 memos is a subsection of the Notes section along with subsection annotations as presented in Figure 12.a.

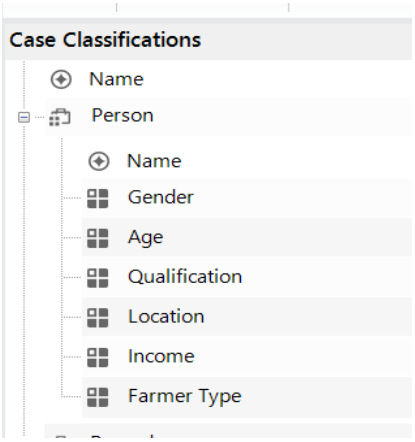
Memos and annotations were used in this research study as can be illustrated in Figure 12.b. The 66 interview transcripts were analyzed easier benefiting from the memos and annotations. They are also used to keep important information in the footnotes where they can be easily accessible. The coding cycle descriptions were noted as annotations in memos. Figure 12.b illustrates how a relevant section of the document is highlighted and annotations made in the footnotes. The text is highlighted, right clicked, and the drop-down menu appears then the code selection option is selected. The user then chooses the case that the extract should belong to.

C. Attributes, Values, and Classifications

According to Tonin et al. [188] NVivo attributes are created to contribute to the accuracy of structural elements which assist in the differentiation of one concept from another. Taking into account what the concept is and also the concept representation. Attributes are characteristics of a source aspect or the node which has or will have an impact when analyzing data [189]. The study research project animal traceability.nvp in NVivo creates multiple cases. A sample extract of the created case person is shown in Figure 13.a and it contains attributes that relate to the case person, sex which gives the options male or female and age: 20,25,30, gender, qualifications, and source of income which all have options to select from. Dollah et al. [177] explain how each of the attributes will have a set value and a corresponding classification that is the collective name assigned to the node. In the ontology-driven approach to animal traceability project the attributes correspond to characteristics of the Person and Records as illustrated in Figure 13.a and Figure 13.b.



(a) Sample of attribute created in NVivo of case Records.



(b) Sample of attribute created in NVivo of case Person.

Figure 13 Sample Attributes

Each of the attributes listed has values attached to it. Table 14 lists the attributes of case Person with the values attached to it. These values were used to run queries that are discussed in the next chapter of this study. An overall case classification spreadsheet of the 66 respondents was generated by the creation of a demographics table using Microsoft Excel and importing it into the node classification, as presented in a sample of six respondents in Table 14.

Attributes	Values
Gender	Male, Female

Age	Under 24, 25-34, 35-44, 45-54, 55-60, over 60
Qualification	Master's Degree, Bachelor's Degree, Diploma, Certificate, High School, Below high school
Location (Districts)	Kgatleng, Kweneng, Central, Southern, Kgalagadi, Ghanzi, South East, North East, North West
Income	Full-time farmer, Part-time farmer, and Job, Job, Farm manager
Farmer Type	Commercial, Communal

Table 14 Sample attributes and values of case Person.

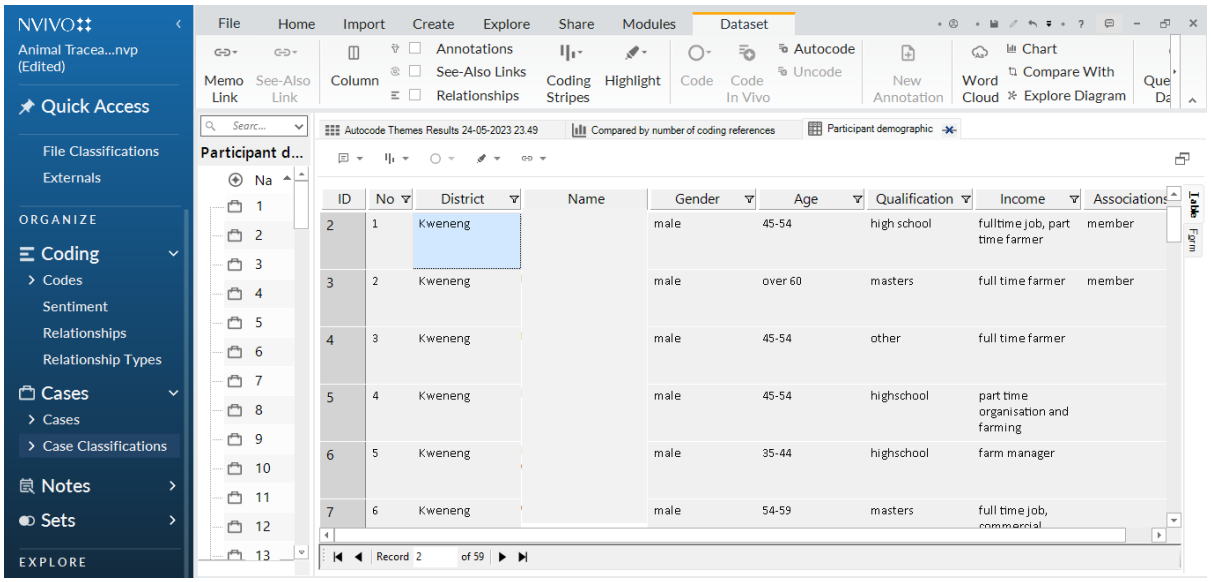


Figure 14 Sample presentation of the imported demographic of respondents.

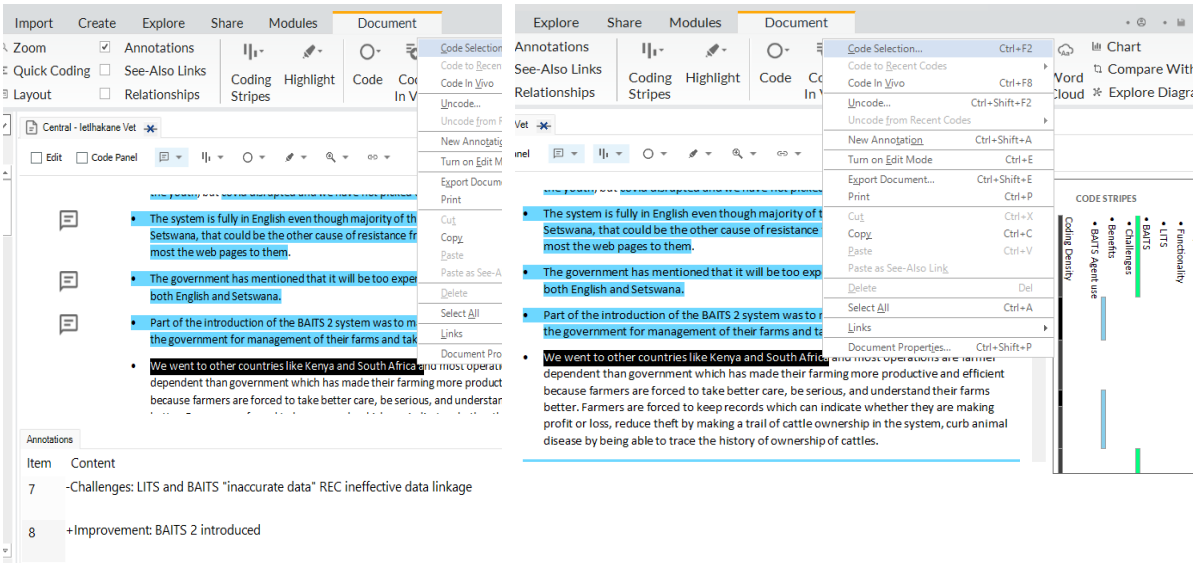
3.9.1.2 Stage 2: Topics

The second stage of analysis in NVivo is the deduction of content and themes that will illustrate patterns in the text data and provide a better understanding of the research topic. This stage is crucial for extracting concepts that will be used for formalization in the development of the proposed ontology-driven framework. Stage 2 includes coding the interview material to group related concepts under containers called nodes. Coding was done both in the software and manually following coding text data guidelines [186]. The main objective of this topic is to answer research questions one and two, which are looking for user requirements and components for framework development.

A. Identifying Topics

According to Saldana [186] a code is a short phrase or word in a qualitative inquiry that significantly attaches a summative attribute to a language or visual portion of data. The coding process in NVivo occurs when the researcher right-clicks a word or sentence and selects code as illustrated in Figure 15.a. The data to be coded usually comes from interview documents, field notes, photographs, journals, internet sites, literature, and email correspondence [186].

After the material has been coded NVivo provides a function to highlight the text selected and assign a specific colour to the coding stripes. These denote the topics that are obvious and originate from nodes formulated [185]. The coding stripes are displayed on the right side of the transcript showing the node that the content coded relates to. The initial themes and concepts were generated from the interview transcripts in this study with efforts from coding stripes shown in Figure 15.b. Hutchison et al. [190] found the coding stripes function is exceptionally useful in early conceptual development. He further states that coding stripes offer a good visual overview of the way coded nodes may relate to each other. Coding stripes allow visualization of sections of text or documents to see the additional nodes coded to the specific text selected.



(a) Sample of coding of selected text in an interview transcript.

(b) Sample of coding stripes of an interview transcript.

Figure 15 Sample for coding

B. Creating initial nodes

Before the creation of the initial nodes coding was first done manually following the guidelines from the book published [186]. Saldana [186] recommends a two-cycle coding process that should be performed on the data collected (interview transcripts) before moving electronically. The first cycle of coding is performed on a part of data that ranges from one-word, full paragraphs, to a whole page [186]. In the second cycle of coding the coded portions can be the same, increased or reduced passages, or reconfiguration of the codes. This research employed magnitude and In Vivo coding in the first cycle process. The eclectic coding was performed first to refine the first cycle before moving to the next cycle. The second cycle of coding the research used evaluation coding for the categorization of the data.

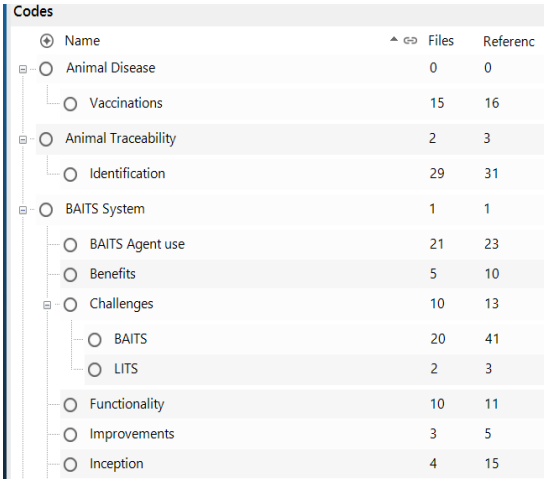
As indicated after gaining experience from doing manual coding and developing an understanding of the data analysis foundations, the knowledge is applied to chosen CAQDAS specifically NVivo 12 software [186]. The CAQDAS software will not perform the coding for the researcher, which is why the manual coding process is first completed. Flick [191] agree that coding for the program is just attaching tags to texts or sections of passage. Furthermore, the software is essential to effectively store, manage, reconfigure, and organize data for human analytic reflection. In the analysis of the research, coding was used to connect items to nodes. Labels with meaning were attached to sections of data and whenever a new concept was found it was tagged and a node developed or stored to an existing similar node. For the generation of queries, it is crucial to form associations of data with their sources, node classifications serve this purpose. The research for an ontology-driven approach to animal traceability created sources and references with the names of respondents and their respective districts. This is to enable query generation [185]. Figure 16.a illustrates the first initial nodes created after manual coding was completed. Figure 16.a shows the nodes when they are not expanded, they have sub-nodes which can be seen in Figure 16.b showing an expansion of the nodes into sub-nodes. It includes the files and references the data belongs to, crucial for associations to permit queries to be generated.



Codes

Name	Files	Reference
Animal Disease	0	0
Animal Traceability	2	3
BAITS System	1	1
Cattle count	30	30
Cattle Market	19	19
cattle rustling	3	3
Farming	0	0
Intergration	5	5
Location	43	45
Movement	17	17
Records	1	1
Suggestions	13	14
Tracking	5	5

(a) Sample of Initial nodes created from coding.



Codes

Name	Files	Referenc
Animal Disease	0	0
Vaccinations	15	16
Animal Traceability	2	3
Identification	29	31
BAITS System	1	1
BAITS Agent use	21	23
Benefits	5	10
Challenges	10	13
BAITS	20	41
LITS	2	3
Functionality	10	11
Improvements	3	5
Inception	4	15

(b) Sample of nodes expanded showing sub-nodes.

Figure 16 Sample for nodes

3.9.1.3 Stage 3: Analytic

The analytic stage of the research is important for the content and thematic analysis of this study. This is the main area where merging of the initial nodes created in stage 2 takes place and running the NVivo queries to get visual results. Maureen, [180] state that the analytic stage extensively explores the nodes' complex features. The merging of nodes to make hierarchies started by moving, merging, and where necessary renaming the initial nodes developed in the previous stage. The hierarchy arrangement is done for a better analysis of the coding with the use of queries. After working on the nodes, the final structure of the layout hierarchy is generated and the final node structure of this research study incorporates nodal hierarchies Initially the project started with 13 main nodes and 32 sub-nodes, which is a total of 45 nodes.

A. Data Sets

For the extrapolation of themes of high order from categories that are lower, the creation of sets has to be carried out states O'Neill [185]. Furthermore, [176] explains the creation of data sets is performed for the relations of 2 nodes that are different from each other. Hutchison et al. [190] concur, theoretical development is greatly advanced by the creation of data sets, and nodes can be grouped with data set tools. The research developed data sets with the use of the themes from the material that was coded. One unit of data is queried using data sets. It enables generation of queries by the study of wider categories. These are used to deduct concepts that are used in the formalization of ontologies.

B. Models and Relationships

At this point of the research, the study will concentrate on utilizing analytically and productively the tools available in NVivo. Focusing on techniques to utilize tools in analysis software will give provision for the facilitation of developing themes, concepts, and categories [178]. AlYahmady and Abri [178] further suggest that the author should also use various numerical data that was found during the study, it can be demonstrated with the use of figures and sketches for the facilitation of reading the findings. Models will be created from the deduction of the data sets developed. According to Shafique and Mahmood [192] the development of models is said to be a research method that is most effective. He further states that it helps in describing, testing, understanding, and predicting complex events [192].

C. Visualization

A visualization tool in NVivo was used in this research. It was used to get a clearer image of the concepts and themes that emerged throghout the analysis of coding and recording. Figures and charts presenting percentage coverage of codes and concepts of the study assist in visualizing each respondent's contribution and extract results from the cases. Jackson and Bazeley [176] affirm that summary visuals of the data presented in diagraphs and charts are developed in database software to provide the vision of data relationships and patterns. A great number of data can be condensed into a single visual page.

D. Using Queries

Sections of the data in the text transcript of the project that are case specific can result in relevant important information using Queries [185]. This study used queries to search for answers based on the research questions. Multiple query options provided by NVivo are found in the explore tab as seen in Figure 17. Queries have provisions for the investigation of relationships among categories and concepts searching in the coded data elements simultaneously. Jackson and Bazeley [176] also asserts that queries are useful particularly at the last stages of analysis to assist in sorting through attributes and node values that were created in the project. NVivo offers various types of queries to use word frequency, matrix coding, project maps, and hierarchy tree maps.

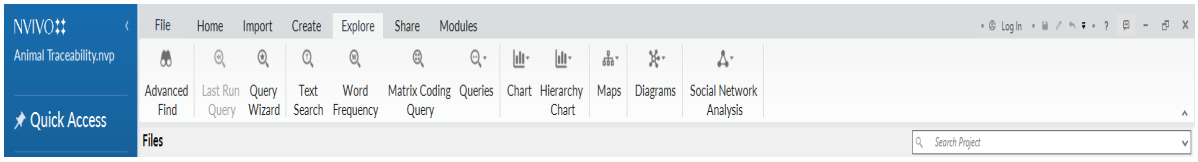


Figure 17 Explore panel with query options from NVivo software.

E. Running queries

The exploration process of complex parts of nodes is described as running queries [185]. Also sorted easier with NVivo [176]. The job of the analyst at the end is to review the data produced from the queries and interpret that data in relation to the research questions. Jackson and Bazeley [176] listed 3 main categories of queries that can be performed: Text mining, Theory-building, and Clerical queries. Text mining queries are used to search text documents for words and phrases that are specific. Category Text mining provides queries word frequency queries and text search queries. Exploring relationships between items is performed by theory-building queries, like the matrix coding query, coding query, and compound query. The last is clerical queries that assist in managing components in the analysis process by providing useful output. The Clerical queries are called the group query and coding comparison query. To run a query for this study the selected query button was pressed, the results were generated, and a figure was generated. The research used the word frequency query, matrix coding query, and cross-case query for deeper understanding.

F. Word frequency query

The word frequency query is used to catalogue up to 1000 words frequently used words in the data provided [176]. Word frequency query along with the word cloud query were used in the research on the interview transcripts. This is a tool used for mapping or exploring words for content analysis. The author further states that the tool is a wide net that can capture a majority of the text like participants' names, numbers, and ages. Word frequency helps the research in deducing word patterns worth exploring to derive the concepts that will be created in the formalization stage of the proposed ontology-driven development.

F. Matrix Coding queries

Matrix coding query is used in the research for its function to provide direct information from the research question's perspective. According to experts Jackson and Bazeley [176], a comparison of subgroups from the database is made with a matrix coding query, and by double-clicking on the cells access to qualitative information is provided. As a mixed research study detailed findings of both numbers and content are valuable results from the query. Results in the form of charts and tubular form are produced in the representation of data, references coded, and node cells with percentage data. The matrix query cross-tabulation of paired items was displayed in matrix form. They assisted in the identification of the multidimensional properties of the study, thus finding answers to the research questions.

G. Cross-case query analysis

According to Maureen [185] the cross-case query is more complex than the other queries. It analyses by comparing and differentiating among and within groups of respondents in the study. The best use is

comparing cases on unique factors and continued refinement on other factors like time, using the filter button to sort specific column items as required [176]. It employs matrix coding query or framework analysis to process data. Cross-case query has the advantage of comparatively focusing on single cases while preserving their uniqueness, unlike other methods that merge and group cases. Accordingly, cross-case queries are employed in the research. The queries were created on the basis of the research questions and findings. The matrix chart produced enabled the comparison of cases looking at specific properties, and then the search was further refined considering and comparing factors from other cases.

3.9.1.2 Conclusion

This is the last stage of the qualitative data analysis where nodes have been created, manipulated, and recreated. With the provision of the query tools by NVivo software, queries were generated, and results were displayed in the form of figures, tables, and graphs. The figures, tables, and graphs are a result of data deduction from the analysis, and theories and concepts will be developed to answer the research questions and deduce components used for the development of the ontology-driven framework.

A. Developing Theories and Concepts

The qualitative analysis of the mixed data analysis was effectively carried out using the CAQDAS called NVivo version 12. The software helped in the organizing of data collected to assist in analysis that will draw reliable and clear conclusions. Manual tasks of analysis are reduced greatly by using the software, this allows the researcher to focus on discovering tendencies, recognition of themes, and conclusions derived with less time [178]. The software provided the study with the option to manage data, produce separate ideas, perform data queries, display visual models, and report the data. AlYahmady and Abri [178] advises researchers carrying out qualitative analysis to pursue the software processes to lighten the vague, lengthy time, and muddled tasks. The research material was well documented and imported into the NVivo tool for constant evaluation, analysis, and synthesis of the information. This was done to derive thoughtful discussion and insightful evaluation of the research findings. With the use of the visuals that were generated throughout the analysis stages. The results deduced in each of the four stages were then interpreted to feed into the concepts for the development of the ontology-driven framework.

3.9.2 Quantitative Data Analysis

Mixed data analysis is a combination of both quantitative and qualitative data analysis, this research carried out quantitative data analysis using descriptive statistics. The initial stages of analysis are the descriptive analysis methods used for the summarization and description of data [146]. It is used for the establishment of a comprehensive guide to where and what are the impacts and causes of the animal traceability challenges. The interview data and the themes generated in the qualitative analysis were used for the descriptive statistics analysis. Microsoft Excel was the software tool used to calculate and generate figures and graphs for quantitative data analysis.

For the summarization of data in an organized manner providing a description of relationships among variables in a sample is described as descriptive statistics [175]. To see the general trend of the data and use the data to fit models of statistical analysis tools of descriptive analysis should be employed [175]. To calculate and evaluate where the center of the frequency distribution lies three measures were employed, mean, mode, and median. Scherman et al. [187] say to further analyze the themes generated it is advisable to include the range of scores, the minimum, the maximum, and the standard deviation. Calculations for both central tendency and range were used in the research which generated graphs and figures.

3.10 Chapter Summary

The research design and methodology discussed the research approach, methodology, and analysis choices that were taken in this study. Each of these choices was explained and reasons were given for each selection. This study chose to use the explanatory research design and mixed methodology which included both the quantitative and qualitative methods. Timing of methods employed the concurrent triangulation design where both methods occur at the same time. Data collection was done using interview and questionnaire tools. The tools were taken through a process of validation and reliability.

Ethical procedures for the data collection tools were prepared, submitted to the ethics committee, and approved. Data analysis for both quantitative and qualitative data was conducted separately using different tools.

The data analysis focused on the methods that were used to conduct analysis of the data collected. For quantitative data analysis, this study chose to use central tendencies from descriptive statistical analysis tools. The central tendencies calculated numerical data using mean, mode, median, maximum, minimum, and standard deviation. Qualitative data used NVivo software for content and thematic analysis. The textual data is sorted, synthesized, and analyzed providing visuals. Pictorial visualization of data is important in summarizing large data and displaying it in a clearer and more understandable way. The results from the data analysis are presented and discussed in the findings and discussion Chapter 4 of this study.

4 CHAPTER 4: FINDINGS AND DISCUSSION

4.1 Introduction

The purpose of this chapter is to provide analysis of the data collected in Chapter 3. The concepts from this analysis will form a basis for the design of the proposed framework and ontology detailed in Chapter 5. This study addresses the lack of frameworks in guiding the development of animal traceability platforms. This study proposes an ontology-driven framework to address the challenges of animal traceability in Botswana. A gap between the stakeholders and the current system in animal traceability was identified. Data collection was conducted at the stakeholder locations to gather information that would answer the study research questions. The two research questions below are addressed in this chapter. The chapter includes the results and discussion of the major findings related to animal traceability to ascertain if the new data supports or is in contradiction with existing information.

RQ2. What are the key components of an interoperable and knowledge-driven framework for animal traceability?

Section 4.2 presents the results of the quantitative research with a descriptive discussion of each item presented. There are subsections to section 4.2 for each statistical figures and tables shown. The next section 4.3 illustrates the qualitative results from the data collection. The results are presented, and a discussion is provided of the results discovered. The last section concludes the chapter.

4.2 Quantitative Research Results – Descriptive Statistics

This section sets out the descriptive statistics of the quantitative research. Descriptive statistics is used to summarize the quantitative data to gain insight by deducing potential trends in animal traceability. The results are gathered from a sample of 66 respondents who were made up of farmers, veterinary officers, matimela officers, the MOA, and BMC representative. The data from 66 respondents who completed questionnaires were captured and analyzed to produce the results that follow.

The quantitave data is used to give empirical evidence on the practical behaviour of the framework. The data is used for supporting evidence-based modelling of the animal traceability framework. The data further informs the animal traceability ontology structure and scope by identifying animal traceability trends and frequent behaviours. Additionally, identifying the ontology concepts and relationships. Quantitave data helped in recognising the animal traceability norms to use in ontology development, such as “cattle identification”, “record keeping”, and “tracking movement”.

4.2.1 Response Rate

The study intended to cover as much of Botswana as possible to produce results that were more district representative. The respondents were classified into the nine districts of the country and the response rate of each district was recorded. The response rate is of livestock rearing respondents and employees who deal with animal traceability who volunteered to take part in the study per district. The districts are as shown in Figure 18: Kgalagadi, Central, Ghanzi, Kgatleng, Kweneng, South-East, Southern, North-East, and North-West. Figure 23 illustrates that Kgalagadi district had the highest response rate from the districts at 23%. Ghanzi district follows behind with a response rate of 18% of respondents. The next district is the Central district with 17% of respondents volunteering to participate in this study. The respondents of the Kweneng district made up 9% of the respondents in the study. Kgatleng district and Southern district had an equal representation of 11%. Likewise, the Northeast and Northwest Districts with 5% equal representation of respondents. Trailing behind with a very low 3% is the Southeast District of Botswana, with the lowest response rate.

The results show that the Kgalagadi, Central, and Ghanzi respondents were more interested and willing to take part in a study to help improve their way of cattle farming, tracking, and traceability. These three areas also showed to have more cattle farms than the other areas. The researcher received a list of farmers from the veterinary office in some of the districts and then randomly called the people to check their availability and willingness to participate in the study. In all district areas, the researcher moved around the farm areas and made door-to-door requests for voluntary participation. For other districts like the South-East and Southern, the researcher had to fully rely on going to farm areas and requesting participation from the respondents in the field. This study found that the South-East and Southern District of Botswana had fewer cattle farms than the other districts which contributed to the low participation rate. The three areas Kgalagadi, Central, and Ghanzi with the highest response rate were districts that had more cattle farms and people that were practicing farming than formal professions. Most of the farmers in the three highest response rates were full-time farmers who relied on farming to sustain themselves and their families.

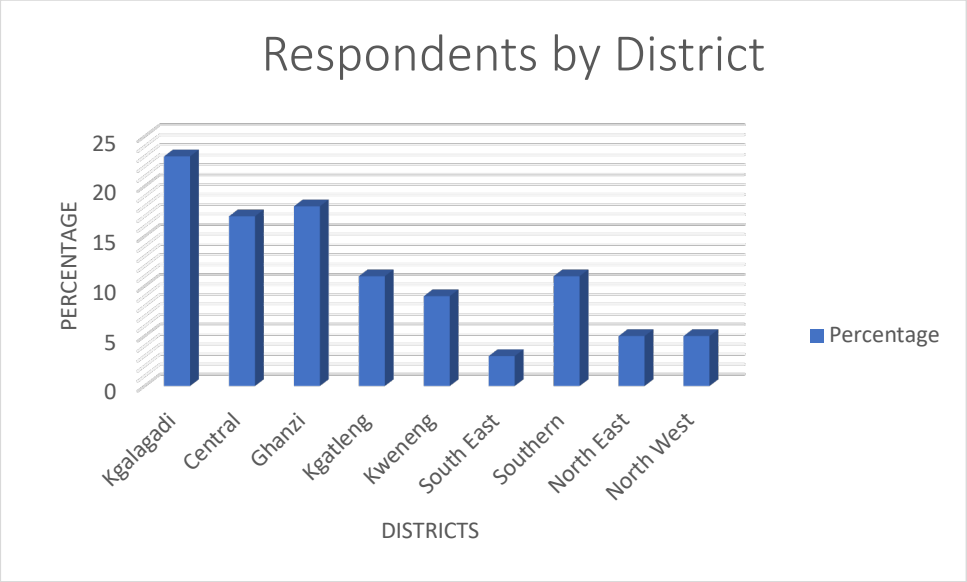


Figure 18 Response rate by district of respondents.

4.2.2 Respondents Demographic

A. Age of Participants

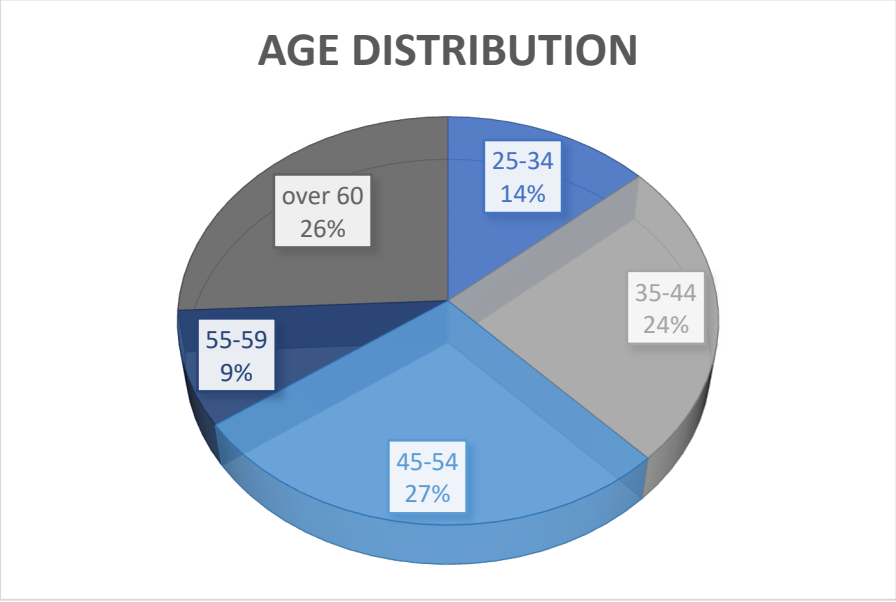


Figure 19 Age of the respondents.

Figure 19 captures the age of the respondents who voluntarily took part in this study. As can be seen, 27% of the research respondents are between the ages of 45-54. This would suggest that the middle age group is more involved in animal traceability practices than the other age ranges. However, this percentage was increased by the government officials (Veterinary, Matimela, MOA, and BMC) who took part in the study. When basing the age distribution on farmers the 26% of over 60 is more accurate. More of the retired population is actively cattle farming, thus using animal traceability practices more than the younger groups. The groups 55-59 at 9% and 35-44 at 24% are the part-time farmers and fully employed population. The part-time farming and fully employed population are the farmers who are at office jobs during the week and go to the farm on weekends. The farmers are usually called weekend farmers because throughout the week they do not know what is happening at their farms, sometimes they skip some weekends to go to work functions and family events. However, they are trying to start and get more accustomed to farming practices before they venture fully into farming.

The youngest age distribution of 25-34 has a low percentage of 14% representation of respondents. This study found that the youngest age group is not actively involved in animal traceability practices because some are still fresh from university and trying to find their paths, while others lack the finances to start any farming practices. The government has initiatives in place to assist young farmers in venturing into cattle farming, but the lack of experience and paperwork processes are still leaving the younger generation lagging behind. The younger group needs to put forward more effort in cattle rearing in the country as the future generation, to contribute to the development and sustainability of Botswana cattle farming.

What can be observed from Figure 19 also is the fact that over 60 age is at the top and most of the respondents saying they are retired and fully committed to farming after retirement is concerning. It is concerning that it seems like the population looks at cattle rearing as a hobby that one ventures into after employment when they are old to have something to do. This also applies to weekend farmers as they practice cattle rearing as a hobby for popularity and backup plans after retirement. Animal traceability challenges stem from these types of observations. Animal traceability technologies cannot be fully applied and used effectively in Botswana because the population that practices more farming is older and technology illiterate. The lack of full commitment to cattle farming leaves deployed systems that are ineffective because the users are not aware the systems are the problem. One respondent believes that the government was to blame for the failing animal traceability because they do not do visitations to their farms. Another respondent said, “We cannot sell or buy cattle as we want with the animal traceability in place because the system is always down”. The respondent said that meant the network, unaware that the system functionalities are also a challenge.

B. Gender of Participants

To ensure a lack of bias in a research study it is ideal for the research to attract 50-50 males and females. However, this was not possible due to the uneven gender distribution of cattle farmers in Botswana. The gender distribution within the study was skewed. The researcher randomly chose respondents by moving around farming areas and other areas using a list from veterinary but there were still more males than females found to take part in the research. A veterinary respondent states that this could be caused by the traditional stereotypes surrounding cattle farming. He said that traditionally Batswana believe that farming is a masculine practice that is supposed to be practiced by men to provide for their families and the women receive the produce to prepare and feed the family. That is why this study had a female response rate of only 16.7% (11) females against a high rate of 83.3% (55) of males as illustrated in Figure 20. The lower female participation at 16.7% highlighted gender-specific barriers in the adoption and management of traceability practice.

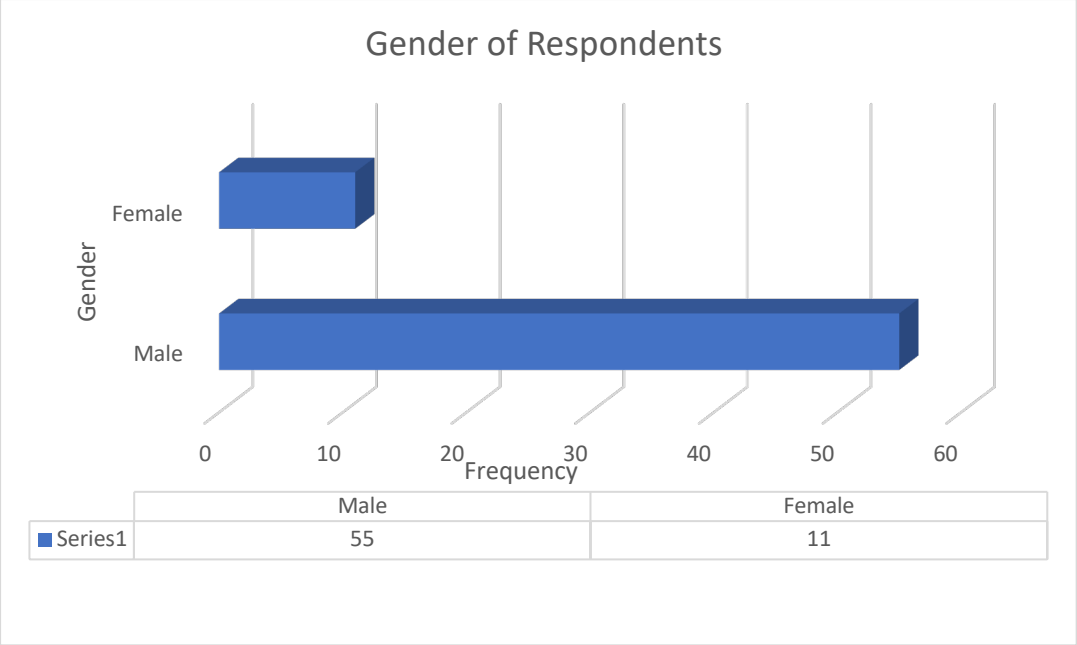


Figure 20 Gender distribution of respondents.

C. Level of Education

This demographic is important in understanding the relationship between animal traceability and level of education. That is understanding the level of education of the people who are using the current animal traceability system, and what future improved systems developed should be aware of during system development. The level of education will also assist in understanding the opinion about whether the current system is good or not and worth using based on the level of complexity. The level of system complexity is based on the functionality of navigating the system, the syntax and semantics of the system, and the language used in the system.

The result of the level of education represented in animal traceability is presented in Figure 21. As illustrated in Figure 21 the greatest proportion by far of the population have qualifications below high school (45%). These results are aligned with the age distribution of over 60 of the research population because the majority of the over 60 have low-level qualifications. Respondents with high school qualifications followed with just 18%, with diplomas at 12% and certificates at 11%. Bachelor's 8% and master's degree 6% make up the lowest percentage for respondents' qualification. The highly qualified participants were respondents employed at various organizations. Only four full-time farmers had qualifications above a bachelor's degree. This shows us that the current systems deployed did not take into account the level of education of the users of the system.

According to a farmer respondent they tried to use the system, but it was too complex, and they failed, which led to them spending money to get their transactions completed. Another farmer stresses that because of their low level of education, they cannot fully understand English, and the system only communicates in English. As stated by a BAITS agent, the complexity of the system has made it impossible for anyone without a diploma qualification to use it seamlessly. This tells us that the level of education not being accommodated by current systems adds further weight to this research objectives.

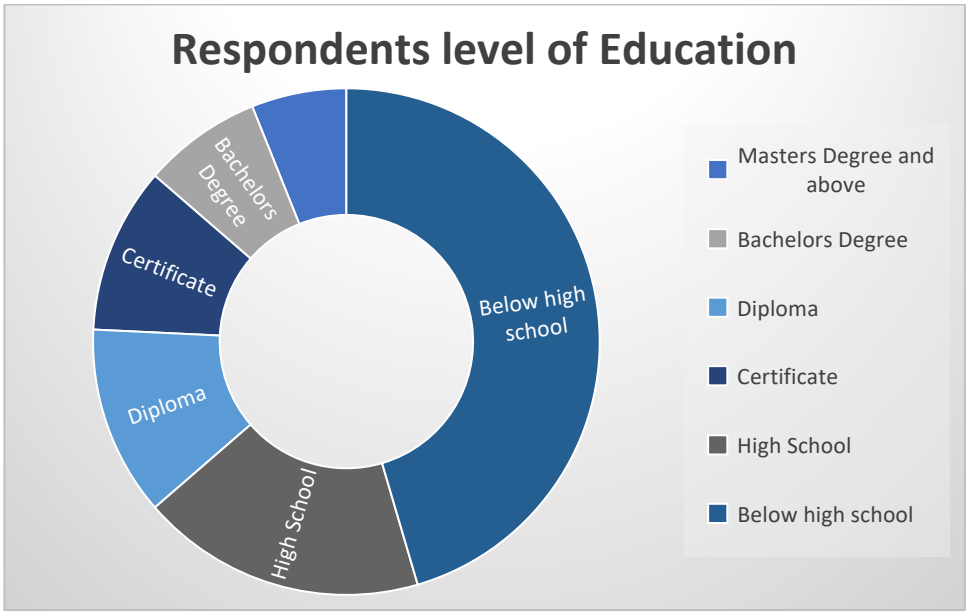


Figure 21 Level of education representation of respondents.

D. Respondent's Source of Income

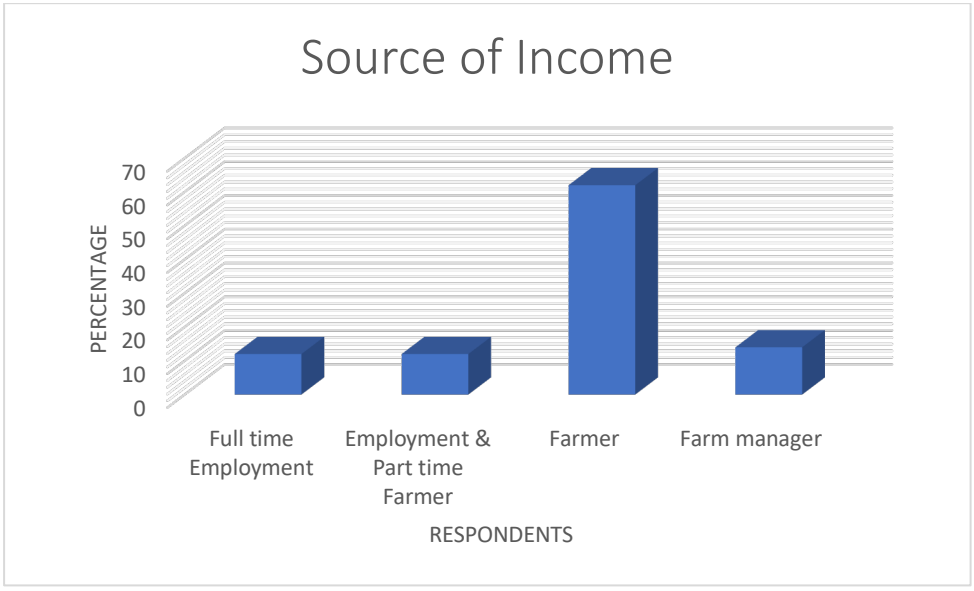


Figure 22 Respondents source of income.

The source of income is an important representation that impacts the country's economy and the social well-being of the community based on cattle rearing. Successful cattle rearing is reliant upon the effectiveness of animal traceability. Animal identification, traceability, and tracking are crucial socially and economically to Batswana farmers. For a farmer to sell or buy their cattle to other farmers they have to use an animal traceability system. According to the BMC representative a farmer makes high profits by selling their cattle to European markets through the BMC. To sell the cattle to the European markets the farmer has to be actively using the animal traceability system to meet the regulations. The challenges in the current and only available system in Botswana for animal traceability recognized by the government hinder farmers from getting the best profits from their produce. Figure 22 indicates that many of the respondents depend on farming to support themselves and their families at a high 62%.

Farm managers who work full-time at the farms make up 14% of the respondents. 76% of people depend on the successful sale of cattle for their family's livelihood.

The other respondents 12% work full-time at various organizations that are involved in animal traceability in Botswana. The other respondents, with 12% of participants are both farmers and employed. This indicates that livestock rearing can be practised in multiple ways, and each is affected by animal traceability socially, economically, and financially. The high rate of people dependent on cattle rearing as a source of income emphasizes the importance of this research in providing an ontology-driven framework that will guide in the development of effective and functional animal traceability systems. The systems will serve farmers in transacting cattle to get high profits to feed their families and increase the economy of Botswana.

E. Respondents Cattle Count

The respondents indicate that they have experienced cattle theft throughout their farming experience. Table 16 reports the different cattle thefts and cattle found by respondents in each district. The number of cattle owned by the respondents is also recorded for each of the districts. From Table 15 results it is apparent that the respondents from Ghanzi have experienced the highest cattle theft at 277 stolen cattle and only found 12. Kgalagadi with the highest count of cattle also has high numbers of stolen cattle at 209 cattle with a low 26 cattle found. 181 is the number of cattle that were stolen from the respondents in the Kweneng district, and they got back 11 cattle. The Kgatleng respondents have suffered the loss of 125 cattle and have found none of them to date.

Respondents from the Southern district experienced theft of 95 cattle with only a return of 3 cattle. Over the years the Central respondents have not found any of their lost cattle out of the 68 stolen from the representation population. Northeast and Northwest districts have experienced theft of 35 and 32 cattle thefts and none of the cattle were found. The Southeast district has experienced the lowest theft of only 1 cow, and none found, this is also because they have the lowest number of respondents. Cattle theft is the highest concern of farmers when it comes to animal traceability, they would like a tracking mechanism alongside the traceability system. The respondent's request addresses the first objective of this study, finding the requirements for animal traceability that will be incorporated in the ontology-driven framework.

DISTRICT	CATTLE THEFT	CATTLE FOUND	CATTLE COUNT
Kgalagadi	209	26	8620
Central	68	0	454
Ghanzi	277	12	2052
Kgatleng	125	0	115
Kweneng	181	11	1047
South East	1	0	24
Southern	95	3	249
North East	35	0	120
North West	32	0	150
Total	1023	52	12831

Table 15 Cattle count, theft, and found distribution

.

F. Respondents Cattle Tracking

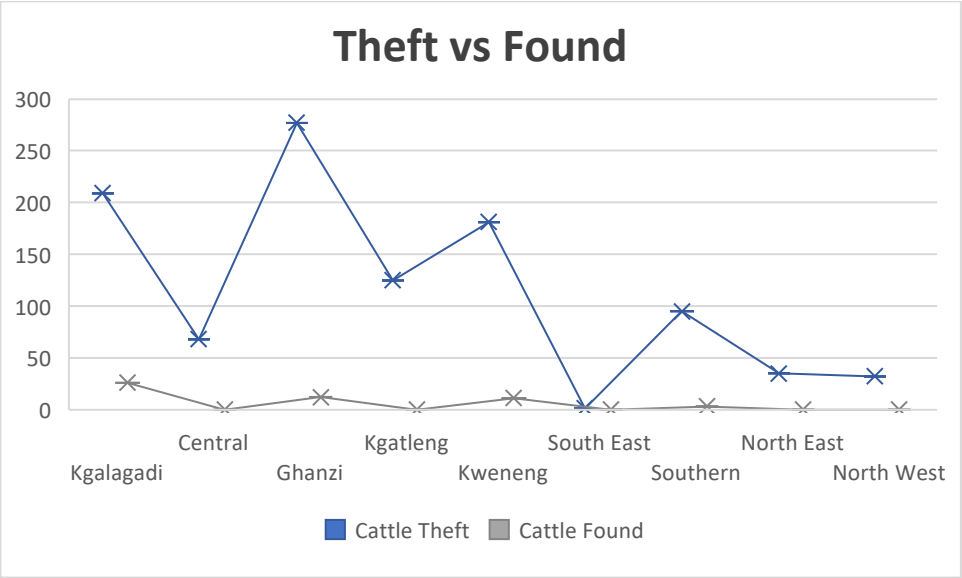


Figure 23 Cattle theft vs cattle found by district.

The MOA representative explained that the type of animal traceability that is conducted in Botswana by the BAITS system is a paper tracing of cattle from farm to fork. They do not perform any form of cattle tracking. For animals to be sold and profit made they have to be present in the farm, protected, and their physical location known at all times. Animal tracking is an important aspect of animal traceability because if there are no animals then there is nothing to trace. Batswana farmers have no mechanism for tracking their animals. One of the farmers described their tracking as opening for the animals to go grazing in the forest then in the evening or next morning they wait for them to come back. If some of the cattle do not come back, then the herd man walks around the farm area to look for them and inquiries from the neighbors. The farmers use the traditional way of tracking cattle by word of mouth, which has proven to be ineffective over the years as cattle theft rises.

Each district experiences cattle theft at different rates. Figure 23 depicts the different rankings of cattle theft versus cattle found by each district. The Blue line shows the rate of cattle theft showing Ghanzi to be the highest and Southeast to be the lowest. Cattle that have been found are shown by the grey line depicting Kgalagadi to have the highest number of cattle found and 4 districts being Central, Southeast, Northwest, and Northeast with the lowest cattle found. The results explicitly show that the paper trail animal traceability system by the current systems is not enough to effectively manage cattle production in Botswana. This is more evident from Figure 24 which captured the overall impact of cattle theft in Botswana at a rate of 95%, versus the low 5% cattle found rate. This study suggests an ontology-driven framework that will incorporate animal tracking devices to curb cattle theft in the country that is causing distress to farmer's lives.

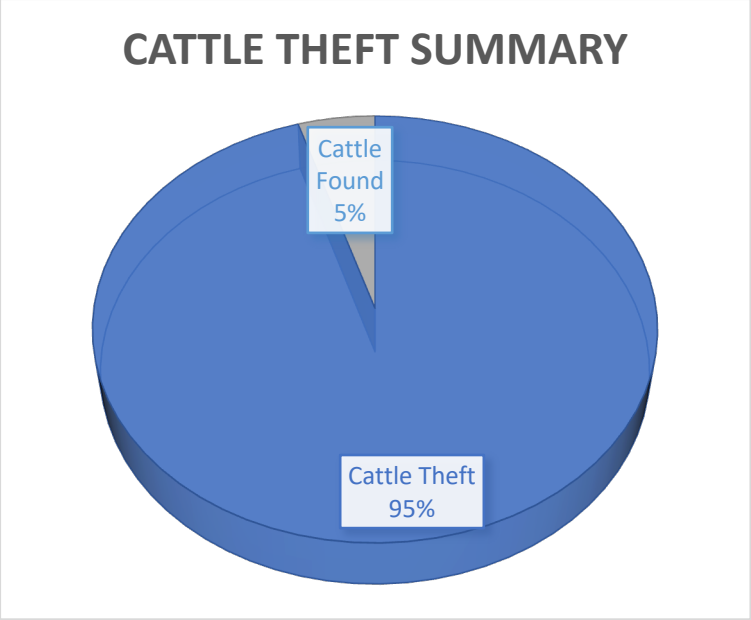


Figure 24 Cattle theft vs cattle theft in Botswana.

G. Respondent's outlook on animal traceability

The findings from the collected and analyzed quantitative data centered on respondents' perceptions of the challenges related to animal traceability in Botswana. Participants were asked to rate their perceptions on various aspects of animal identification, tracking, and traceability methods as illustrated in Figure 25. More than half of participants (55%) indicated that the current methods used by farmers are inadequate, despite 42% making efforts to address these issues. Farmers' associations, established to support various agricultural activities including traceability, were criticized for their ineffectiveness; 56% of respondents expressed that these associations were more about talk than action.

The Matimela program, a government initiative designed to manage stray cattle and prevent theft by housing unclaimed animals in kraals, was also viewed negatively. A significant 85% of respondents believed that the program was ineffective, mainly due to the policy of auctioning unclaimed cattle after 15 days without offering compensation to rightful owners. This has led to dissatisfaction among farmers who often found their cattle sold off without proper notification.

Regarding the BAITS, only 17% of respondents found it useful for identifying cattle, especially in cases where ear tags were still intact. However, 39% felt that BAITS was largely ineffective because enforcement officers and Matimela program administrators rarely utilized it to trace cattle ownership. Consequently, 67% of respondents believed that law enforcement was not doing enough to combat cattle theft. Furthermore, 44% of respondents expressed uncertainty about the usefulness of BAITS, viewing it primarily as a tool for purchasing ear tags and transferring cattle ownership rather than a comprehensive traceability solution.

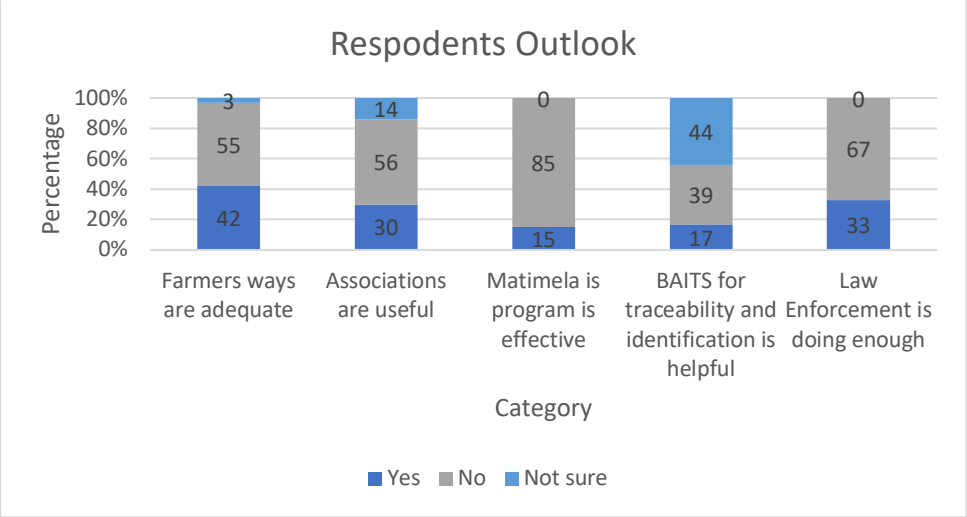


Figure 25 The outlook of respondents to animal traceability challenges.

4.2.3 Central Tendency of Animal Traceability and Tracking Effects

Central tendency is the value that depicts the whole series of data as one measurement. The three fundamental measures of central tendency are the mean, median, and mode. Mean is the total sum of values in a dataset divided by the total number of observations. The median is the value in the middle of the distribution when data values are ranked in order from the highest to lowest. Mode is the most common value in a dataset [65]. Table 16 presents the data set results computed using the measures of central tendency mean, standard deviation, and median of the total number of cattle affected by the challenges of animal traceability in each district. Mode was not used because there was no common value in the data set. The maximum number of cattle affected from a district is 277 from Ghanzi. The maximum number of 112 is also from the Ghanzi district. The minimum of affected cattle is represented by South-East with 0. The results are affected by the low rate of respondents from the district. A mean of 39.6 is captured from Ghanzi, with a joint median of 35 from Kgalagadi and Ghanzi. 34.0 high standard deviation from Ghanzi district. The results are affected by the rate of respondents, as the more respondents answering the questionnaire the higher the challenges recorded. These values are out of a total of cattle owned at 8620 from all districts. The central tendency guides the animal traceability framework by modelling the key common practices. Such as the main districts affected ineffective animal traceability causing high theft rates. This will help define property restrictions to frequently update records in the ontology development process. The data will help to prioritise the domain areas of animal traceability, like the acceptable delays for animal movement reporting (property :reportWithin) constraints.

Number	District	Total	Max	Min	Mean	Median	SD
1	Kgalagadi	209	50	14	34.8	35	13.2
2	Central	68	30	3	17.0	17.5	9.7
3	Ghanzi	277	112	10	39.6	35	34.0
4	Kgatleng	125	50	2	17.9	17	14.2
5	Kweneng	181	70	2	30.2	17.5	28.8
6	South East	1	1	0	0.5	0.5	0.5
7	Southern	95	35	10	19.0	10	11.1
8	North East	35	15	10	11.7	10	2.4
9	North West	32	20	10	15.0	15	5.0

Table 16 Central Tendency statistics for animal traceability.

4.3 Qualitative Research Results

This section presents the results from the analysis of the data collected with the use of NVivo 12 software. In-depth interviews with respondents were conducted to find answers to the research questions. Criteria were set for the selection of respondents using purposive sampling. The aim was to get 60 respondents, but the study ended up with 66 respondents. The results from the interviews, after data analysis was performed, are presented and discussed in this section.

The animal traceability knowledge is not available explicitly, but is available in the form of interviews, experience practices, and documents. The NVivo tool was crucial in carrying out the data analysis of the data to explicitly provide the knowledge needed for the development of the animal traceability ontology. NVivo organized and categorised the data to easily identify the concepts, themes, patterns, and relationships that recur. By coding and querying in NVivo, concepts that are related in practice were revealed, like an animal being vaccinated by a herd boy. The relationship is translated in the animal traceability ontology into an object property. The key entities that the stakeholders mentioned often were tracked, such as cattle, movements, eartags, and tracking. Which guides the structure of the ontology by prioritizing the areas most problematic for the stakeholders. The main terminologies and concepts relevant to the domain were also extracted by NVivo to be used when building the glossary of terms of the ontology. Ontology validation can be supported by the codes and themes from NVivo being used to design competency questions, which will be used to check if the ontology answers the specific domain questions.

4.3.1 Themes and Concepts

The data analysis performed in Chapter 4 uncovered patterns and themes that emerged to be prominent in animal traceability for the identification of relevant concepts and relationships to design the ontology-driven animal traceability framework, and further develop the animal traceability ontology. According to Loeb et al. [193] there is a variety of visualizations available to use in highlighting and presenting the main findings in research data. She further states that the objective of visuals is to assist the readers in better understanding the main data patterns. The results of this study are summarized and presented below in the form of tables, graphs, and figures. These are described and discussed to provide further details of the results. The interpretation of the data patterns using NVivo methods of coding, case classifications, word frequency, word cloud, tree map, and coding matrix. The first results are the pattern interpretations from NVivo software.

A. Word Frequency

Word Cloud

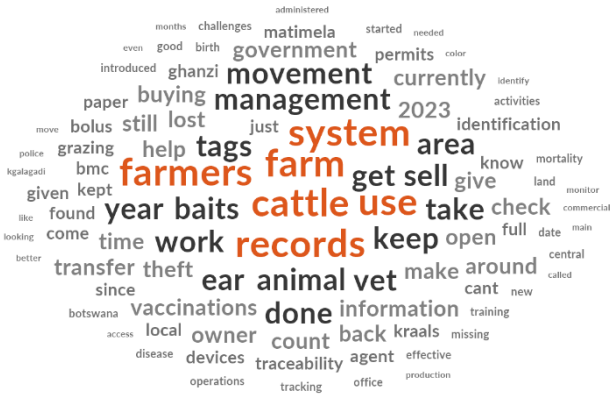


Figure 26 Word Cloud for animal traceability data.

According to DePaolo and Wilkinson [194] a word cloud displays a visual representation of words in text data that are used most frequently identified as keywords, viewed in different colors and sizes. Additionally, a researcher can see unique patterns of phrases and words by conducting an examination of the word cloud frequencies or the lack of patterns in the data. This study used NVivo software for data analysis and one of the results was the word cloud as depicted in Figure 26. NVivo filters the most common English words (as, a, it, and I) and they are excluded from the results. The word cloud in Figure 26 displays a graphical illustration of text that represents respondents' knowledge of animal traceability.

The pattern of words investigated from the word cloud representing responses showed the understanding of animal traceability key concepts by respondents, also identifying common mistakes that have been made. The size of the words show the frequency of the concepts, the largetse being the most frequent concept.The main key concepts identified in the cloud are cattle, farmers, farm, system, use, and records depicted in orange with the largest font from the other words. The key concepts that will be useful in the development of the ontology-driven framework are farmers, farm, system, and records. The cloud clearly shows that the main stakeholders in animal traceability are the farmers; as such, the framework should revolve mainly around them. Records are another prominent concept that is illustrated that should be explored further when added to the framework. The concept of records aligns with the feedback from veterinary and BMC respondents, who stated that records are the main aspect a farmer has to maintain and produce to show effectiveness from farm-to-fork cattle rearing. The word Farm represents farm management, concepts included in the ontology-driven framework that are related to good farm management, will result in animal traceability platforms that will yield high cattle production.

B. Cluster Analysis

Cluster analysis is used in this study to divide the text data into useful and meaningful groups. Cluster analysis is used in this study to divide the text data into useful and meaningful groups. The coding similarity of the cluster analysis procedure in NVivo carries out complete-linkage hierarchical clustering. A table is built with rows of files and codes (appearing in the diagram), and columns are the characteristics chosen to cluster by. Characteristics include word similarity, coding similarity, and attribute value similarity. The two data sources' i,j similarity index is computed between each pair of items (rows) in the table using:

$$sij = 1 - dij / \max(dij)$$

Where: dij is the distance between data sources i and j.

Equation 1 Cluster analysis algorithm

The Pearson correlation coefficient (-1 = least similar, 1 = most similar) is chosen considering that it is scale-invariant, and also appropriate for content coded to multiple themes. NVivo conducts sequential collections until the cluster requested cut-off, at a default of 10, is reached. When the cluster cut-off is reached, a display of a dendrogram using an iterative multidimensional scaling algorithm is provided.

The cluster analysis in Figure 27 is produced by NVivo using an algorithm (equation 1) that group similar animal traceability concepts that are based on co-occurrence. The groupings created by NVivo are listed in Figure 27. Figure 27 is a large picture of many connected and related groupings that NVivo found meaningful and useful. The concepts are the words, cases, and nodes from the data sources. The clusters reveal similarities and relationships between the concepts and themes, also identifying the data structures. These will be used as the foundation for the animal traceability ontology conceptual modelling of classes and subclasses. The classes and subclasses that will be defined in the ontology development processes are "Building glossary of terms" and "Taxonomy".

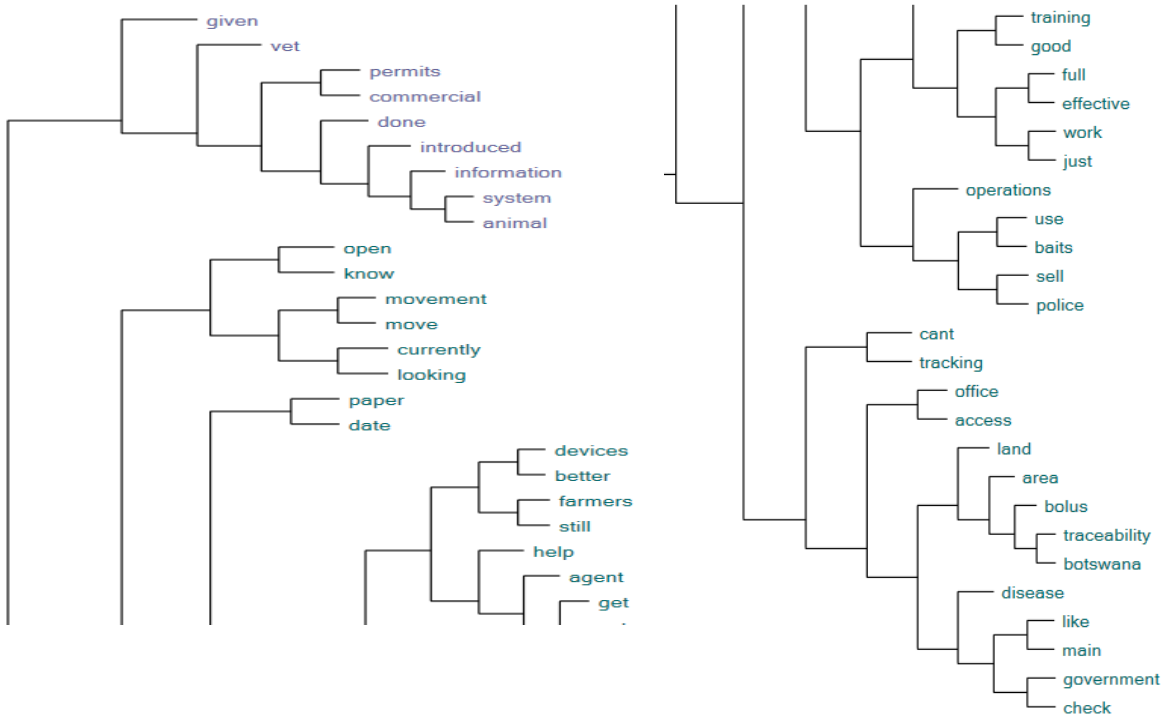


Figure 27 Cluster groupings of animal traceability data.

Validation of the generated clusters by NVivo was carried out in multiple ways. There was tracing to the source data by re-examining the cluster automatically generated in the original transcripts. The manual checking uncovered matches between the automatically generated clusters and participants own phrasing. The initial thematic codes were participant checked by reviewing some participants, and they confirmed that the relations captured participants viewpoints accurately. The steps taken to validate the cluster analysis animal traceability output results are important for moving the analysis from tool output to researcher mediated interpretation. For example the feedback of clusters (Figure 27) like operations, which formed BAITS, Sell, Police, and Use is aligned with the responses from the veterinary and MOA participants about their practices in animal traceability. Second cluster Tracking, which encapsulates access, land, area, traceability, government, and disease was found to be in relation with the farmers responses to animal tracking questions. The third interesting cluster is vet, permits, information, system, and animal which related the animal traceability operations betwen the farmer and veterinarian. The last cluster identified includes devices, farmers, help, and agents which is more aligned with farmers and BAITS agents. These clusters will contribute to the conceptualization stage of ontology development (Section 5.3.2.5).

C. Coding Matrix

Coding matrix is a method based on coding and interpretation of data transcripts that originate from in-depth interviews [197]. He further states that the process of coding outcomes creates groups of categories along with several steps that are predetermined. The interpretations from the respondents using the matrix were made to answer the research questions. Coding was done on the animal traceability interview transcripts to synthesize the content in the analysis phase of this research. The cross-tabulation Figure 28 is the animal traceability results grid that shows the cases that are coded to which nodes, depicting density, presence, and frequency of the animal traceability coding intersections.

The NVivo coding matrix is a visual tool that was used to display the intersection between the different nodes and other attributes such as demographic groups, cases, and themes. Patterns are shown in the qualitative data by counting how many times certain codes co-occur within the same dataset elements, that is, the interview transcripts, documents, and reports. The coding matrix algorithm works on a binary

and frequency logic, operating in the internal data indexing system of NVivo. A representation of the algorithm is as seen in equation 2.

```
For each (Data Unit):
   If (Node A is coded) AND (Node B is coded):
      Increment cell (A, B) by 1
End If
```

Equation 2 Coding matrix algorithm

The algorithm starts by using the dataset to index all coded references; there is a link between each code and text segment. The query compares the two dimensions: rows (codes) and columns (other codes, attributes, and cases). The computation by the algorithm checks for each document, does the node A and node B occur both. Then counts how many times the co-occurrence happens in the same document. The results as seen in Figure 28 shows each of the cells containing a number that represents the count of intersections. How repeatedly coding a row node and column node happen in the same content unit.

When coding animal traceability interview transcripts about the experience of farmers with various aspects of traceability systems. A comparison of what different case attributes have said about an issue. The rows were the cases and columns aspects of traceability systems (animal, cattle, farm, system, and tags) as seen in Figure 28. The interview transcripts reference concepts and the cell shows a number, for example Figure 28 displays case attribute 1 (farmer) the matric shows the farmer frequently links to "farm 9" times and "tags 6" times.

The objective of the coding matrix is to find confirmation of the conceptual model contents and structure of the study based on academic knowledge [197]. This applies to this study as the Figure 28 results were further used to produce clearer visuals illustrated in Figure 29. The illustration was produced based on the main coded items: animal, cattle, farm, system, and tags.

	A : animal	B : cattle	C : farm	D : system	E : tags
1 :...	2	3	9	1	6
2 :...	0	3	1	0	2
3 :...	2	0	3	1	1
4 :...	0	1	1	0	2
5 :...	0	0	0	0	1
6 :...	2	3	1	0	4
7 :...	1	1	6	0	0
8 :...	0	2	2	0	2
9 :...	1	1	6	0	2
10...	0	1	2	0	3
11...	0	1	0	0	1
12...	0	0	0	0	2
13...	0	0	2	0	2
14...	1	0	0	0	2
15...	1	1	4	0	3
16...	0	0	1	0	2
17...	0	0	3	0	0
18...	1	0	2	2	2
19...	1	1	3	2	0
20...	1	2	2	0	1
21...	0	0	4	0	1
22...	0	1	2	1	3
23...	0	0	4	0	1
24...	0	1	1	4	4
25...	0	0	1	0	2
26...	1	0	4	0	1
27...	0	0	1	0	0
28...	1	0	5	1	2
29...	0	1	0	0	1
30...	0	3	2	0	1
31...	0	0	0	0	2
32...	2	2	5	0	3
33...	1	0	1	0	0
34...	1	0	3	4	0
35...	0	1	2	1	1
36...	2	2	8	1	1
37...	5	1	9	3	2
38...	0	0	1	0	1
39...	0	0	2	3	0
40...	0	1	1	0	2
41...	0	1	3	0	0
42...	0	1	0	0	1
43...	0	0	2	0	1

Figure 28 Coding matrix of the five main concepts deducted from the text data.

The validation of the generated coding matrix, Figures 28 and 29, by NVivo was done through the interpretation of the patterns. It is important to question if the pattern is a reflection of a genuine thematic link, for example if the case attributes and nodes (animal, system, tags, records, and BAITS system) capture both animal traceability specific knowledge and general knowledge. An evaluation of the importance and frequency was also considered; issues may present lower frequencies but carry disproportionate influence on the success of traceability. That is why this study did not solely rely on the matrix frequencies, risks of underestimating low-frequency but high-impact factors. Figure 29 presents "cattle count" as one of the high frequencies. However, the number of cattle a farmer owns is not a high-impact factor in animal traceability. The frequencies "animal traceability", "location", "cattle market", "records", and "BAITS system" computed show a genuine reflection of key factors in animal traceability.

NVivo software generated the coding matrix in Figure 29 from the results of the data file illustrating better visuals from Figure 28 and the results were interpreted. Respondents complained about systems being created in the city and not considering the environment that systems will be working in. The current systems need internet access to function which is rarely or not at all available in most farms. However, as satellites reach even the extremely rural areas future systems should consider using them for connectivity. Tags are also shown to be dominating in the coding matrix figures they are a crucial property of animal identification and tracking. Respondents expressed that they are currently using the mandated ear tags bought from the government and personal tags that they make for themselves. Personal tags are produced because the government ear tags are lacking in information that is needed for animal identification. The veterinary representative pointed out that the ear tags have RFID that can be read using special equipment and information registered about the cattle will appear. That is a feature that is not being put to use because farmers cannot afford to buy the equipment, and some do not know about it. This study suggests tags that do not need special equipment to read, that are satellite-based and the information goes straight to users' system, tags that encompass both identification and tracking of cattle.

The results of Figure29 are also a visual produced from the coding matrix in NVivo software , but it has identified other items of equal importance from the in-depth interview transcripts. The main items identified by the results of Figure 29 are cattle rustling, records, animal traceability, BAITS system, location, cattle market, and cattle count. Farming respondents expressed their dismay and pain of cattle theft that negatively affects them and brings down their cattle production. They explained how most of their farm areas are not secure, because their farm area does not have enough grazing land they have no choice but to let their cattle out of their farm into the open grazing land to feed. Cattle in the open fields make them vulnerable to thieves and they have no methods of monitoring the cattle while out. Animal traceability is the main concept of this research and the contents that are involved are as seen in Figure 29 records, location, and cattle count. For the profitable sale of cattle, a farmer needs to have data records that are up to date, meet the profitable market requirements like BMC, EU regulations, and effectively use an animal identification, tracking, and traceability system.

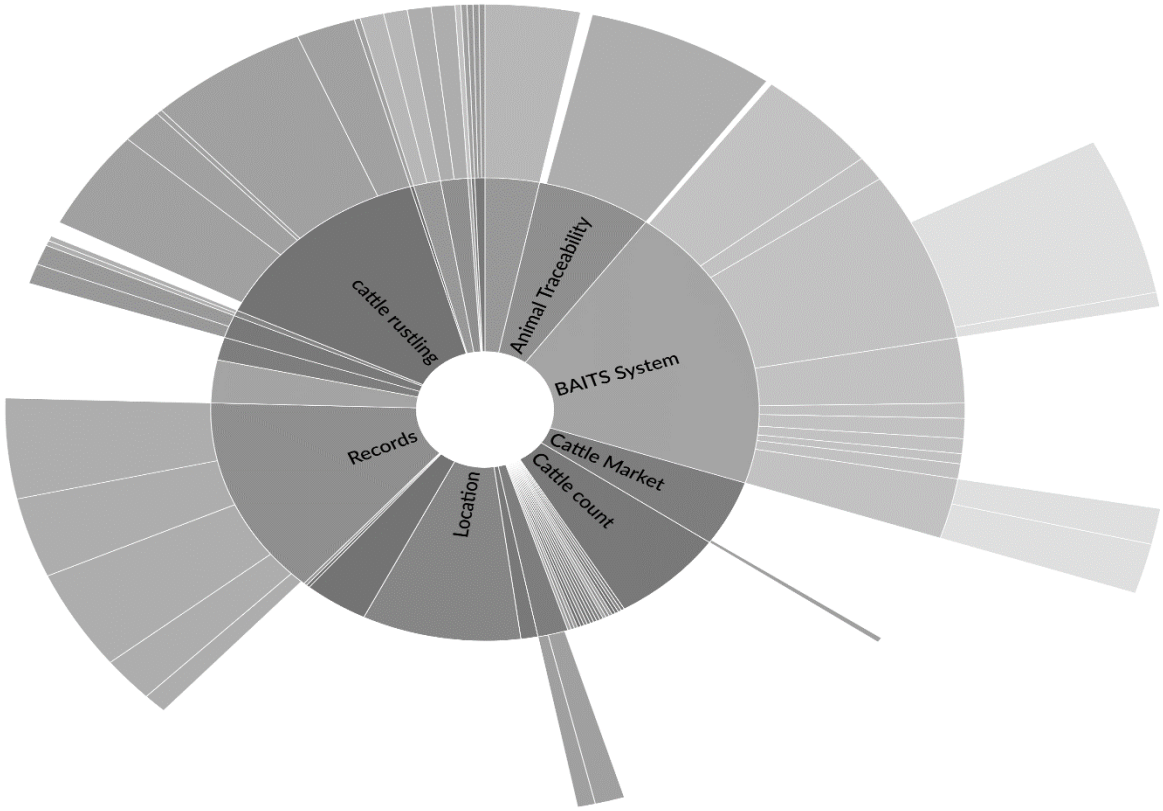


Figure 29 Coding matrix illustration of interview results.

D. Hierarchy Tree Map

Further exploration of the data file using the codes created was done using the tree map query in NVivo software. The coding hierarchy provides a visualization of the animal traceability node structure as a nested and size-weighted image. It reveals the parent-child relationships of the nodes, a thematic hierarchy. This is crucial in showing how the animal traceability themes are structured and where emphasis on ontology concepts should be when developing the ATO. The hierarchy tree map is specifically useful when refining the taxonomy of the ATO, when determining the top-level classes and their subclasses.

The hierarchy tree map visualises a hierarchy, an NVivo interactive diagram displaying the relative size and structure of coding between nodes. The data in the hierarchy shows a set of nested rectangles that vary in size, as seen in Figure 30. Representation of how the data is distributed across parent and child nodes. Figure 30 shows each rectangle representing a node; the size of each rectangle corresponds to the number of coding references that are relative to others. The tree map analyses the animal traceability hierarchies in a single view and compares relative sizes of the differing aspects of the data.

The hierarchy tree map was computed using the proportional area mapping algorithm based on space-filling visualisation. Equation 3 computes node sizes based on the derived number of coding references. First, the total references are calculated, then the area proportion is computed. The results are used to render the rectangle sizes proportionally and apply recursive subdivision for the child nodes.

For every parent node:
Total references = Parent References + Sum of Child References.
Then compute the area proportion:
Area proportion = (Node References / Total References in the dataset).

Equation 3 Hierarchy tree map algorithm

The tree map, as seen in Figure 30, appears as horizontally oriented rectangles which are subdivided by size into major categories and their subcategories. The display conveys the part-to-whole relations. Jadeja and Shah [195] state that hierarchical data is presented on 3D or 2D tree map displays based on attribute values. Furthermore, the users are provided with visualizations so they can manipulate the categories and save each subtree. Each of the categories of the tree is displayed in terms of levels in the hierarchy; the bigger the size of the category, the higher level it is. This study has expanded the main tree Figure 30 and manipulated it to display further sub-trees as shown in Figure 31 of the system, Figure 32 farm, and Figure 33 cattle.

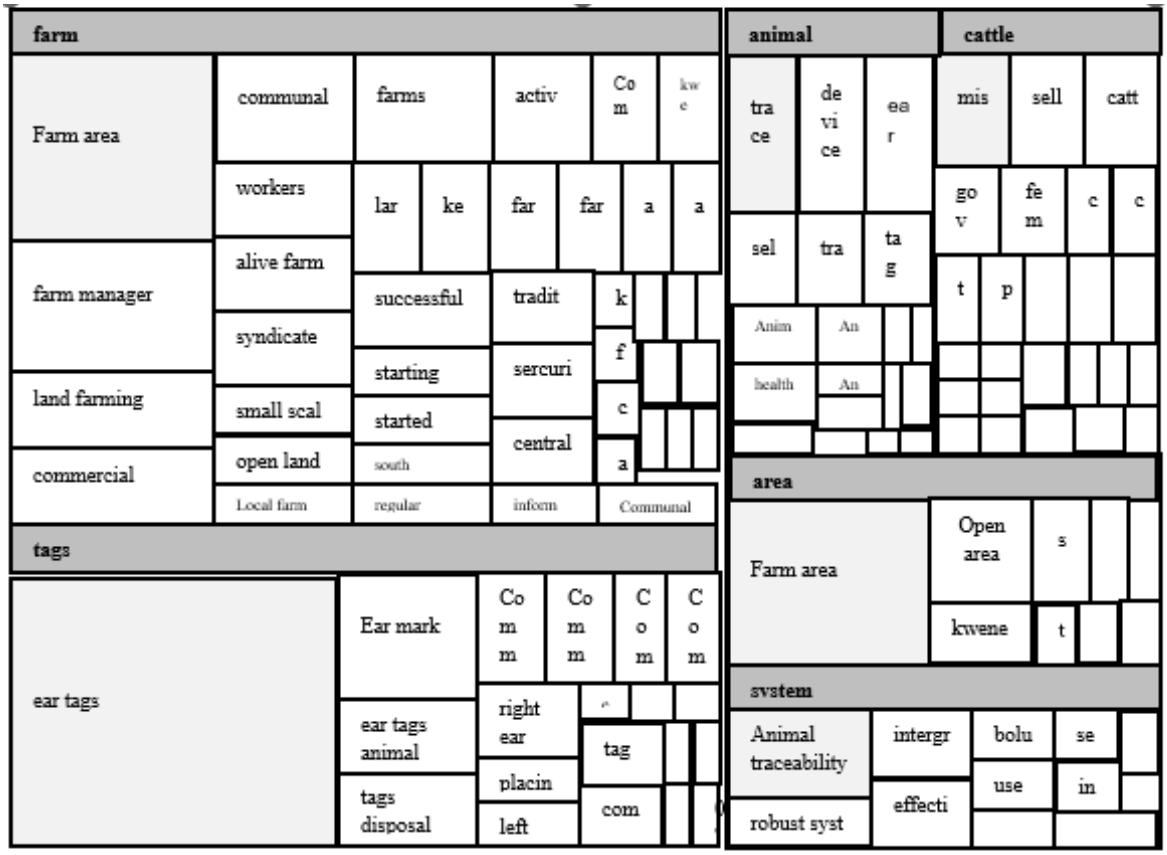


Figure 30 Main tree hierarchy map from animal traceability data.

Figure 30 is the highest level of the category in the main tree of animal traceability data. The sub-category is labeled farm and under farm, the themes that were identified are depicted in Figure 31 also sized according to high level being the big rectangles and small levels being the little rectangles. The main components in farm are animal traceability, farm manager, land farming, and commercial farming. From the data collection activities, this study found that the majority of the farms had farm managers who resided full-time at the farm while the owners worked white-collar jobs. This shows us that farm managers play a big role in animal identification and traceability. Farm owner respondent said farm managers help in keeping the basic day-to-day running of the farm but getting them to use or keep information on the technical running of the farm is a challenge. He further explains finding farm managers who are literate or willing to follow specific farm management guidelines is difficult. Another farm owner gave their farm manager a book to start keeping records of cattle profiles and activities, but they only used it the first week and never went back. The interview transcripts from commercial farmers show that they have more success in managing their farms. As their sole purpose is to run the cattle farm for business, they hire farm managers who have the skills to do the necessary paperwork and physical work at the farm. Commercial farmers can pay their employees better money, which is why they get better employees.

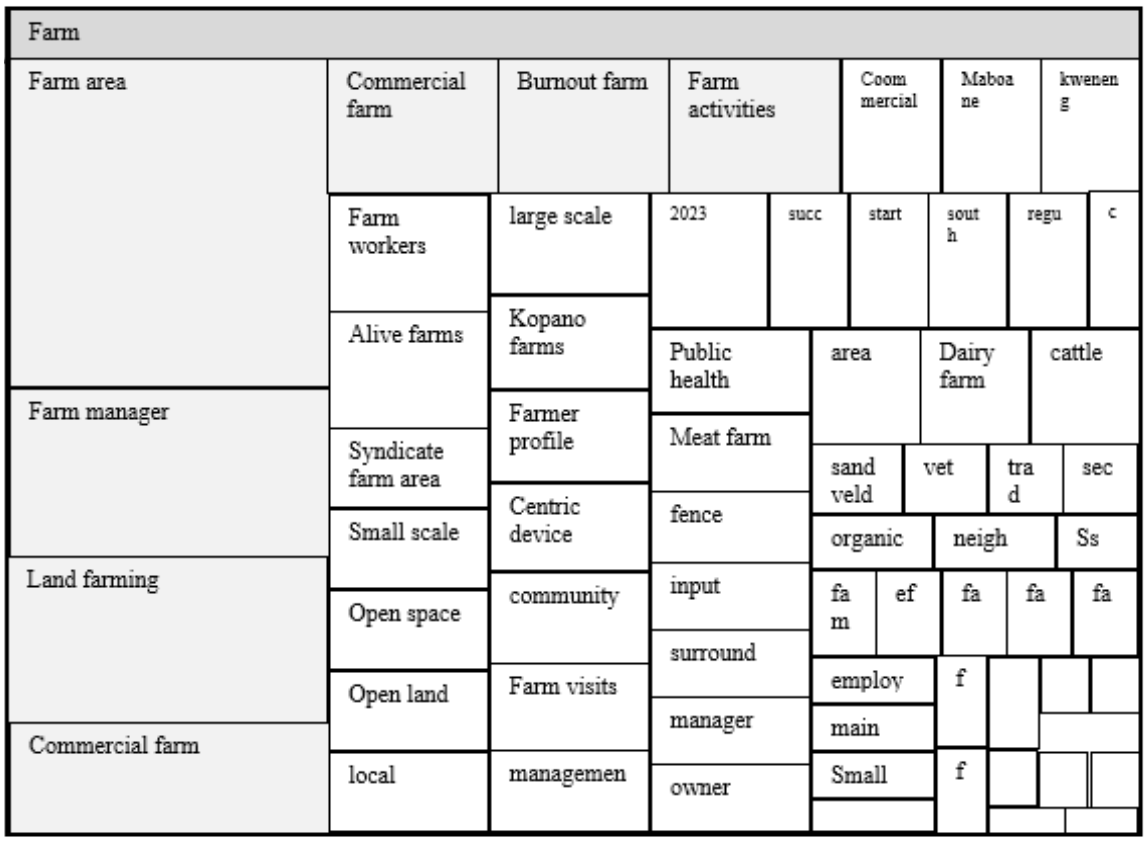


Figure 31 Extract category 1 from the expansion of the "farm" component of the main tree hierarchy map Figure 30.

The second category is the system depicted in Figure 32, it represents all the contents that are related to system mentioned by respondents. The highest categories in system are animal traceability system, robust system, integrated system, and effective system. This study found out that currently there is no integrated system for all the stakeholders of animal traceability. BAITS is the current system, but it does not provide functionalities to all stakeholders except veterinarians who have higher control. The police, BMC, and Matimela officers have access to view, which has not been effective for the parties involved. For example, the matimela representative said they get provisions to re-register the stray cattle they find and give themselves ownership to be able to sell it. Provided that the system has all ear tag numbers and cattle ownership details the system should be able to allow matimela office to search for the ear tag number of the stray cattle and contact the owner to collect it. The ontology-driven framework will provide guidelines for developing animal traceability platforms that will integrate all stakeholders and systems that are robust and effective as countermeasures to the challenges presented by users.

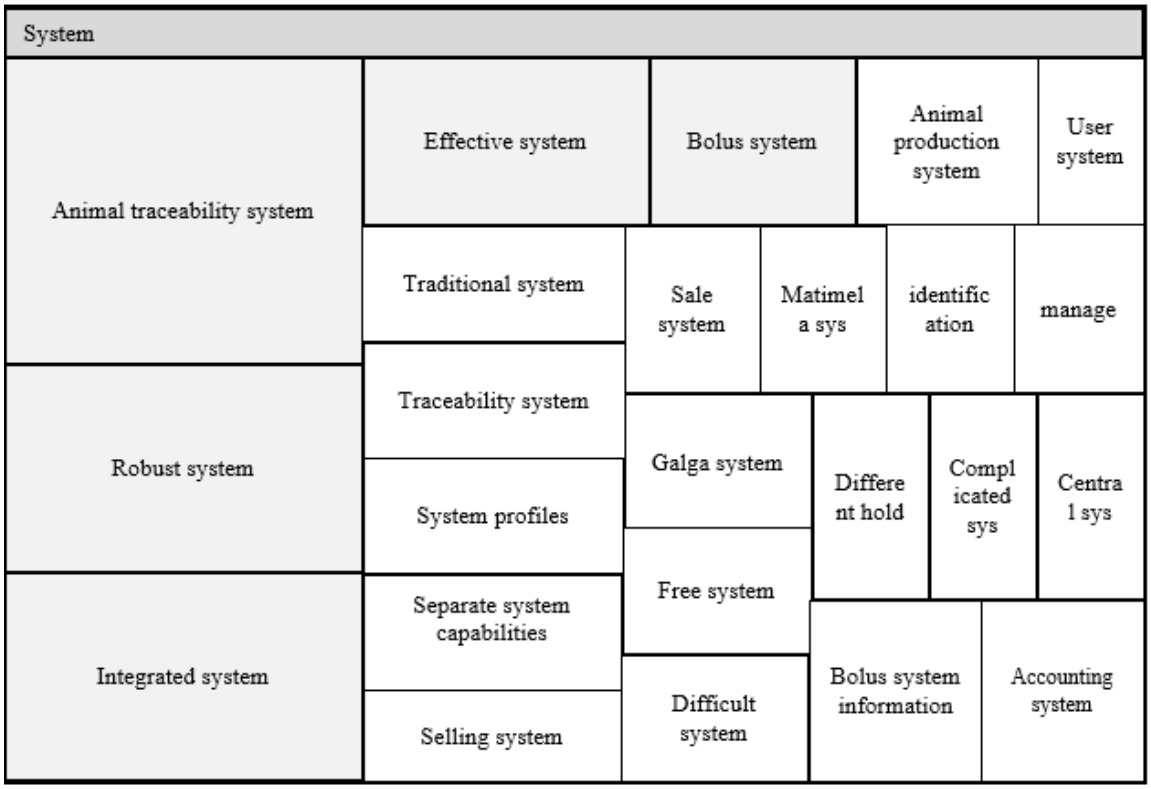


Figure 32 Extract 2 from the expansion of the "system" component of the main tree hierarchy map Figure 30.

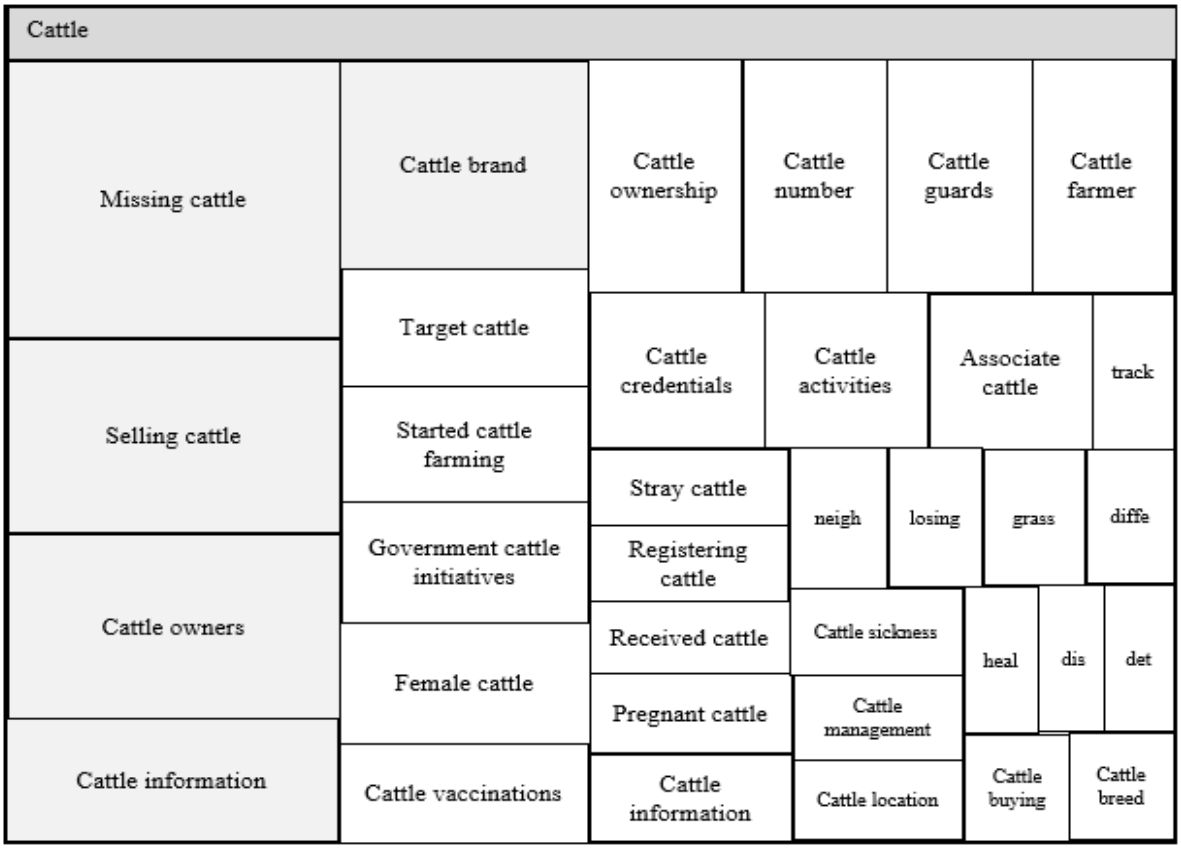


Figure 33 Extract 3 from the expansion of the "cattle" component of the main tree hierarchy map Figure 30.

Cattle are the main component of this study, Figure 33 represents the components generated by NVivo software as sub-categories for the cattle category. The values are important as they will be used and included in the development of the ontology-driven framework. Components like cattle information represent the importance of giving full animal profiles in the system. Missing cattle brings to attention the fact that a method for tracking cattle locations should be incorporated with animal traceability platforms. BMC respondents said that the most profitable cattle are zero-tooth cattle, which provide farmers with high profits. As seen in Figure 33, target animals, zero-tooth cattle, are the main targets for thieves. Selling cattle, cattle owners, and cattle brands point to farmers, further proving that farmers are the main stakeholders of animal traceability.

Hierarchy tree maps, Figures 30, 31 and 32, generated from NVivo, present visuals of the distribution and structure of the emerging theme from stakeholder feedback on animal traceability systems. The analysis results were validated manually based on animal traceability stakeholder feedback. The most prominent nodes include “farms”, “tags”, “animal”, “cattle”, “area”, and “system”, presented as the large rectangles. This indicates coding references at a high frequency in relation to these challenges. The issues that were frequently discussed by stakeholders are effectively highlighted in Figure 30, providing a diagnostic tool for the identification of dominant concerns. However, upon further evaluation, while the hierarchy tree gives clarity on the volume of themes, there is no direct representation of the depth of the themes. Figure 30 shows animal and cattle being separate prominent nodes, but they are similar. Cattle and its sub-nodes reflect the breadth of coding instead of the criticality of the issue. More structurally important challenges like “records”, appearing less frequently coded, are arguably more significant in the success of animal traceability systems.

Regardless of these limitations, the hierarchy tree map is valuable in effectively exploring and scanning coding density to validate that the distribution of the themes is in alignment with the expectation based on the study's research questions. It was useful in guiding the identification of priority animal traceability areas for intervention in animal traceability. Cross comparison was also conducted with the other NVivo outputs (word frequency, cluster analysis, and coding matrix) to refine the challenges and suggested solutions of animal traceability.

E. Project map

The project maps are an NVivo output from the animal traceability transcripts, used in this study to find patterns and relationships from the data to find possible structures that can be extracted to develop the ontology-driven framework. The project maps, Figures 34 and 35, were built by user-directed input with the system-assisted layout optimization. NVivo organizes and manages the spatial layout using an algorithm for the nodes and their relationships. The manual development of project map to enable a conceptual and interpretive representation of the animal traceability landscape. Project map is used to synthesize and visualize the interconnections among thematic cases, data sources, and nodes associated to the feedback from stakeholders on animal traceability. Figure 34 illustrates how the main themes, “Cattle”, “Market”, “BAITS system”, “Farming”, and “Records” are linked to specific stakeholders, like farmer, veterinarian, BMC, and MOA. This visual helps in clarifying the individual feedback was distributed across the various thematic areas.

It is important in visualizing model relationships, stakeholder interactions, and workflows. When building the animal traceability domain models, the project map is useful. Furthermore, assisting in planning the ontology relationships, such as concept linking of “Herdboy”→ “Farm”→“Cattle”.

Two thematic maps are produced by NVivo, Figure 34 and Figure 35, as project map results for the animal traceability data. Project maps were useful in informing the decision-making process of the overall functionality of the ontology-driven framework. It will assist in identifying suitable components and relationships of relevant factors for specific concepts in the ontology conceptual stage. Wilson and Bruni-Bossio [196] describe the application of visual project mapping as a method that is successful for laying out logical and mental processes for a project. The visual project mapping in Figure 38 has indeed

been useful in laying out logical and mental processes for the development of the ontology-driven framework for this study, with the visual of the parent and child relationships.

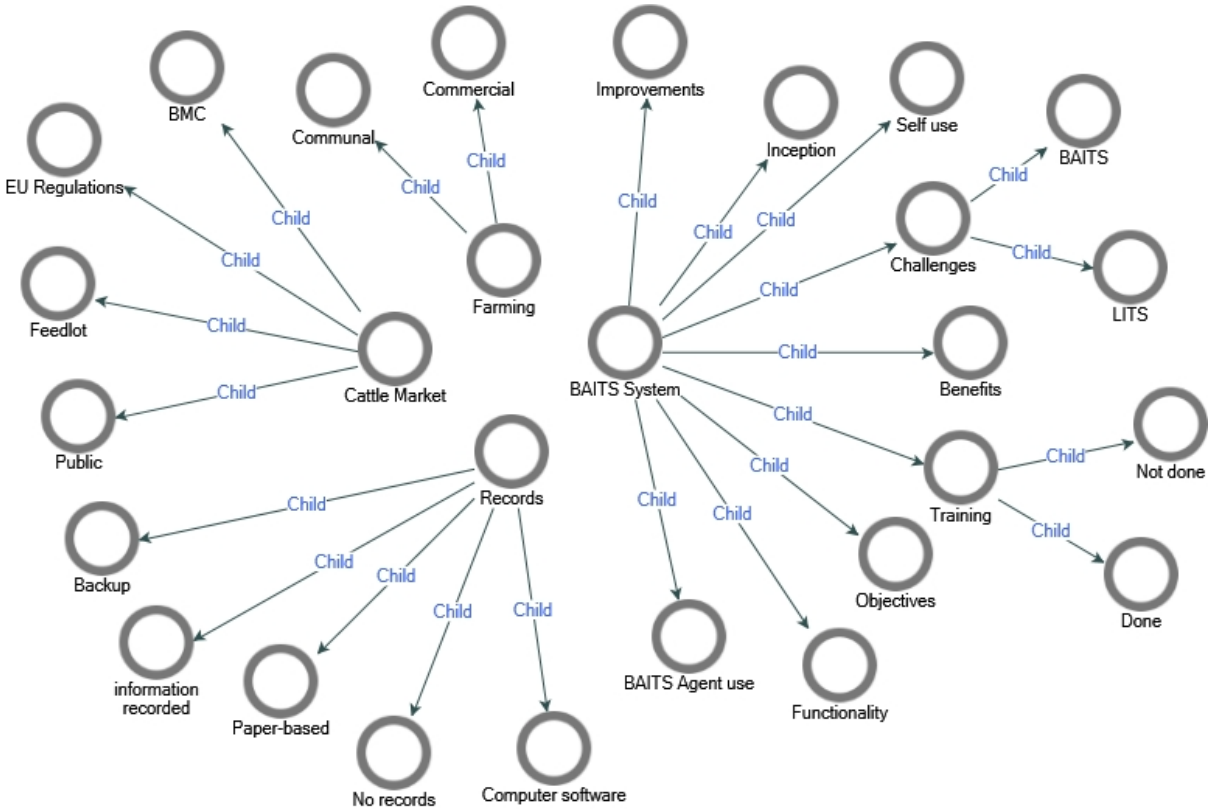


Figure 34 Project map of the four main components generated by NVivo software.

Large data can make it difficult to deduce the relations and themes. Figure 34 has provided a clearer and more detailed pictorial of collected animal traceability data. The project map harnesses the vision of this study to understand the large data clearly. The main concepts identified Farming, Cattle Market, Records, and the BAITS system are depicted with each of their child relationships. The main concepts are also linked to the codes that were generated per district as illustrated in Figure 34. It is seen that unlike in Figure 34 where we see just the links of the parents and children, Figure 35 shows that the main concepts are also linked together through the codes and district interview transcripts. The concepts and their links will feed into the development of the ontology-driven framework.

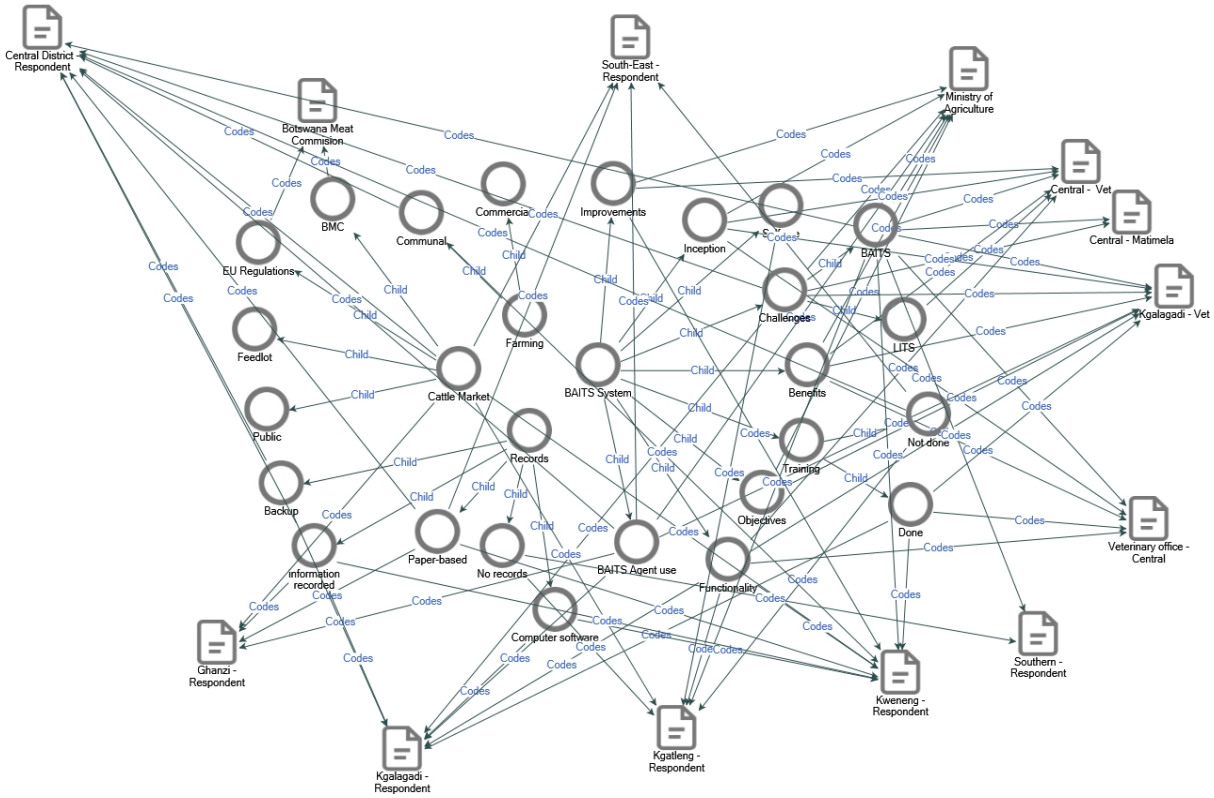


Figure 35 Project map of the four main components, including the coding relationships from each District.

The relationships presented in the project map are researcher-defined and not computed from an algorithm like the other visuals. There is a reflection of the analytic logic manually derived rather than a natural emergent connection. Project map is important in helping analytic transparency, overall structure communication of the findings in a coherent format, and identification of coding logic gaps. Overall, the map was used in internal analysis and external display of the coding decisions and coherent concepts within the study's wide exploration into animal traceability challegens.

4.4 Chapter Summary

The findings and discussion chapter presented the results of the data collection that was deduced in the research data analysis chapter of this study. The analysis chapter synthesized data using NVivo software and Microsoft Excel. After the data was deduced, the tables and figures were presented and this chapter. The demographic findings of age, gender, source of income, and animal traceability outlook were first discussed based on the respondents' feedback. Visuals generated by NVivo, like word frequency, coding matrix, project maps, and hierarchy tree maps, were presented and discussed. The visual presented valuable insights into the co-occurrence of key themes within stakeholder feedback on animal traceability systems. NVivo brings insight to the qualitative analysis for the ATO design. Each of the NVivo features strategically helps in the design of the ontology-driven animal traceability and the animal traceability ontology. The word cloud is the foundation of the actual data used as concepts for ATO, such as cattle, farmers, and records. The project map and cluster matrix derive logical relationships, used in the binary relation of the ontology. The tree map provides structure to the taxonomy. Lastly, coding matrix ensures input from the stakeholders is represented properly. The visuals were discussed based on the thematic patterns and contents illustrated, which were concerning the study topic. Both quantitative and qualitative results were discussed, and concepts were identified that answered the study's research questions. The next Chapter 5 proposes the ontology-driven Framework with derived information, patterns and contents provided in the Findings and Discussion of Chapter 4.

Chapter 4 main results that address the problem statement and align with the objectives start with the response of the current BAITS system according to the stakeholders. 39% believe that the BAITS system is ineffective and not useful to them in the traceability of their cattle and management of their farm. Only 17% of the stakeholders said BAITS has been helpful, but only in the identification of cattle when lost (if the ear tag has not been cut) and the sale of cattle. However, for the sale of cattle, the farmers are mandated to put ear tags on cattle to be sold, or the sale will not be permitted. This research also found out that stakeholders (focusing on farmers, law enforcement, and matimela officers) have been struggling with stray cattle, where the cattle go missing, are found by law enforcement, then taken to matimela kraals, only to be sold off after 15 days if not claimed. The stray animals are sold off even when they have ear tags that contain information about the farmer. The issue is that there is no interoperability in BAITS, so law enforcement and matimela officers have no way they identify the owner of the cattle. These results were found from the quantitative research methods.

The data analysis observations underscore the necessity of triangulating matrix outputs with other NVivo tools, such as word frequency analysis and cluster diagrams, as well as returning to the raw data to contextualize patterns. This critical engagement strengthens the analytical rigor of the study, ensuring that recommendations for improving animal traceability systems are grounded not only in visible patterns but also in the depth and complexity of stakeholder experiences. Data results from the analysis tool NVivo were also key in informing the design of the ontology-driven animal traceability framework and the development of the animal traceability ontology. The framework will be grounded on the actual scope and structure for the stakeholders. Furthermore, ensuring that the ontology is built based on the practices and conceptulisations of the stakeholders. Making sure the ontology presents an interoperable and accurate knowledge representation. For example, if multiple interviews talk about “recording the diseases that affect the cattle” and “keeping a record of all the cattle they own”. The ontology will be guided to add concepts “infectedBy” and “belongsTo”

5 CHAPTER 5 PROPOSED ONTOLOGY-DRIVEN ANIMAL TRACEABILITY FRAMEWORK & ONTOLOGY DEVELOPMENT

5.1 Introduction

The purpose of this chapter is to provide a detailed description of the proposed framework. The chapter also describes the design of an animal traceability ontology. The design of the framework and ontology are based on concepts found as a result of the analysis provided in Chapter 4. Chapter 5 presents the proposed Ontology-Driven Conceptual Framework for Animal Traceability in Section 5.2 and the Animal Traceability Ontology in Section 5.3. The previous Chapter 4 Findings and Discussion illustrated theories and concepts that emerged from the data collected from the animal traceability stakeholders. The concepts from Chapter 4 and information from Chapter 2 (Literature review) are used to guide in identifying the relevant and necessary components for the development of the ontology-driven animal traceability framework. According to Al-Baltah et al. [45] ontology is used to support semantic interoperability between data sources that are heterogeneous. Furthermore, ontologies will ensure that the data meaning that is exchanged and shared is understood by both the communicating parties. The main focus when developing the ontology framework is to guide in the creation of animal traceability systems that are interoperable and knowledge-driven. The framework will provide a foundation for creating traceability systems that is extensible and flexible, with ontologies that will allow for expansion and updating over time adding new technologies and data. The proposed conceptual framework addresses

> RQ3: How can an ontology-driven approach address the challenges of interoperability and integration in animal traceability systems?

The next Section 5.3 shows a step-by-step process of the development of an animal traceability ontology based on the information from Chapter 2 (Literature review), concepts from Chapter 4 (Research findings), and guiding principles of Section 5.2 (Proposed Ontology-driven Animal Traceability framework). It was important not only to develop an ontology-driven framework but also the animal traceability ontology because the ontology will further offer semantic reasoning, semantic clarity, interoperability, and standardization. The animal traceability ontology development is in line with the

> RQ4: Which of the existing ontologies can be reused or created from scratch to develop the animal traceability ontology, given the defined components?

Ontology data stores were thoroughly analyzed and checked for similar ontologies to adopt and incorporate that have similar features required by the animal traceability ontology. None of the existing ontologies found offered similar features that could be incorporated into this research. This research opted to develop the animal traceability ontology from scratch following one of the ontology development methodologies. Corcho et al. [96] state that currently ontology technologies are mature enough to develop ontologies from scratch using the available methodologies, languages, and tools. The study chose the methodology called METHONTOLOGY, Protégé development environment, and OWL language for the development of the animal traceability ontology. Chapter 2 (Literature review) provided justification for choosing METHONTOLOGY, Protégé, and OWL.

5.2 Proposed Ontology-Driven Animal Traceability Framework (ODF-AT)

The main objective of the research is to develop an interoperable ODF-AT. The Framework will primarily guide the development and management of animal traceability platforms. Figure 36 is an illustration of the research objective. It is the ODF-AT that can be used as a guiding principle to develop animal traceability systems, websites, applications, and databases. The framework can also be adopted to develop systems in multiple fields like healthcare, cyber-crime, and education.

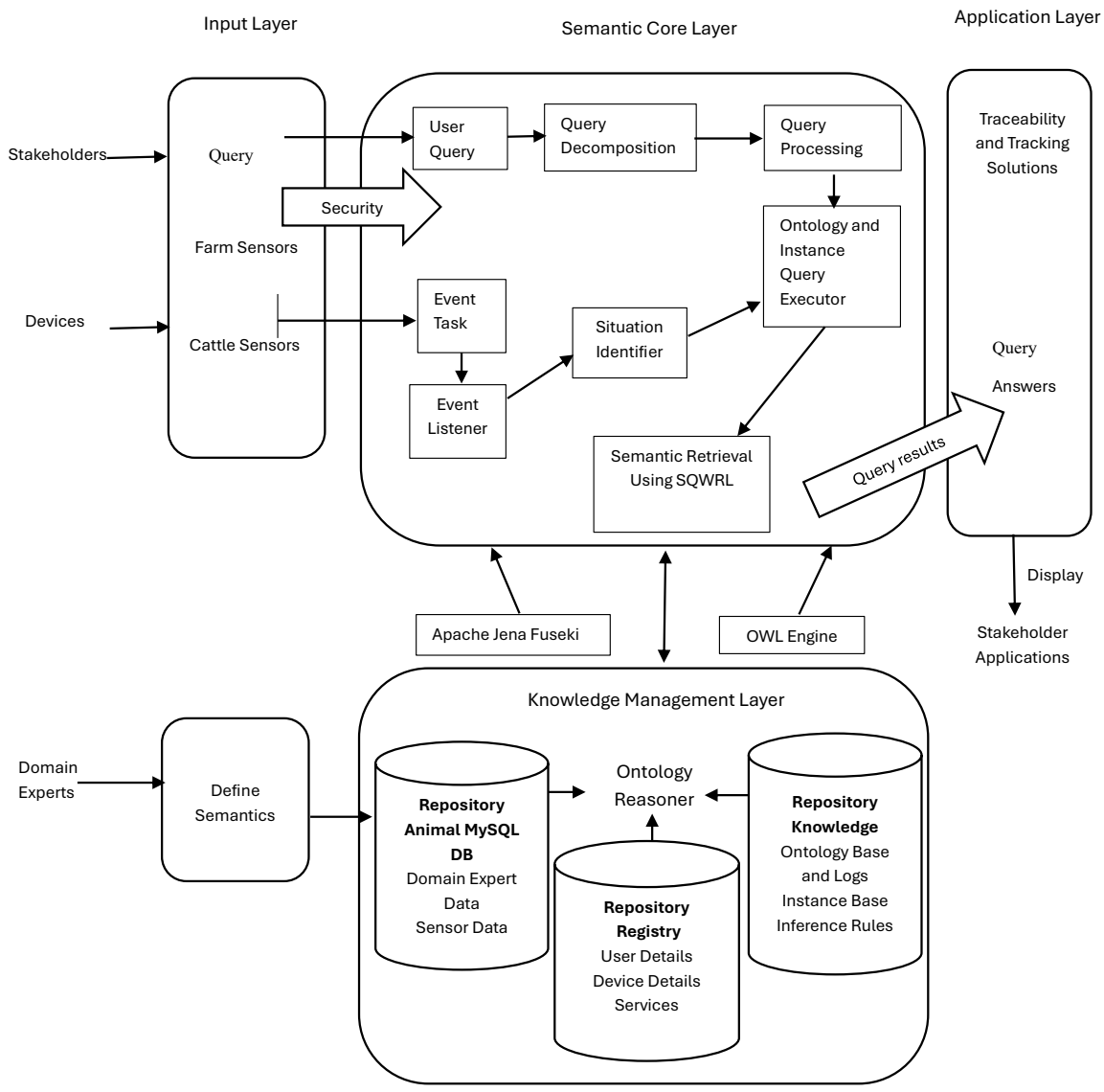


Figure 36 The Ontology-driven Framework for Animal Traceability.

This framework offers a standardized, integrated, semantic approach to managing traceability data, addressing the evolving challenges in the field. Comprising four layers − input, semantic core, knowledge management, and application − the framework facilitates seamless data flow and knowledge management.

5.2.1 The Input Layer of the ODF-AT

The input layer as depicted in Figure 36 is fed by different sources, the first source is the stakeholders who directly input data or make a query to the graphical user interface. The stakeholders for the system

are the farmers, veterinary, MOA, BMC, Matimela officers, and Police officers. The other input source is the sensor devices placed on the cattle and the farm for monitoring and triggering alerts. The information and alerts from the devices feed into the input layer. The last source that provides information to the system is the domain experts, the people who are considered to be experts in the field of animal traceability. Relevant and useful information is derived from the experts sent to the knowledge management system.

Farm location which includes the farm environment plays an important role in the development of a system. The location and environment of a farm influence the type of sensor devices that can be used in the area. The current difficulty that farmers are facing is being able to use any type of system while at the farm because they are in very remote areas. Developers should consider developing systems that can function with satellites to accommodate areas that do not have internet connectivity. The type of farm – commercial or communal - also influences the sensor devices used and the priority of components in the systems. Commercial farms that focus on business as their core objective of rearing cattle have different needs than a communal farm that rears cattle to feed and sustain the family. The type of sensor device that is implemented connected in the farms also depends on the needs of the farmer and their location. A communal farmer will need a tracking sensor device more than a commercial farmer because communal cattle go out of the farm area to graze by themselves which makes them more prone to theft. The commercial cattle graze in the farm-enclosed area together with feed provided by the farmer so they are better monitored than the communal cattle. The commercial farmer can use sensors to keep count, monitor cattle health, heat period, behavior monitoring, and track the movement of cattle for disease control. These sensors are also important to communal farmers with the addition of tracking devices to prevent theft. The current animal traceability system in Botswana does not incorporate any type of sensor device to obtain important information about the cattle.

5.2.2 The Semantic Core Layer of the ODF-AT

Within the semantic core layer, the process begins with a security authentication check for users logging into the system as seen in Figure 36. The security check is used to make sure the stakeholder's private information is secure. Authentication is also used to determine who the user is and what privileges they have in the system. Each user has different viewing and editing privileges in the system. If the user has authorization the query proceeds to the next layer where the query will be processed, if not an error message is returned to the user. In the semantic core layer, the queries from the user and alerts from devices are managed. The query from the user is first registered in the system. Query processing starts the decomposition of the query by transforming it from a high-level query to a relational algebra query. The query is then checked if it is semantically and syntactically correct. The decomposition involves semantic analysis, simplification, and query restructuring.

The results from the query decomposition are processed in the query processor which then passes the query to the ontology query executor to check the requested query. The ontology query executor carries out four steps, it searches through the cache to check if there is an existing plan of execution for the query. If no execution plan has been parsed, then the query parser parses it through the semantic retrieval tree using SQWRL to the knowledge management layer. SQWRL is used as the query language for defining the set of Structural Query Language (SQL) like query operations from the ontology query executor. SQWRL constructs specifications for retrieval of information stored in the ontology base. Semantic retrieval using SQWRL is effective and useful in combining ontological properties and concepts with instances, SQL statements, and production rules. With the query statement, the animal traceability system interface between the user's query intention and knowledge base is set up.

5.2.3 The Knowledge Management Layer of the ODF-AT

The knowledge management layer includes the MySQL database repository, the ontology reasoner, and the ontology base with associated logs shown in Figure 36. The knowledge management layer is created

to describe the composite situation triggered and their animal traceability-related concepts. The first repository is a database for sensor data and domain expert data. All the information that is read by the sensor devices is stored in the repository along with the information that is inputted by experts in the field of animal traceability. The data is stored as raw data and MySQL performs the first manipulation of transforming the data into more organized structured information. The second repository is used for the ontology base and logs. Ontologies are stored independently in the ontology base represented in OWL files. The ontologies are manipulated by the administrators and knowledge is deducted from them. Instances of the ontology relations and concepts are also stored in the instance base. The instances are used to represent and store different concepts like OWL files, electronic documents, and HTML pages. They are important for extracting new ontology constructs, and decision-making for adding and deleting ontology constructs.

The repository also contains the ontology logs which provide a list of actions and executed operations that occurred in the ontologies. The logs are used for the recovery of ontologies in case of data/system corruption or inconsistencies. The ontology recovery occurs by the system reviews the ontology logs queries that are uncommitted and then rolls back changes to the queries made. The third repository is a registry of user details, device details, and services. The user details registry delivers a semantic description of the preferences and profiles of the system users. Relevant information pertaining to the status and profile of the sensor devices is provided in the device details registry. The last is the service registry which supplies functional services that occur in accordance with the location of the service deployment based on the link between the event and suitable action services.

The ontology reasoner also known as the semantic rules engine is an automated reasoner used to infer logical consequences from the set of axioms written in OWL. The reasoner provides mechanisms for performing the requested operations using inference rules from the ontology. For the animal traceability framework, the ontology reasoner carries out services of query answering. It is used to reason about the knowledge representation in the ontology, the input is from the animal traceability repositories. First, it checks if the structure exists that satisfies the axioms by associating the reasoning with the animal traceability ontology and the query, if it fails a warning message is sent. If it passes it continues to the next phase to retrieve the user query results based on the satisfaction of the inference rules. Basically, the options at the user interface convert the user's request to SQWRL query statements and at the ontology reasoner, the reasoner finds and extracts the information required from the animal database, registry repository, and ontology base then returns the query results to the user. The reasoner is important to the developer as it can ensure that the animal traceability model is always logically consistent in the classification.

Apache Jena Fuseki is part of the animal traceability framework, it is a Simple Protocol for RDF Query Language (SPARQL) server that is run as a standalone server or an operating systems service. According to Chokshi and Panchal, [197] the creation of the SPARQL standard allows users and developers to concentrate on the things they would like to know and not on how the database has been organized. Apache Jena Fuseki is utilized for information applications that are connected and building the semantic web. SPARQL supports querying and updating. The OWL Engine facilitates communication between the querying user and the answering server. A protocol for the question-answering communication between agents using knowledge representation is defined by the OWL Engine. It works with the ontology reasoner to manage the animal traceability queries.

5.2.4 The Application Layer of the ODF-AT

The results of the query processed within the knowledge management layer are then forwarded to the application layer (Figure 36). The application layer computes the query and displays the answers to the stakeholders through the Graphical User Interface (GUI). It is like the outcome phase. The systems and applications developed are seen in this layer. The systems are handed over to the users and the user can

experience the system through the graphical user interface created. As seen in Figure 36 the outcome phase is connected with the stakeholders because they are the end users of the developed system.

Another important aspect to be considered is the language the system will communicate in. In Botswana the current systems have language barriers with their users, the users understand the native language Setswana, but the system communicates fully in English. The developer should provide communication both in Setswana and English to cater to all users. When the system is completed after being trialed and tested, the developer has to deploy the system to the users with detailed descriptions and manuals. Training for users should be conducted as one of the unsuccessful objectives of the current system in Botswana. The training courses are important to inform the users and teach them ensuring that the system is being effectively used to assist its users. The developer can also identify problems in the system during the training and fix them in the early stages. Maintenance of the system is important as a component that will continue throughout the use of the system. New updates and fixes will be conducted in the maintenance step.

Developers can adopt the already developed animal traceability ontology in section 5.3 and use the components to develop their system/application. Ontologies are used to provide semantic interoperability of agreed concepts for mutual communication of systems in animal traceability, and integration of information from the different stakeholders that all come together to form an ontology for animal traceability. The ontology provides scalability meaning that any new information that a developer wants to add can be easily added without making any substantial change. Reusability is very important in the era where technology is growing each day, the vocabularies and semantics in this ontology are reusable not only for animal traceability but also for other fields. This saves developer time and resources and allows for further growth of new ideas to be added in the future.

5.3 Ontology Development Approach

The main outcome of this research is an artfact in the form of an ontology. In this regard two main methodologies would be appropriate to follow: Deseign Science Methodology and the METHONTOLOGY. Design science methodology (DSM) is a broad, iterative approach for creating innovative products or systems to solve real-world problems, generating both the artifact and the knowledge about its use. Other hand, the METHONTOLOGY methodology is a specific, structured methodology used to build ontologies, which are explicit representations of knowledge, by providing detailed, quality-oriented phases and activities for their systematic development. The METHONTOLOGY was chosen over the DSM because the DSM provides a general research paradigm for the creation of artifacts, whereas the METHONTOLOGY is a more detailed and ontology-focused methodology for constructing the knowledge bases such as ontologies.

5.4 Animal Traceability Ontology (ATO) Development

As stated in Section 5.3, the animal traceability ontology was developed using the METHONTOLOGY methodology. Al-Baltah et al. [45] suggest using METHONOTLOGY methodology to implement ontologies from different domains, believing it provides effective guidelines and better ontology products than the other methodologies. The ontology development tool used is Protégé 5.1.6 for the creation and visualization of web ontologies using the ontology language OWL.

This research proposes and develops the animal traceability ontology to serve as a guiding principle for developing animal traceability systems. For countries like Botswana, Namibia, Malawi, and Kenya, which have been experiencing challenges with animal traceability systems (presented in Chapter 2, Literature review). Developers will need detailed steps and documentation to successfully integrate the ontology. METHONTOLOGY is chosen as the ontology development methodology because it has a development process that is systematic and comprehensive, with step-by-step phases – knowledge acquisition, specification, conceptualization, formalization, implementation, evaluation, and

documentation. METHONTOLOGY is suitable for developing the animal traceability ontology for developers and researchers who require being part of the development process before jumping to implementation, support for tools that make software engineering easy through clear documentation templates, ontology lifecycle management, and an evaluation checklist. The evaluation process ensures the ontology is semantically coherent by validating competency questions, checking consistency, and using metrics for quality assurance. These make it possible to easily deploy the animal traceability ontology in real-world systems.

5.4.1 Animal Traceability Core Features

5.4.1.1 Identification

Identification is a form of uniquely identifying one cow from another cow. Identification is carried out and is important to all the animal traceability stakeholders. The most common method of identification for cattle is ear tagging. The ear tag contains a number that identifies the cattle owner and a unique number for each cow. The ear tag that contains a radio frequency identification device (RFID) technology is called an electronic ear tag. The electronic ear tag stores information about the cow, the sex, gender, color, identification number, location, and owner. Botswana sells electronic ear tags to farmers, but they are underutilized because the RFID inside the tag is not being used. The majority of the farmers do not know it exists, they only use the visible numbers of the tag. This is because farmers do not have and cannot afford ear tag reader devices. Farmers have also created their own ear tags that have the animal owner's details like location, name, and phone number. A mark is also used by the farmers on one side of the ear by making a unique cut that identifies the owner of the cattle.

For any form of tracking or traceability to happen the first step is identification that can uniquely identify the cattle. The presented ontology-driven framework proposed the use of electronic devices to effectively identify cattle. Technologies like wearable devices, tracking collars, tracking tags, sensor devices, and the Internet of Things can be adopted by farmers to effectively collect the data that will be used for the ontologies. The current manual collection of data is flawed because there is a lot of human error that affects the results of animal traceability.

5.4.1.2 Tracking

Tracking of cattle is being able to record and store periodically time location of cattle movement using GPS tracking devices. Tracking is beneficial to farmers in providing time location of cattle when they are gone out for grazing. To study the movement patterns of the animals around the area, quickly identifying the location of cattle if they are lost or stolen, determining the location of a cattle if it died in the forest, and identifying the last movement of cattle before they got sick to determine disease infected areas. For farmers who work full-time and only go to their farms on weekends tracking will keep them involved remotely when they are not at the farm, they can better communicate and solve daily challenges with their farm managers. With trackers integrated with ontology systems police, BMC, and matimela can easily determine the location and movement history of the cattle without relying on word of mouth from the farmer. It will verify if indeed the farmer is the owner of the cattle. Tracking devices play a vital role in the ontology-driven framework in terms of the information that is fed to the ontology for knowledge management and interoperability. Respondents from MOS relayed that tracking is currently not a priority to them like traceability and they do not practice it. Furthermore, tracking benefits the farmers more than MOA because their main aim is animal disease. Farmers in Botswana are also not benefiting from the tracking devices because some of the farmers do not know about them and where they can find them, some of the devices are expensive for an average farmer.

5.4.1.3 Traceability

The only technological activity that is practiced in Botswana is animal traceability called the BAITS system. BAITS is provided by the government of Botswana. Traceability is used to track the lifestyle

of the cattle from farm to fork. That is the location where the cattle were born and other locations the cattle moved around before it was slaughtered. The food intake of the cattle, the disease the cattle caught and how it was treated, the vaccinations the cattle receive, and other supplements given to the cattle. The information is important to ensure the safety of the cattle meat that is being eaten by people. Currently, it is a struggle for farmers to record this information and provide it to the relevant stakeholders because the system is place is not effective nor user-friendly for farmers to record their data. The ontology developed in this study provides more crucial data that should be recorded and the user-friendly develop user friendly systems. The study presented the challenges with current system and proposed countermeasures that will assist developers and researchers when adopting the ontologies to develop animal traceability platforms.

5.4.1.4 Farm Management

The components of the animal traceability framework include important farm management aspects that all farmers need to adopt in order to run a profitable and manageable farm for the best animal traceability results. Farm management includes record management of cattle, employees, finances, inventory, and farm visitors. Productivity maintenance of the farmland and cattle.

5.4.1.5 Interoperability

The current system BAITS does not provide integration and interoperability for the stakeholders. Ontologies are proposed for their ability to effectively counter these challenges in animal traceability. Not all the stakeholders use the current animal traceability system they have developed their own systems even though they are all dealing with the thing "cattle". MOA provides veterinary services and farmers with BAITS. Police and BMC only have viewing privileges, only BMC checks while the Police contact veterinarians when they want cattle information. BMC uses a system called integration for its animal traceability activities. Matimela employed a private company this year 2023 to develop a traceability system for their cattle activities. Most farmers use Microsoft Excel, which does nothing productive for the traceability of their animals. The proposed ontology-driven framework provides integration of an interoperable animal traceability system that all stakeholders can coincide and work together under one system to yield the best result in cattle rearing.

5.4.1.6 Ontology Reuse

Ontologies can be reused, thereby save time, money, energy, and resources of the developer. The vocabulary that has been developed for ontology will be published online for developers and researchers to adopt. The study encourages system developers and ontologists to reuse the ontology in their studies in animal traceability and other fields of study. The study can be carried on with other innovative techniques to improve animal traceability further.

5.4.2 Ontology Development - METHONTOLOGY

This is the ontology development section for animal traceability using the METHONTOLOGY methodology. The METHONTOLOGY stages used for ontology development were adopted [96], [101], [198], [199], [200]. Figure 37 depicts the stages that were carried out during the ontology development process. The methodology has three main categories: management activities, development activities, and support activities. All the activities coincide together throughout the development process. Each stage of the activities is illustrated and described with relevant tables and figures that represent animal traceability data and concepts.

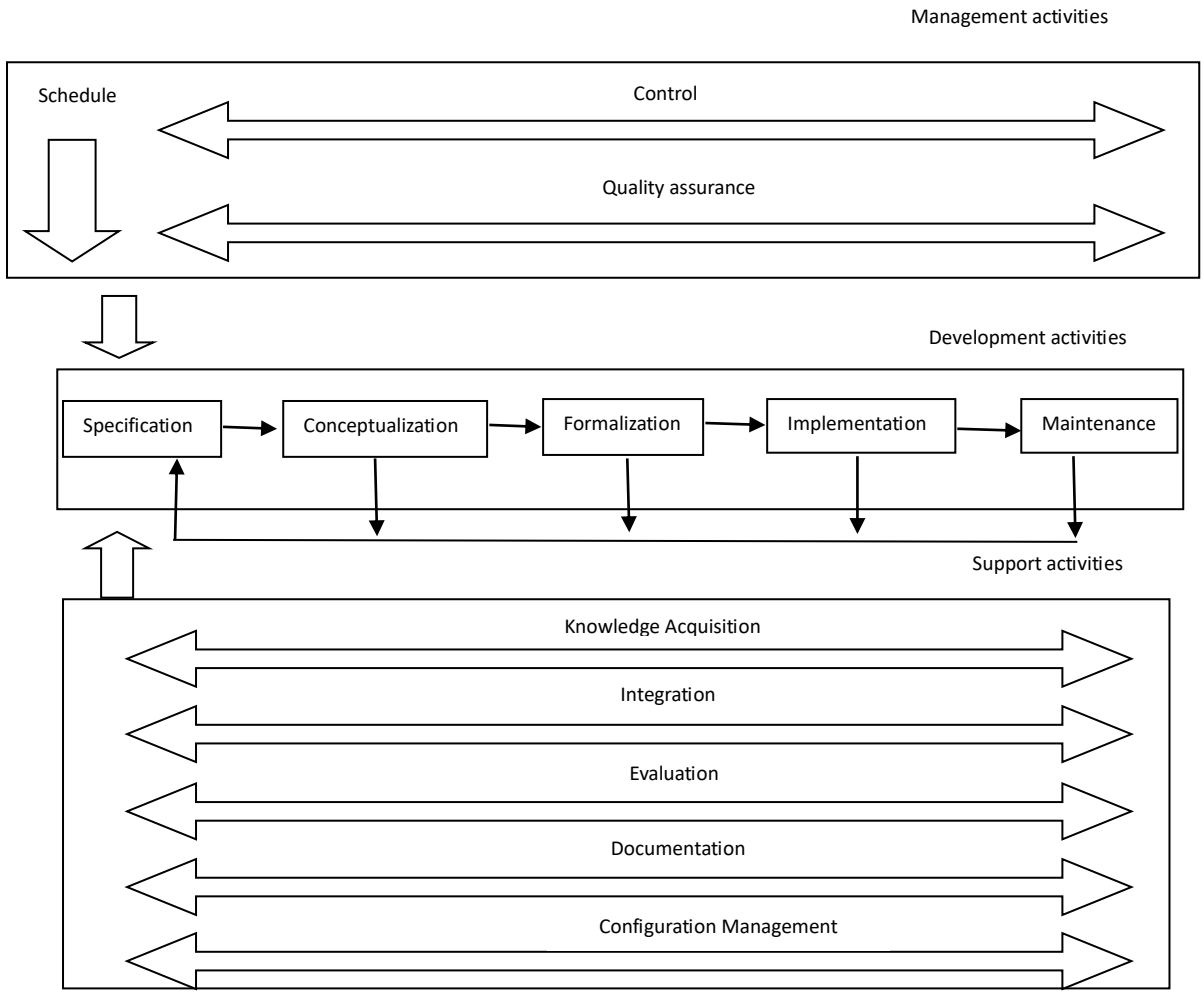


Figure 37 Ontology Development Activities as proposed by METHONTOLOGY [75].

5.4.2.1 Ontology Management Activities

The ontology management activities include planning, control, and quality assurance. The first activity planning was carried out prior to the development of the ontology-driven framework for animal traceability. It was used to identify the tasks that were necessary to develop the ontology-driven framework. The tasks were arranged with the time and resources needed to fulfill them. The tasks can be seen in the specification stage of ontology development. The control activity was used to ensure that all the tasks that were planned were completed as intended. The last activity in ontology management is quality assurance which checks the quality of every methodology output and assures that it is satisfactory. Throughout the ontology life cycle the control activity and the quality assurance activity will be performed.

5.4.2.2 Ontology Development

The development activities begin after planning has been completed starting with the specification. Before the specification knowledge acquisition is conducted first to gather the knowledge from experts that is relevant to building the ontology. Most of the management activities and the support activities ran simultaneously with the development activities. Acquisition and evaluation are conducted in the conceptualization stage of development. The conceptualization is supplied by information from the knowledge acquisition activity. Evaluation is carried out mostly in the conceptualization stage to prevent errors when moving to the implementation stage. The integration process is not performed in the implementation stage only but throughout the development of the ontology.

5.4.2.3 Knowledge Acquisition

Contextualization and empirical for the knowledge acquisition phase is provided by the output from the data analysis and findings phase of this research (Chapter 3 Research Design and Data Analysis and Chapter 4 Findings and Discussion). In this activity, the aim was to acquire knowledge from stakeholder experts in animal traceability used in Botswana. Data obtained from interviews, and literature analysis provided insights into existing practices, stakeholder needs, regulatory guidelines, and best practices in animal traceability as illustrated in Figure 38. The use of NVIVO to review and interpret data from the stakeholder interviews and literature uncovered key patterns, entities, and gaps in animal traceability. This information analysed shapes the empirical foundation for the knowledge acquisition phase to define and formalise the relevant domain (animal traceability) concepts, constraints and relationships. Furthermore, ensuring that the model developed is contextually accurate and semantically rich. Providing an accurate reflection of the real-world practices of disease control, movement tracking, regulatory compliance, and animal identification.

5.4.2.4 Specification

According to Sawsaa and Lu [201] the specification stage starts by analyzing the purpose of the ontology then the goals, strategy, boundaries, and scope are identified. He further states that specification is important for avoiding problems of inconsistencies among classes of ontology. As depicted in Figure 38 the specification stage addressed the three main points. The purpose of the ATO was to address the challenges associated with non-interoperable animal traceability systems in Botswana. The goals of the ATO were identified as follows: (1) to enhance interoperability between different stakeholders involved in animal traceability, (2) to standardize vocabularies and data management practices among stakeholders, (3) to enable seamless data integration and exchange, (4) to derive value from traceability data for knowledge-driven economic opportunities, and (5) to support effective knowledge management for accurate and rich data about animal traceability.

The boundaries of the ATO were defined by its focus on representing knowledge directly related to animal traceability, excluding unrelated areas such as non-animal agricultural products or broader environmental concerns. The ATO's scope encompassed enhancing knowledge representation and extraction, promoting reusability, and supporting cross-platform exchanges among various stakeholders involved in animal traceability. These stakeholders included entities involved in the lifecycle of cattle, such as farmers, veterinary services, Matimela officers, BMC, the Ministry of Agricultural Development and Food Security, law enforcement agencies, and the public.

The ontology is intended to be a high-level ontology-driven framework that is capable of integration into developer applications and software, that has the potential to be applicable and adopted in a wide range of animal traceability platforms. From the acquired domain knowledge, requirements specification was added as a set of competency questions that the knowledge base created should answer from the developed ontology. The competency questions assist in understanding the research scope and in the validation stage of the ontology development using an animal traceability system/application developed. Competency questions were designed based on the codes and themes from the results of the data analysed by NVivo tool. The competency questions are created for developers who will adopt the Ontology to test the effectiveness of their developed animal traceability systems/applications, checking if it answers domain-specific questions. The selected competency questions from the many are as follows:

1. What is the holding name and district location of the farmers? How many cattle reside in each of the holdings per farm?
2. Who are the employees in the farm?
3. How many cattle have been tagged in each holding? Brand, traceability tag, and personal tag?
4. What is the birth rate and death rate of cattle monthly in a holding?
5. What is the loss rate of cattle per holding per district monthly?

6. What tracking mechanism is used by farmers? Does the cattle tracking device provide the cattle coordinates?
7. What vaccinations have been administered to cattle? Individually, date, ID, vaccination serial number?
8. Which animals are sick? Diagnosis, attending person, solution?
9. What are the audit results, to give farmers approval to sell to EU profitable markets?
10. What is the role of each stakeholder in the system?

We researched and checked to find similar ontologies in this domain that could be reused for this study, unfortunately, no ontology that exists was found that was suitable for this study. Technologies that were used in the development of the animal traceability ontology include Protégé 5.6.1 ontology development environment, OWL and RDF query language, Pellet Reasoner, SPARQL, and Apache Jena Fuseki for hosting the ontology and running SPARQL queries.

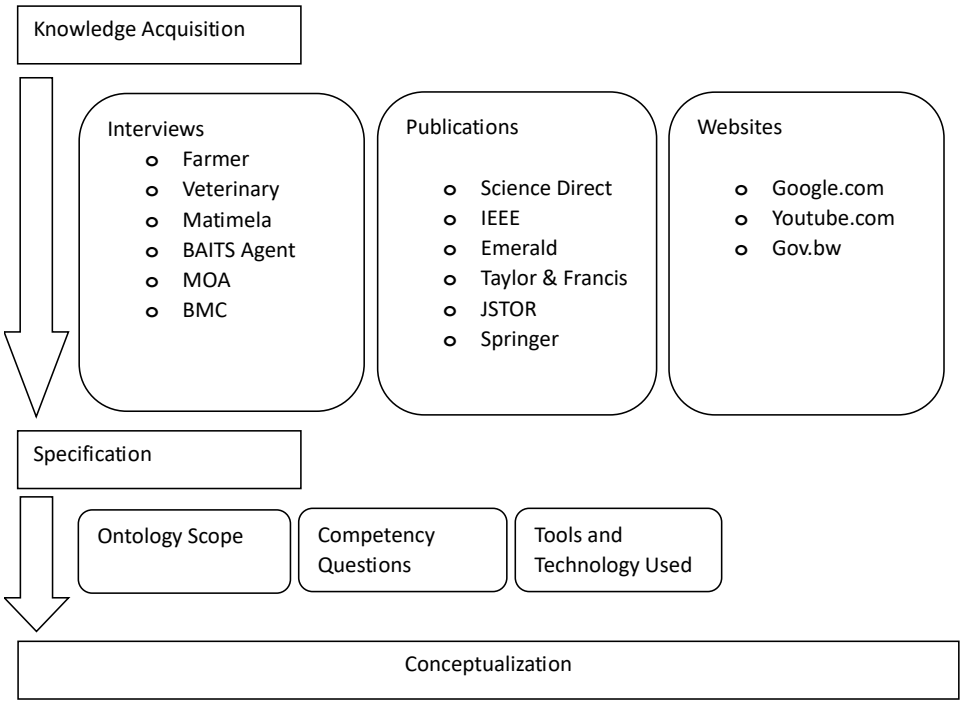


Figure 38 Knowledge acquisition and Specification activities of the animal traceability ontology.

5.4.2.5 Conceptualization

In the conceptualization phase, the focus was to take the acquired knowledge to arrange and structure it using tables and graphs. The tables and graphs include the glossary of terms table, taxonomy, binary relations diagram, concept dictionary table, binary relations table, instance attributes table, class attributes, formal axioms table, and instances table as depicted in Figure 39. Zakaria et al. [202] describe the conceptualization stage as the change of informal data to semi-formal specifications with the use of representations based on graph and tabular notations guided by METHONTOLOGY. The study used knowledge acquired to create meaningful models using concepts, instances, attributes, their relations, and axioms. The models make the ontology development easier to understand. Figure 39 depicts the steps that were taken to perform the conceptualization of the animal traceability ontology. The First step is building the glossary of terms as presented in Table 17.

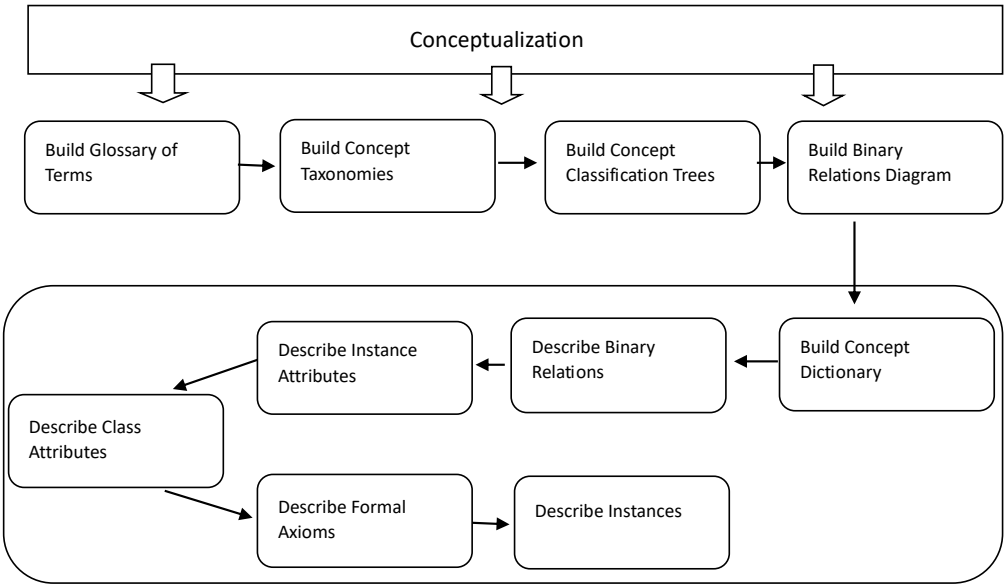


Figure 39 The conceptualization steps for developing animal traceability ontology based on METHONTOLOGY.

A. Building the Glossary

The output from NVivo software in the data analysis phase plays a critical part in informing the construction of the glossary of terms. Thematic analysis and systematic coding was facilitated by NVivo of the stakeholder interviews and animal traceability documents. Recurring concepts, relationships, and terms specific to the domain were identified and categorised. The themes and codes serve as the primary input for building the glossary, ensuring the terminology in the ontology reflects the specific domain language that is grounded in empirical evidence. Table 17 lists the terms along with their description and synonyms. The next tables list the class data properties from the glossary of terms, with a description for each property, which are the sub-components. The glossary of terms is important to conduct at the beginning of the development to define and describe the terminologies that will be used in the activities involved in building the ontology. The full glossary of terms for the ATO is available at Figshare: (https://figshare.com/articles/dataset/AnimalTraceability_rdfAnimal_Traceability_Ontology/24461224?file=42977926) and Github: (https://github.com/tmokgetse/Animal-Traceability-Ontology)

Name	Description	Synonym
Cattle	The animal in the ontology	Cow
Farm employee	A person employed at the farm, a manager or herd boy	Herd man/boy
Farmer	The person who owns the cattle and farm	Rancher
Organization	The various professional offices that are involved in animal traceability	company
Veterinarian	Government employee who provides various services to farmers for their cattle activities	Vet, doctor
Police	Government employee who takes charge and investigates cattle theft	Law enforcement officer
Matimela	Stray cattle	Missing cattle
Matimela officer	Government employee who deals with stray cattle	-
Ministry of Agriculture	Government office that provides ear tags and animal traceability services	-
Botswana Meat Commission	The organization that buys cattle from farmers that meet the regulations	-
Market	The businesses that farmers go to sell their cattle, e.g feedlot	Vendor

Identification	A form of uniquely identifying one cow from another	Recognition
Tracking	A technique for gathering the location of cattle.	Trail
Traceability	A technique of tracing animals from farm to fork	Trace
Farm	A place where cattle are reared	Holding
Records	Documents kept about cattle, and practices on the farm	Documents
Disease	A sickness that affects cattle	sickness
Vaccination	A drug administered to cattle to prevent diseases	Immunization
Farm management	Activities that are carried out at the farm day-to-day	Holding management

Table 17 Glossary of Terms for Animal Traceability Ontology.

Cattle Class and Farmer Class Data Properties of Glossary of Terms

Table 18 depicts the glossary of terms for the class Cattle on the left side of the table and class Farmer on the right side of the table. Each data property for both classes is provided with a description of the term

Cattle Data Property	Description	Farmer Data Property	Description
Ear tag number	unique number on the ear tag of a cow for identification	National identity number	A unique government appointed number that identifies a farmer
Date of birth	The day the cow was born	First name	The name that identifies a farmer
Ear tag color	A color on the tag that identifies cattle by groups of breed, gender, location, or owner	Last name	The last time that identifies a farmer
Brand ID	The unique number branded on the cow belonging to the farmer	Age	Age of the farmer
Color	The skin color of a cow	Gender	Sex of the farmer. Male or Female
Sex	The gender of a cow, male or female	Address	The current physical address of the farmer
Breed	Type of breed of the cow	Contact	The current mobile number used by a farmer
Arrival and Depature dates	The date a cow arrives and leaves the farm		
Calves Count	The number of calves given birth to by a cow in a year	Keeper ID number	The unique number that identifies the farmer
Market	The place/person where the cattle is sold	Nationality	Nationality that the farmer belongs to
Price	The price that the cow was sold/bought	Date of birth	The day the farmer was born

Table 18 Glossary of terms for the class Cattle and class Farmer.

Farm Employee Class and Tracking Class data Properties of Glossary of Terms

Employee Data Property	Description	Tracking Data Property	Description
National identity number	A unique government appointed number that identifies an employee	Cattle ID	Identification number of cattle used for tracking the cow
First and name	The name that identifies an employee	Farmer ID	The identification of the owner of a particular cow being tracked
Age	Age of the employee	Location	The exact location of the cow at the time it is tracked
Employment history	The previous places of work of an employee	Device name	The name of devices used to track cattle
Address	The current physical address of employee	Serial number	The serial number of the tracking device
Date of birth	The day the employee was born	Last updated	The last updated information about the cow
Contact	The current mobile number used by employee	Temperature	The recorded temperature of the cow
employee ID number	The unique number that identifies employee from others in the farm	Signal strength	The signal strength of the tracking device at the location of the cattle

Nationality	Nationality that the employee belongs to	details	The details of the cows were recorded at their exact location
Date employed	The date the employee started working	History	The history record of each cow tracked
Last day	The date the employee stopped working	Track ID	The identity number that identifies the tracking history of the cow

Table 19 Glossary of terms for class Farm Employee and class Tracking.

Table 19 depicts the glossary of terms for the class Employee on the left side of the table and class Tracking on the right side of the table. Each data property for both classes is provided with a description of the term.

Traceability Class and Disease Class data Properties of Glossary of Terms

Table 20 depicts the glossary of terms for the class Traceability on the left side of the table and the class Disease on the right side of the table. Each data property for both classes is provided with a description of the term.

Traceability Data Property	Description	Disease Data Property	Description
Cattle ID	The unique number that identifies a cow	Type	The name of the disease affecting cattle
Animal history	The transfer of ownership history of each cow	Cause	Event, condition, or characteristic that plays an essential role in producing an occurrence of the disease
Current owner details	The owner of the cattle name and location	Treatment	Description of how the cow was treated and name of drugs used
Current owner location	The current farm location of the cattle owner	Dosage	The amount/quantity of drugs administered to the sick animal
New location of owner	The transfer location of cattle from one location to another	Cattle ID	The identification number of the sick animal
New Owner detials	The identity of the person who has purchased the cattle	Location Found	The location where the animal was found sick

Table 20 Glossary of terms for the class Traceability and class Disease.

Organizations (Vet, MOA, BMC, Police) and Visitations Class data Properties of Glossary of Terms

Table 21 depicts the glossary of terms for the class Organizations on the left side of the table and class Visitations on the right side of the table. Each data property for both classes is provided with a description of the term.

Organizations Data Property	Description	Visitations Data Property	Description
ID	Identity number of the organization	Full Name	First and last name of a visitor entering a farm
Name	Name of the person/organization	Contact	The phone number of the visitor
Address	office address of the organization	Reason	Detailed reason for the visit to the farm
Contact	Phone number of the organization		
Service Provided	The type of service provided by the organization		

Table 21 Glossary of terms for the class Organizations and class Visitations. Glossary of terms for the class Organizations and class Visitations.

Vaccination Class and Farm Class data Properties of Glossary of Terms

Table 22 depicts the glossary of terms for the class Vaccination on the left side of the table and class Farm on the right side of the table. Each data property for both classes is provided with a description of the term.

Vaccination Data Property	Description	Farm Data Property	Description
Name	The name of the vaccination	Farm ID	The unique number that identifies each farm holding
Brand	The type of brand of vaccination	Location	The GPS coordinates of each farm location
Serial number	The serial number on the vaccination container	Zone	The district in each farm belongs and disease status of each district. Either green for safety or red for danger
Date bought	The date the farmer bought the vaccination from the shop	Farm details	The name and contacts of farmer
Date administered	The date the vaccination was administered to the cattle	Audit results	The results from veterinary after performing an audit at a farm for export regulations compliance
Expiry date	The expiry date written on the vaccination	Female Count	The number of female cattle a farm is holding
Dosage	the amount that is administered per cow	Male Count	The number of male cattle a farm is holding
Cattle ID	The identification number of each cow vaccinated	Total Count	The total number of cattle in a holding
Details	The use and reason for administering vaccines. Diseases it is preventing	Lost	The number of cattle that have been lost
Schedule	The dates each cow should be vaccinated	Stolen and Dead	The number of cattle that have been stolen and dead

Table 22 Glossary of terms for the class Vaccination and class Farm.

Financial and Inventory Class data Properties of Glossary of Terms

Table 23 depicts the glossary of terms for the class Financial on the left side of the table and the class Inventory on the right side of the table. Each data property for both classes is provided with a description of the term.

Financial Data Property	Description	Inventory Data Property	Description
Budget	The amount of money the farmer plans to use for each month	Item ID	Serial number of machine and tool
Overhead Expenditure	The fixed amount of money the farmer spends each month	Item Name	The name of the machine and tool
Variable Expenditure	The amount of money a farmer uses per month changes	Item Price	The price of the machine and tool were bought
Profit	The amount of money made by farmers after deductions	In farm	The tools that are currently on the farm
Contingency	The amount of money kept for emergencies	Borrower	The name of the person who has borrowed a tool a machine
Petty Cash	The amount of money spent on small unexpected daily	Date Borrowed	The date the machine or tool was borrowed
Salaries	The amount of money paid to employees each month	Date Returned	The date the machine or tool was returned
Equipment	The amount of money spent on equipment each month	Lost	The machines and tools that are lost
Feed	The amount of money spent on cattle feed each month	Stolen	The machines and tools that have been stolen
Vaccines	The amount of money spent on cattle vaccines each month	Out of Farm	The full list of machines and tools that have been borrowed or sent for repair

Transportation	The amount of money spent on the transportation of cattle each month	Farm area	The village/town the farm is located
Sale and profit	The total sales made monthly	Farm name	The name of the farm tools belong
Employee ID	The employee identities that are paid monthly and per job		
Cattle ID	The identity of all cattle sold and bought		

Table 23 Glossary of terms for the class Financial and class Inventory.

Market Class data Properties of Glossary of Terms

Table 24 depicts the glossary of terms for the class Market. The data properties for the class are provided with a description of each term.

Market Data Property	Description
Cattle ID	The identity of all cattle sold and bought
Farmer ID	The identity number of the farmer making cattle transactions
Quantity	The number of cattle transacted
Cattle price (per cow)	Price set for each cow transacted
Total	The total amount paid for cattle transactions
Sale dates	The dates cattle were bought and sold

Table 24 Glossary of terms for the class Market.

B. Building the Concept of Taxonomies

The development of concept taxonomies was carried out to classify the glossary of terms into hierarchical structures. In the animal traceability taxonomy, Figure 40 depicts the taxonomy properties: disjoint decomposition, exhaustive decomposition, subclass, and partition. Disjoint decomposition means that an instance can only belong to one subclass, and the subclasses are mutually exclusive. In Figure 40, "Device" and "Location" are disjoint subclasses of "Tracking". In exhaustive decomposition, the instances of the class should be covered by its subclasses, so each instance should belong to at least one subclass. Exhaustive subclass is depicted in Figure 40, showing that "Farmer" is exhaustively decomposed into "Commercial" and "Communal"; this means every farmer is either commercial or communal. The next property is partition, which is both an exhaustive decomposition and disjoint. In partition, the classes are divided into non-overlapping subclasses that collectively cover all the instances of the superclass. Figure 40 depicts that "Records" is partitioned into "Employee", "Vaccination", "Disease", "Market", and "Financial". Lastly, the properties in the subclass show that the subclasses "Veterinary, BMC, MOA" all inherit the properties of their superclass "Organisation". The animal traceability taxonomy is critical for effectively structuring knowledge hierarchies and facilitating data organization and reasoning that is enhanced in animal traceability systems.

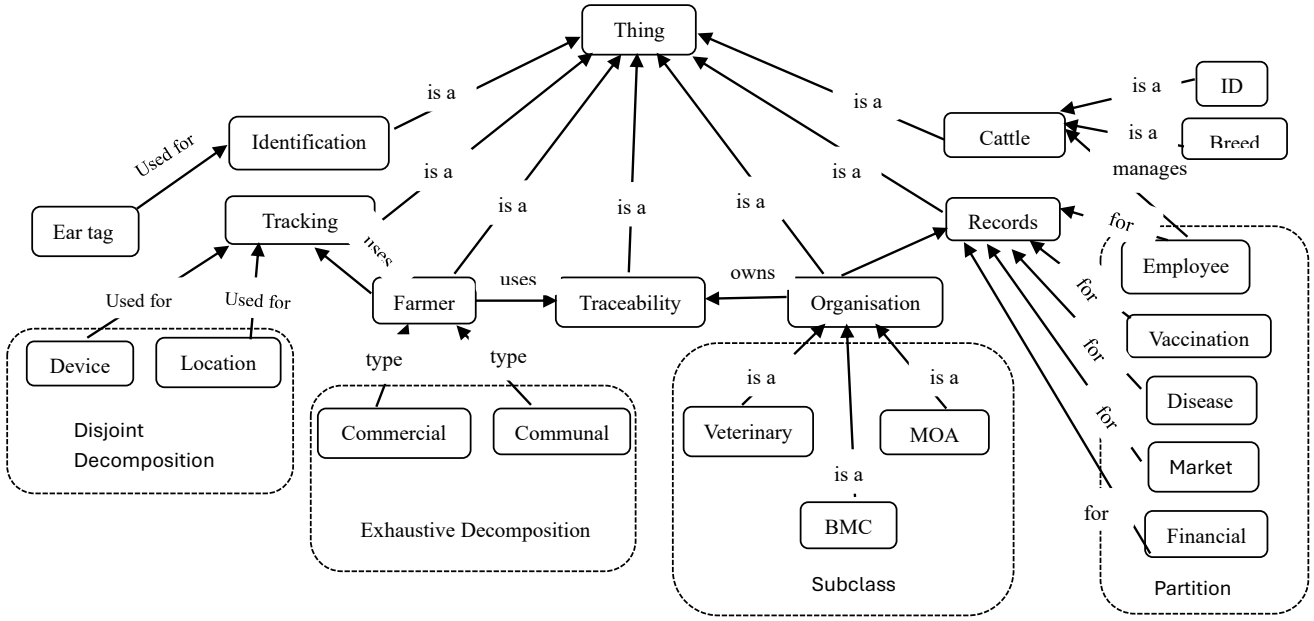
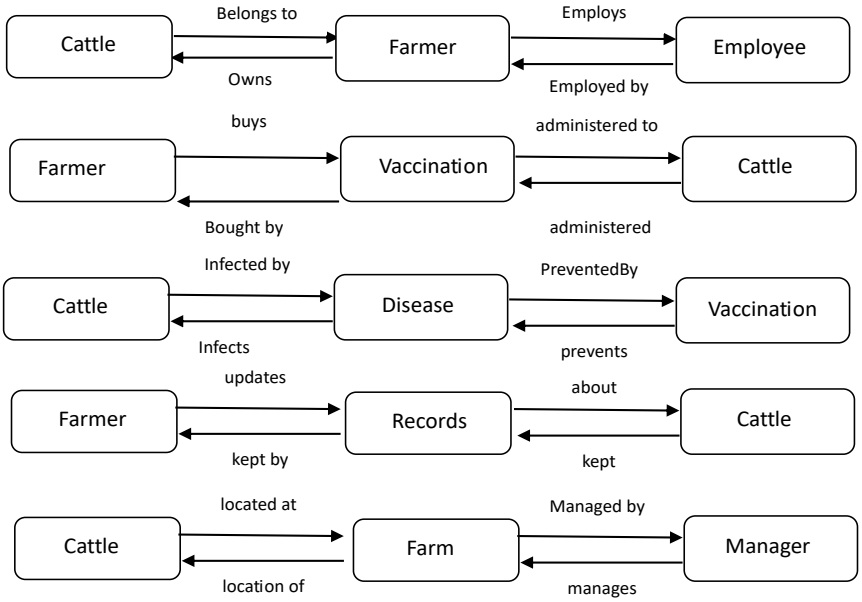


Figure 40 Taxonomy for animal traceability.

C. Developing the ad hoc Binary Relations

The ad hoc binary relations are built next after completing the taxonomy and concept classification trees. Relationships between concepts of the same taxonomy were identified in this step to map interrelationships that will form the required map of knowledge. Figure 41 shows the ad hoc binary relations built for the animal traceability ontology.



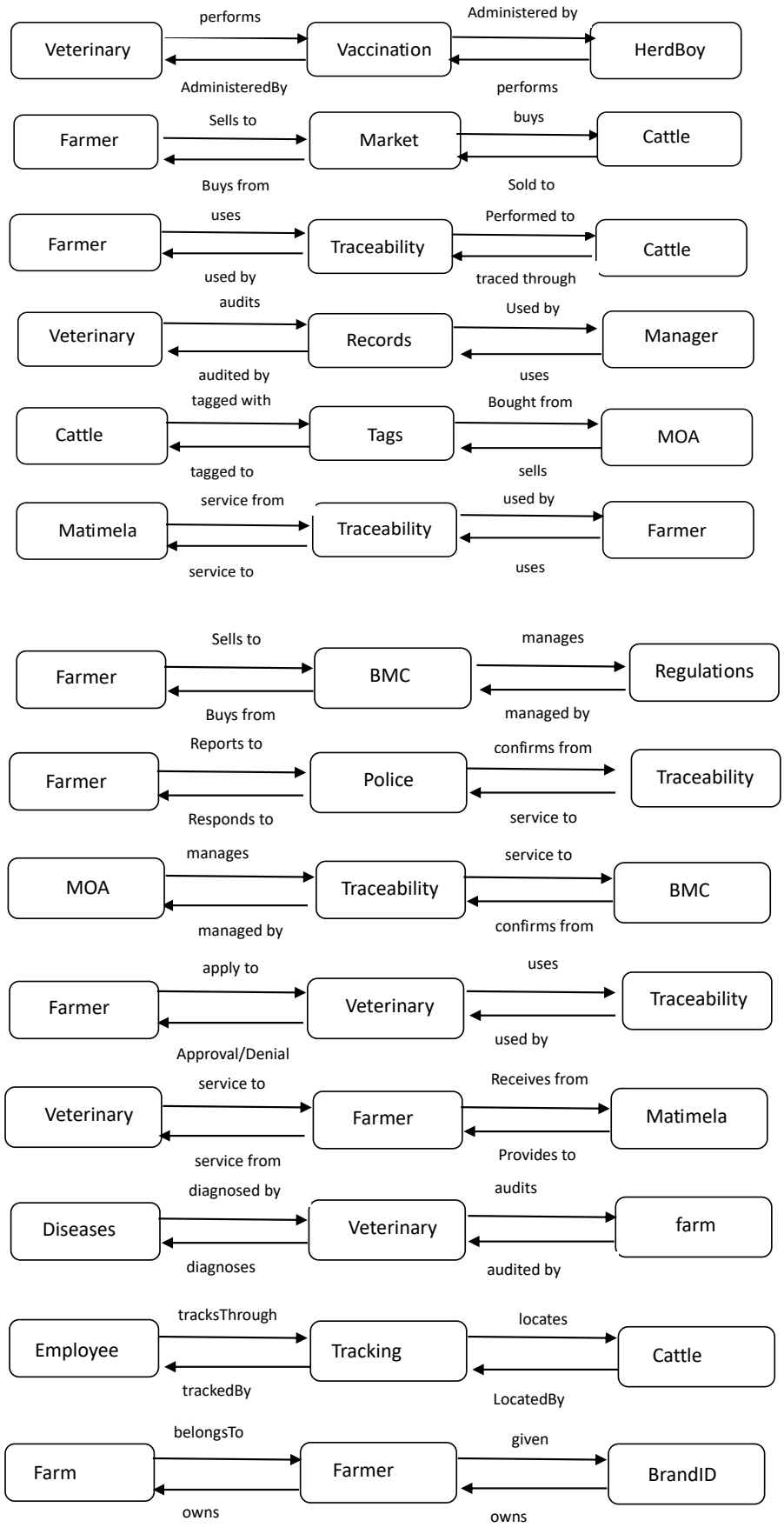


Figure 41 Ad hoc binary relations built for the animal traceability ontology.

D. Semantics of the Binary Object Properties

Table 25 describes the semantics of the binary object properties from the binary relations shown in Figure 41. The object property belongsTo is functional, meaning that each cattle is owned by only one farmer at a time. There is no inverse functional because a farmer can own multiple cattle. It is not transitive, as ownership does not propagate transitivity. There is no symmetric as no ownership is a directional relationship.

Object Property	Domain	Range	Functional	Inverse Functional	Transitive	Symmetric
belongsTo	cattle	farmer	yes	no	no	no
employs	farmer	employee	no	no	no	no
infectedBy	cattle	disease	no	no	no	no
preventedBy	disease	vaccination	no	no	no	no
keptBy	records	farmer	no	no	no	no
locatedAt	cattle	farm	yes	no	no	no
manages	manager	farm	no	no	no	no
performs	veterinary	vaccination	no	no	no	no
buys	market	cattle	no	no	no	no
usedBy	traceability	farmer	no	no	no	no
taggedWith	cattle	eartag	yes	yes	no	no
sells	MOA	eartag	no	no	no	no
buysFrom	BMC	farmer	no	no	no	no
ServiceTo	traceability	police	no	no	no	no
providesTo	matimela	farmer	no	no	no	no
locatedBy	cattle	tracking	yes	yes	no	no
given	farmer	brandId	yes	yes	no	no

Table 25 Semantics of the Binary Object Properties

The object property locatedAt is functional, meaning that an animal is located at one place at a time. There is no inverse functional because a location can house multiple animals. It is not transitive because if an animal is located at a farm, and the farm is in a region, the animal is not directly located in the region. There is no symmetric as location is directional.

The object property employs is not functional as a farmer can employ multiple employees. There is no inverse functional because an employee might work for multiple farmers. It is not transitive because if farmer A employs employee B, and employee B supervises worker C, it does not imply that farmer A employs worker C. There is no symmetric as employment is a directional relationship; if farmer A employs employee B, it does not mean employee B employs Farmer A.

The object property taggedWith is functional since each cow is typically associated with a unique eartag. There is an inverse functional as each eartag is assigned to a single cow. It is not transitive because the tagging relationship does not propagate transitively. There is no symmetric because if cow A is tagged with eartag B, it does not imply eartag B is tagged with cow A

The object property locatedBy is functional because each cow is associated with a unique tracking device. There is an inverse functional as each tracking device is assigned to a single cow. It is not transitive because the location relationship does not propagate transitively. There is no symmetric because if Cattle A is located by Tracking Device B, it does not imply Tracking Device B is located by Cattle A.

The object property given is functional since a farmer assigns a unique BrandID to each cow. There is an inverse functional as each brandID is associated with a single farmer. It is not transitive because the assignment relationship does not propagate transitively. There is no symmetric because if a farmer gives a BrandID to a cattle, it does not imply that the BrandID gives the farmer to the cattle.

A. Building the Concept Dictionary

From the ad hoc binary relations, the concept dictionary is used to specify the properties, instances, and relations for the description of each concept from the taxonomy. According to Santoso [198] a concept dictionary is used to identify the instances of attribute components from the concepts, to be able to fully describe the concept in the form of knowledge. It works as a structured repository for the domain-specific knowledge. The concept dictionary should contain the domain concepts, their relations, instances, also their class and instance attributes [199]. The concept dictionary is built in Table 26, which contains the domain concepts, instances, class attributes, instance attributes, and concept relations. Each concept is given a full description of knowledge using the domain concepts as shown in Table 26. The class attributes, instance attributes, and relations are local concepts; thus, the names given have been repeated in other concepts. In the development of ATO, the concept dictionary is used to standardize the terms and relationships, to make sure there is consistency in semantic descriptions across multiple creation assets and systems.

Concept Name	Instances	Class attributes	Instance Attributes	Relations
Farmer	Farmer_Lesego Farmer_John	Type of control	National ID number First name Last name Age Gender Address Contacts Keeper ID number Nationality Date of birth	Employs Owns Buys Updates sellsTo uses reportsTo applyTo receiveFrom given
Herd Boy	Employee_Neo Employee_Leo	Type of managing	ID number First name Last name Gender Contacts Nationality Date of birth Employment history	employedBy Treats Manages Performs Uses
Manager	Manager_Mpho Manager_Kevin			
Organizations	MOA BMC Veterinary Police Matimela	Type of control	ID Name Address Contact Service Provided	Performs Audits Sells serviceTo manages buysFrom respondsTo confirmsFrom owns uses Approval/Denial

				providesTo diagnoses
Traceability	LITS, BAITS	Type of monitoring	Cattle ID Farmer ID (owner) Current owner name Current owner location New location transfers New Owner details Animal history	usedBy managedBy serviceTo performedTo
Tracking	RFID-Based animal tracking systems, Electronic identification tags	Type of monitoring	Cattle ID Farmer ID Location Device name Serial number Last updated Temperature Signal strength Track ID Track details	Locates
Vaccinations	Foot and Mouth Vaccine, Bovine Viral Diarrhea Vaccine	Type of prevention	Name Brand Serial number Date bought Date administered Expiry date Dosage Cattle ID Vaccine schedule Details	BoughtBy AdministeredTo AdministeredBy Prevents
Financial	Annual Valuation Record, Cash Book, , Payment reciepits	Type of monitoring	Cattle ID Employee ID Farm name Variable Expenditure Overhead Expenditure Profit Budget Contingency Vaccines Salary Sale Transportation Equipment Feed PettyCash Profit	Belongsto updatedBy about
Inventory	Equipment maintenance logs, Livestock inventory	Type of monitoring	Item ID Item name Item Price Borrower Date borrowed Date returned Item lost Item stolen In farm Out of farm	Belongsto updateBy about
Disease	Foot and mouth Bovine Respiratory		Type Cause Treatment	Infects diagnosesBy

			Cattle ID LocationFound Dosage	
Cattle	Cow123 Cow456		ID Ear tag ID Ear tag color Brand ID Owner ID Color Sex Date of birth Location Breed Calve count Arrival date Calve ID Departure date Sold and Bought price Market	trackedBy taggedWith tracedThrough kept locatedAt administered BelongsTo soldTo ManagedBy infectedBy
Visitations	Researcher1, Buyer1, Seller1	Type of monitoring	Full Name Contact Reason	
Farm	Sandveld_farms, Lesego_farms		Farm ID Farmer ID Farm name Location GPS Zone Audit results Female Count Male Count Total Count Lost Stolen Dead Cause of death	BelongsTo auditedBy locationOf managedBy
Market	BMC, Sandveld_farms		Cattle ID Farmer ID Quantity Cattle price (per cow) Total Sale dates	buysFrom buys

Table 26 Concept dictionary used to specify the properties, instances, and relations.

B. Describing the ad hoc Binary relations

In this step, a merger among the relations and attributes that were built in the concept dictionary step is conducted. This is done to incorporate all the components of the animal traceability ontology as illustrated in Relational Table 27. The relational table specifies the name of the relation, the name of the source and its target, and the inverse relation. Table 27 shows a relation called employs, the source is farmer and its target is manager, with an inverse relation Isemployed.

Relation Name	Source Concept	Target Concept	Inverse Relation
RegisterTo	Farmer	Traceability	Isregistered
Employs	Farmer	Manager	Isemployed
Owns	Farmer	Cattle	Isowned
hasTraceabilityInteractions	Organizations	Traceability	hasTraceabilityInteractions
Administers	Manager	Vaccination	Hasadministered
Treats	Manager	Disease	Treated
Tracks	Manager	Cattle	Tracked

Manages	Manager	Farm	managed
employedBy	Manager	Farmer	Employs
Belongsto	Traceability	MOA	Hastraceability
Locates	Tracking	Cattle	Islocated
HasPoliceInteraction	Tracking	Cattle	HasPoliceInteraction
BelongsTo	Records	Farmer	Hasrecords
hasFarmerInteraction	Records	Farmer	hasFarmerInteraction
AdministeredTo	Vaccination	Cattle	Isadministered
Prevents	Vaccination	Disease	Prevented
HasDiagnosis	Disease	Veterinary	Diagnosed
HasSigns	Disease	Cattle	hasSigns
HasFarmerInteraction	Regulations	Farmer	HasFarmerInteraction
soldTo	Cattle	Market	Issold
BelongsTo	Cattle	Farmer	Hascattle
Register	Cattle	Traceability	Isregistered
HasTraceabilityInteraction	Location	Cattle	HasTraceabilityInteraction
HasTrackingInteraction	Location	cattle	HasTrackingInteraction

Table 27 Merger among the relations and attributes built in the concept dictionary.

C. Describing Instance Attributes

This step aims to describe the instance attributes from the concept dictionary in detail. The instance attributes are described in Table 28. In OWL the instance attributes are defined as data type properties when developing the ontology in Protégé. Table 28 shows each data type property row containing a detailed description of the instance attributes with specific fields. The specific fields are the data type property, the name of the concept it belongs to, the value type, and its value range. The presented Table 28 shows a fragment of the animal traceability ontology instance attribute table. Other fields of instance attribute for concepts have not been included in the table because of limited space. The concept Farmer from the table has data type properties ID number, First name, and Age with value type integer and string. Furthermore, listing value ranges for each being ID number 0-10, First name 0-20, and Age 0-3.

Datatype property	Concept name	Value type	Value range
National ID number	Farmer	Integer	0-10
First name		String	0-20
Age		integer	0-3
employee ID	Employee	String	0-10
Phone number		Integer	0-15
Nationality		String	0-20
Name	Organizations	String	0-50
Physical Address		String	0-100
Postal address		String	0-100
Cattle ID	Traceability	Integer	0-10
Farmer ID		String	0-10
Location transfers		String	0-150
Owner transfers		String	0-150
Serial number	Tracking	Integer	0-25
Device name		String	0-100
Full name	Visitations	String	0-20
Reason		String	0-250
Brand	Vaccination	String	0-50
Serial number		String	0-25
Date		Integer	0-15
Expiry date		Integer	0-15
Overhead Expenditure	Financial	Integer	0-12
Profit		Integer	0-12
Budget		Integer	0-12
Item price	Inventory	Integer	0-12
Borrower		String	0-100
Diagnosis	Disease	String	0-150
Cause		String	0-15.0
Treatment		String	0-150

Ear tag number	Cattle	String	0-12
Brand ID		String	0-5
Owner ID		integer	0-10
Location name	Farm	String	0-50
Location GPS		String	0-20
Audit results		String	0-250

Table 28 Description of the instance attributes from the concept dictionary.

D. Class Attribute Description

Class attribute Table 29 describes in detail the class attributes that have been listed in the concept dictionary. The table details each class attribute with its defined concepts and value types. Type of managing class attribute is defined by the concept farm employee, and the value type String is one of the presented class attributes in Table 29. Class attributes are properties that are used for the description of concepts as a whole, they are not inherited by subclasses or their instances [199]. While instance attributes are used for the description of properties that are specific to individual instances of a concept, class attributes refer to the concept as a whole. From Table 28, FirstName and ID are datatype properties for a specific farmer concept, whereas the type of control in Table 29 is a class attribute of the farmer concept. The values of the type of control would be shared, public, or private, which indicates the type of ownership for cattle by a farmer.

Class Attribute name	Defined concept	Value type
Type of control	Farmer, Regulations, Organization	String
Type of managing	Farm employee	String
Type of monitoring	Traceability, Tracking, Records, Financial, Inventory	String
Type of prevention	Vaccination	String

Table 29 Description in detail of the class attributes.

E. Axioms Description

The animal traceability ontology uses OWL 2, and one of the main components of OWL 2 is axioms. Axioms are statements that show what is true in the animal traceability domain. The axioms permit relationships to be discovered among the class expressions [203]. Table 30 shows the class expressions in the form of subClassOf axioms, equivalentClasses axioms, disjointClasses axioms, and disjointUnion axioms. Table 30 presents the class expression, the corresponding axiom in the animal traceability ontology, and its interpretation.

Stating that each instance of one class expression is also an instance of another class expression is a subClassOf, a hierarchy of classes can be created. In Table 30, class “cattle is subClassOf calf, which means every instance of a calf is an instance of cattle. In equivalentClasses, the multiple class expressions are equivalent to each other. In Table 30 the equivalentClasses axiom states that “vaccinatedAnimal :objectSomeValuesFrom(:hasReceived :vaccination)”, which means that an animal is a vaccinatedAnimal if it has received at least one vaccination. DisjointClasses axioms express that multiple class expressions do not have instances that are common, that is, pairwise disjoint. From Table 30 “:farmer and :oragnization” are disjointClasses, which means that an individual cannot be both a farmer and an organization. The last is disjointUnion axiom that defines a class as a disjoint union of multiple class expressions, it covers constraints. The expression “:traceabilityMethod disjointUnionOf: branding, earTagging, GPSMonitoring” states a disjointUnion that means each TraceabilityMethod is uniquely one of RFIDTagging, BarcodeLabeling, or GPSMonitoring, as seen in Table 30.

Class Expressions	Axiom	Interpretation
SubClassOf	:calf :cattle	Every instance of a calf is an instance of cattle
	:communal :farmer	Every instance of a communal is an instance of farmer
	:herdBoy :employee	Every instance of a herdBoy is an instance of employee
	:inventory :record	Every instance of a inventory is an instance of record
	:veterinary :organization	Every instance of a veterinary is an instance of organization
EquivalentClasses	:vaccinatedAnimal :objectSomeValuesFrom(:hasReceived :vaccination)	An animal is a VaccinatedAnimal if it has received at least one Vaccination
	:traceability objectSomeValuesFrom(:providesService :traceabilityService))	An entity is a Traceability if it provides a TraceabilityService
	:cattle objectIntersectionOf(:cow objectSomeValuesFrom(:taggedWith :earTag) objectSomeValuesFrom(:locatedAt :farm))	An individual is a Cattle if and only if it is a cow that has at least one taggedWith relationship to an EarTag and at least one locatedAt relationship to a farm
	:farmCattle objectSomeValuesFrom(:locatedBy :trackingDevice	An entity is farmCattle if it is located by a trackingDevice
	:farmer objectIntersectionOf(:person objectSomeValuesFrom(:owns :cow) dataSomeValuesFrom(:hasBrandID xsd:string))	An individual is a Farmer if and only if it is a Person who owns at least one cow and has a hasBrandID data property with a string value.
DisjointClasses	:cattle :sheep	No individual can be both Cattle and Sheep
	:farmer :organization	An individual cannot be both a farmer and an organization
	:employee :organization	An individual cannot be both an employee and an organization
DisjointUnion	:traceabilityMethod DisjointUnionOf: branding, earTagging, GPSMonitoring	Each TraceabilityMethod is uniquely one of RFIDTagging, BarcodeLabeling, or GPSMonitoring.
	:farmRecord DisjointUnionOf: financialRecord, InventoryRecord, healthRecord	Every FarmRecord is exactly one of FinancialRecord, InventoryRecord, or HealthRecord
	:healthStatus DisjointUnionOf: healthy, sick, recovered	An animal's healthStatus is exclusively one of Healthy, Sick, or Recovered

Table 30 Axioms description of concepts definable in OWL.

F. Description of Instance

The last step illustrates relevant instances from the concept dictionary. The following tables describe each class of animal traceability ontology instance. Each instance table defines the concept name, instance names, attributes of the instance, and their values. The animal traceability ontology has many instance created but because of limited space, one instance is illustrated for each class in the following tables.

Instance table for class Farmer

Table 31 describes the instance Farmer of the instance table of the animal traceability ontology. The instance is of the concept Farmer illustrating the attributes with corresponding values for each attribute as presented in Table 31.

Concept name	Instance name	Attributes	Values
Farmer	Farmer_Lesego	National ID number	54625647
	Farmer_John	First name	Nayang Otlhe
		Age	35
		Gender	Male
		Address	Sandveld farms, Khwee, Central District
		Contacts	267 75668824
		Keeper ID number	5832657
		Nationality	Motswana
		Date of birth	06-05-1988

Table 31 Description of instance Farmer.

Instance table for class Farm Employee

Table 32 describes the instance Farm Employee of the instance table of the animal traceability ontology. The instance is of the concept Farm Employee illustrating the attributes with corresponding values for each attribute as presented in Table 32.

Concept name	Instance name	Attributes	Values
Herd Boy		ID number	469512756
	Employee_Neo	First name	Kealeboga Atang
	Employee_Leo		
		Gender	Male
		Age	40
		Address	Plot 547 Serowe
		Contacts	267 74593624
		employee ID	H301
		Nationality	Motswana
		Date of birth	19-07-1983
		Employment history	Last job at Maboane farms for 2 years. From 20/02/2022

Table 32 Description of instance Farm Employee.

Instance Table for Organization

Table 33 describes the instance Organization of the instance table of the animal traceability ontology. The instance is of the concept Organization illustrating the attributes with corresponding values for each attribute as presented in Table 33.

Concept name	Instance name	Attributes	Values
Organization	MOA	ID	Vet100
	BMC	Name	Central Veterinary
	Veterinary	Address	Plot 215 Palapye
	Matimela	Contact	267 4925100
	Police	Service Provided	Permit approvals and denials

Table 33 Description of instance Organization.

Instance table for Traceability

Table 34 describes the instance Traceability of the instance table of the animal traceability ontology. The instance is of the concept Traceability illustrating the attributes with corresponding values for each attribute as presented in Table 34.

Concept name	Instance name	Attributes	Values
Traceability	LITS, BAITS	Cattle ID	101
		Farmer ID (owner)	54625647
		Current owner name	Nayang Otlhe
		Current owner location	Sandveld farms
		New location transfers	Hardveld farms
		Trace ID	005
		New Owner ID	658426514
		New Owner name	Viola Davis
		Animal history	One owner only

Table 34 Description of instance Traceability.

Instance table for Vaccination

Table 35 describes the instance Vaccination of the instance table of the animal traceability ontology. The instance is of the concept Vaccination illustrating the attributes with corresponding values for each attribute as presented in Table 35.

Concept name	Instance name	Attributes	Values
Vaccination	Foot and Mouth Vaccine, Bovine Viral Diarrhea Vaccine	Name	SUPAVAX
		Brand	MSD Animal Health
		Serial number	G2643 (Act 36/1947)
		Date bought	25-05-2023
		Date administered	26-05-2023
		Expiry date	05-06-2023
		Dosage	2 mℓ
		Cattle ID	101
		Vaccine schedule	3 to 4 weeks after inoculation and lasts for at least 1 year
		Details	For the active immunization of cattle against anthrax, botulism and blackleg. do not slaughter animals for human consumption within 2 weeks of vaccination.

Table 35 Description of instance Vaccination.

Instance table for Tracking

Table 36 describes the instance Tracking of the instance table of the animal traceability ontology. The instance is of the concept Tracking illustrating the attributes with corresponding values for each attribute as presented in Table 36.

Concept name	Instance name	Attributes	Values
Tracking	RFID-Based animal tracking systems, Electronic identification tags	Cattle ID	101
		Farmer ID	54625647
		Location	GPS 22.5515° S, 27.1147° E
		Device name	Oyster Edge
		Serial number	OE52654
		Last updated	0hrs 02mins
		Temperature	38.2 C
		Signal strength	HIGH
		Track ID	1012

Table 36 Description of instance Tacking.

Instance table for Visitations

Table 37 describes the instance Visitations of the instance table of the animal traceability ontology. The instance is of the concept Visitations illustrating the attributes with corresponding values for each attribute as presented in Table 37.

Concept name	Instance name	Attributes	Values
Visitations	Researcher1, Buyer1, Seller1	Full Name	Henry Stan
		Contact	267 72556951
		Reason	To view the cattle for possible sale

Table 37 Description of instance Visitations.

Instance table for Financial

Table 38 describes the instance Financial of the instance table of the animal traceability ontology. The instance is of the concept Financial illustrating the attributes with corresponding values for each attribute as presented in Table 38.

Concept name	Instance name	Attributes	Values
Financial	Annual Valuation Record, Cash Book, , Payment reciepits	Cattle ID	101
		Employee ID	469512756
		Farm name	Sandveld1
		Variable Expenditure	P20,000.00
		Overhead Expenditure	P100,000.00
		Profit	P 50,000.00
		Budget	P25,000.00
		Contingency	P5 000.00
		Vaccines	P1500.00
		Salary	P5000.00
		sale	P30,000.00
		transportation	P2500.00
		Equipment	-
		feed	P10,000.00
		pettyCash	P500.00
		profit	P30,000.00

Table 38 Description of instance Financials.

Instance table for Inventory

Table 39 describes the instance Inventory of the instance table of the animal traceability ontology. The instance is of the concept Inventory illustrating the attributes with corresponding values for each attribute as presented in Table 39.

Concept name	Instance name	Attributes	Values
Inventory	Equipment maintenance logs, Livestock inventory	Item ID	T2
		Item name	Tractor
		Item Price	P250,000.00
		Farm name	Sandveld farms
		Farm area	Mmashoro
		Borrower	-
		Date borrowed	25-03-2023
		Date returned	-
		Item lost	Mr Lefika Nata
		Item stolen	-
		In farm	30-03-2023

Table 39 Description of instance Inventory.

Instance table for Disease

Table 40 describes the instance Disease of the instance table of the animal traceability ontology. The instance is of the concept Disease illustrating the attributes with corresponding values for each attribute as presented in Table 40.

Concept name	Instance name	Attributes	Values
Disease	Foot and mouth Bovine Respiratory	Type	pasteurollosis
		Cause	Bacterial infection, suspected for animal bite or scratch
		Treatment	Pasteurella isolates are susceptible to oral antimicrobials such as amoxicillin, amoxicillin/clavulanic acid, tetracyclines, fluoroquinolones (ciprofloxacin, ofloxacin, levofloxacin, moxifloxacin), and trimethoprim-sulfamethoxazole
		Cattle ID	101
		LocationFound	Sandveld

Table 40 Description of instance Disease.

Instance table for Market

Table 41 describes the instance Market of the instance table of the animal traceability ontology. The instance is of the concept Market illustrating the attributes with corresponding values for each attribute as presented in Table 41.

Concept name	Instance name	Attributes	Values
Market	BMC, Sandveld_farms BMC,	Cattle ID	101
		Farmer ID	54625647
		Quantity	50
		Cattle price (per cow)	P10,000.00
		Total	P500,000.00
		Sale dates	22-04-2023

Table 41 Description of instance Market.

Instance table for Cattle

Table 42 describes the instance Cattle of the instance table of the animal traceability ontology. The instance is of the concept Cattle illustrating the attributes with corresponding values for each attribute as presented in Table 42.

Concept name	Instance name	Attributes	Values
Cattle	Cow123	ID	101
	Cow456	Ear tag ID	5758
		Ear tag color	red
		Brand ID	1435
		Owner ID	54625647
		Color	tshweu
		Sex	Female
		Date of birth	06-05-2021
		Location	Sandveld farms
		Breed	1
		calve count	-
		Arrival date	20-04-2023
		Calve ID	50
		Departure date	22-04-2023
		Sold price	P500,000.00
		Bought price	-
		Market	BMC

Table 42 Description of instance Cattle.

Instance table for Farm

Table 43 describes the instance Farm of the instance table of the animal traceability ontology. The instance is of the concept Farm illustrating the attributes with corresponding values for each attribute as presented in Table 43.

Concept name	Instance name	Attributes	Values
Farm	Sandveld_farms, Lesego_farms	Farm ID	101
		Farmer ID	54625647
		Farm name	Sandveld
		Location GPS	GPS 22.5515° S, 27.1147° E
		Zone	Central, C10
		Zone status	Green
		Audit results	Good
		Zone	Central
		Female Count	2500
		Male Count	500
		Total Count	3000
		Lost	150
		Stolen	260
		Dead	20
		Cause of death	Foot and mouth

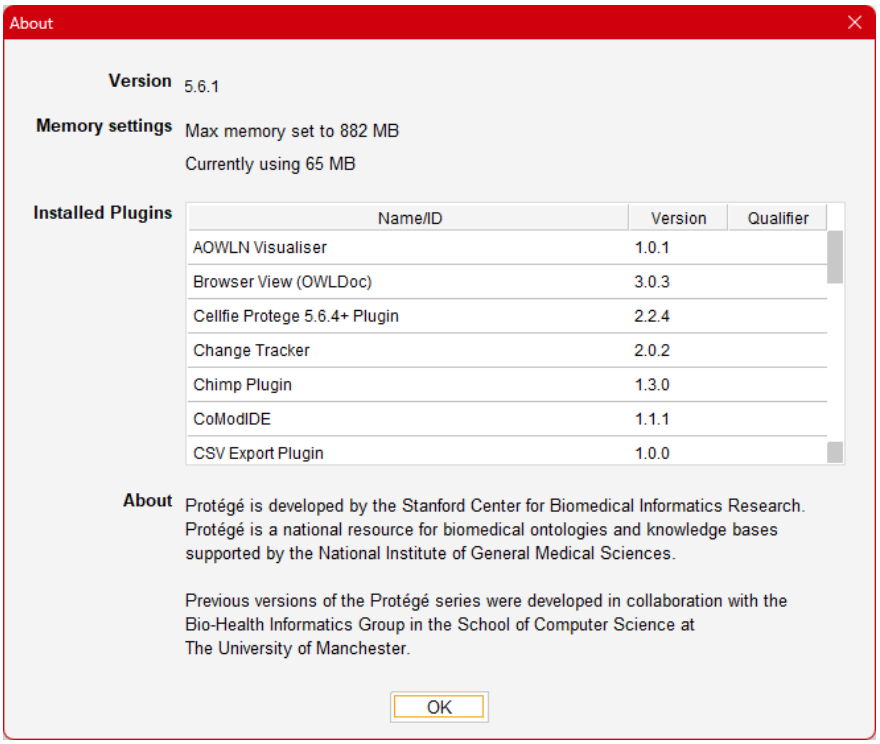
Table 43 Description of instance Farm.

5.4.2.6 Formalization and Implementation

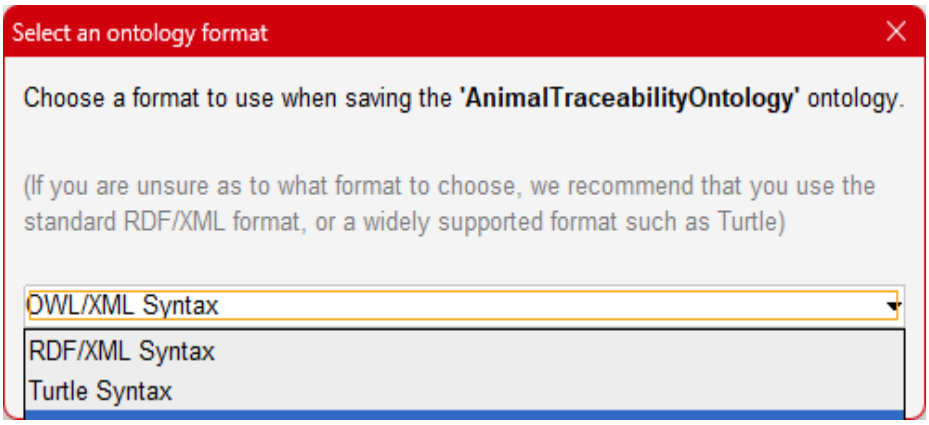
The next stage of animal traceability ontology is the transformation of the conceptual model to a semi-computable or formal model, that is the formalization stage. Figure 43 Below illustrates the steps that will be taken in the formalization stage. This stage is carried out in the ontology development tool Protégé 5.6.1 in OWL language, Figure 42.a shows the Protégé 5.6.1 environment that was used. Protégé is developed by Standaford Medical Informatics at Stanford University [204]. Furthermore, it is an open-source free ontology platform. Protégé enables developers to develop and populate formal knowledge-based and ontology applications straightforwardly. At the start of a new project, it is required

to choose a type of format for the ontology, the study chose the OWL/XML format as shown in Figure 42.b. Throughout the ontology development, the ontology reasoner will run to infer information that is not represented explicitly in the ontology.

The Pellet reasoner in Protégé was employed to provide the following services: consistency checking, equivalence checking, and instantiation checking. To visualize the Animal Traceability ontology protégé plugins OWLViz and OntoGraf were employed. OWLViz result created a knowledge-based graph of the Animal Traceability classes in PNG format. OWLViz is said by Kassani and Kassani [204] to be a highly configurable and powerful extension that provides a graphical description of the semantic relations to help users visualize the classes in OWL ontology.



(a) Protégé 5.6.1 Ontology Development Environment.



(b) Selecting language OWL Syntax.

Figure 42 Protege Startup process

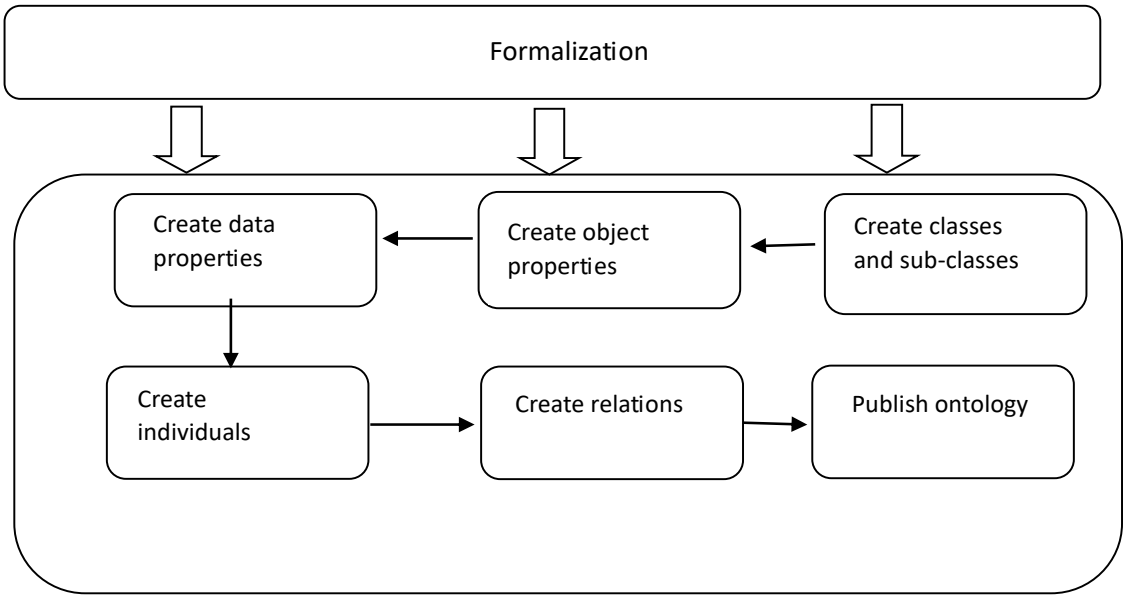
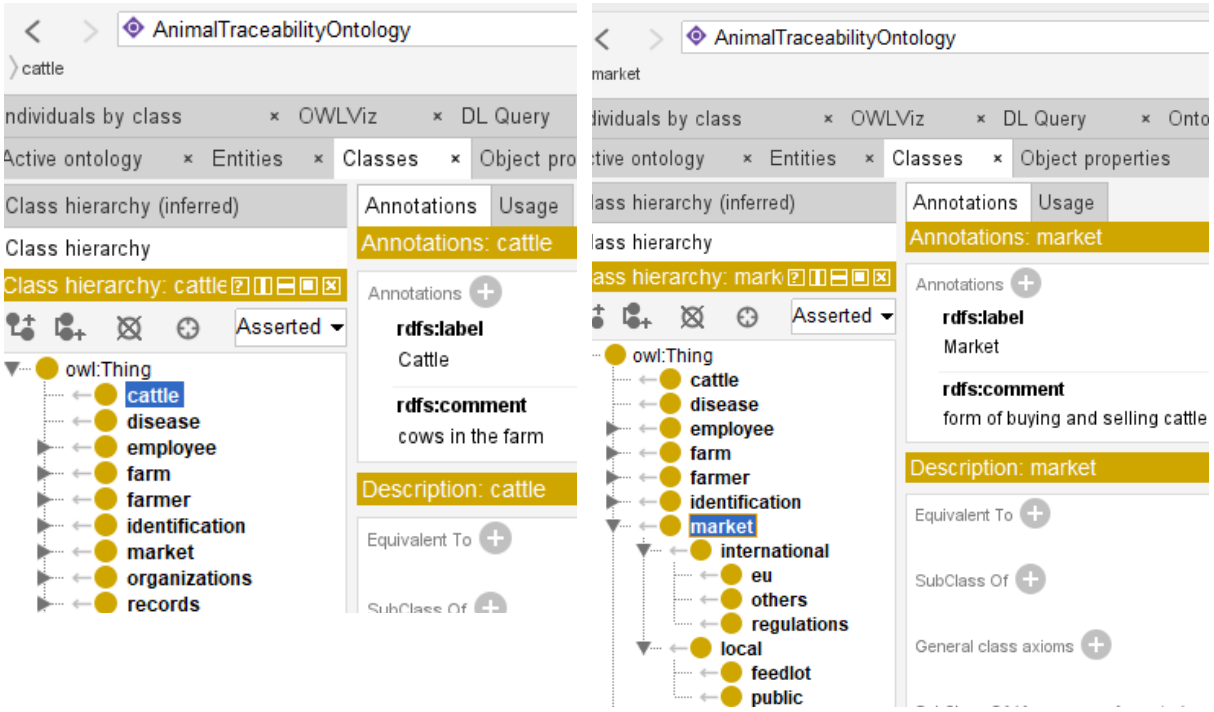


Figure 43 Formalization process for animal traceability ontology.

A. Creation of Classes and Sub-Classes in Protégé

Figure 44.a shows the created class "cattle", which is classified under owl:Thing. owl:Thing is found at http://www.w3.org/2002/07/owl as part of the OWL vocabulary, developed by W3C. Figure 44.b is the created subclasses "international" and "local" which belong to the superclass "market", then subclasses of local as "feedlot" and "public". The second and third-level classes describe more detailed and refined information about the superclass. The superclasses and subclasses have been arranged in an inheritance hierarchy. This was carried out with all the classes and sub-classes that were identified and described in the conceptualization stage. The complete list of classes and subclasses developed in the ontology is seen in Figure 45.



(a) Created class “cattle”.

(b) Created sub-classes of class “market”, as “international” and “local”, and sub-classes of “local” being “feedlot”, and “public”.

Figure 44 Class and Sub-class creation



Figure 45 Complete list of the animal traceability ontology classes, sub-classes, and sub-sub-classes.

B. Creation of Object Properties in Protégé

The following step is the created object properties for the animal traceability ontology. Figure 46 illustrates the object “taggedwith” created in Protégé.

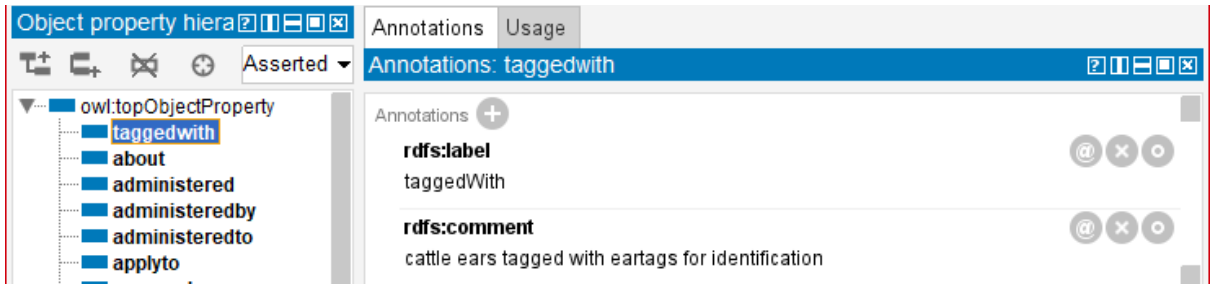


Figure 46 Creating object properties for the animal traceability ontology.

The OWL object properties describe the association between two entities of a given domain. The object “taggedwith” is created as shown in Figure 46, also given a lable and comment. Concepts are linked using the different types of defined properties. Figure 47 depicts the list of all the ATO object properties created in Protégé.

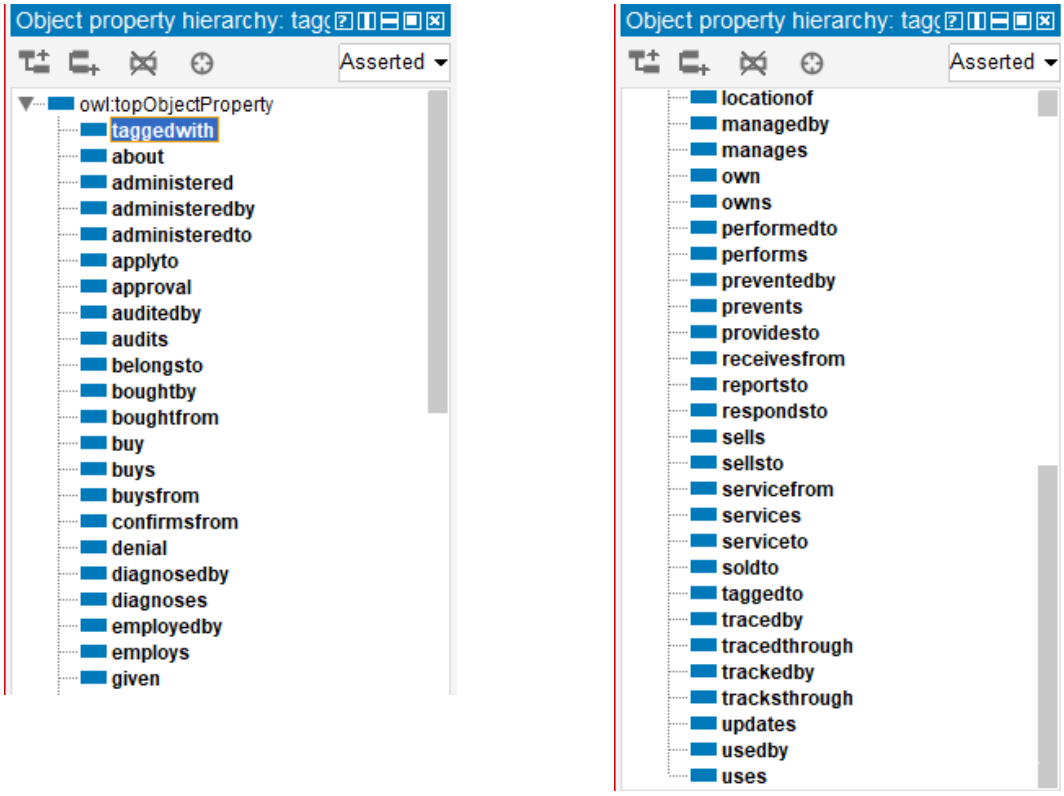


Figure 47 depicts the list of object properties created in Protégé for the animal traceability ontology.

Each of the object properties has defined domains and ranges. Figure 48 shows that the object property "taggedwith" is an inverse of "taggedto", the associations are domain "cattle" and range "identification", and its characteristics are "Functional" and "Inverse functional". Concepts are linked using the different types of defined properties.

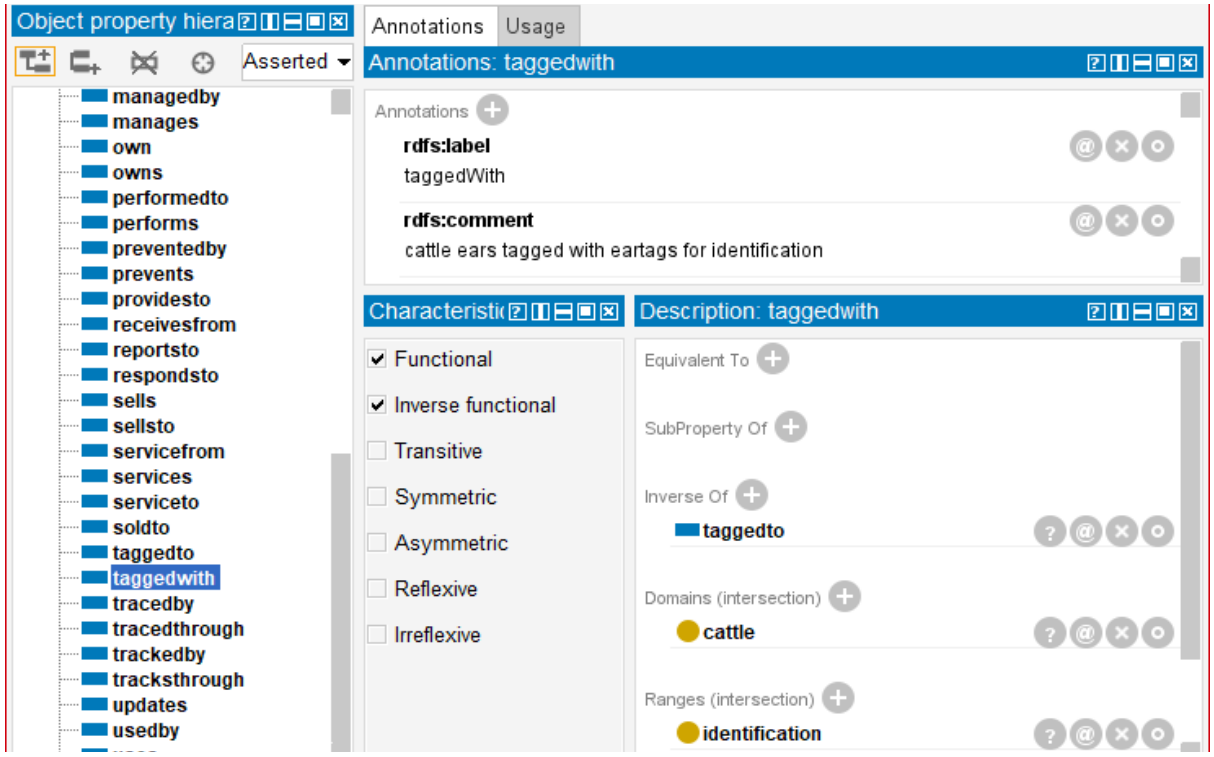
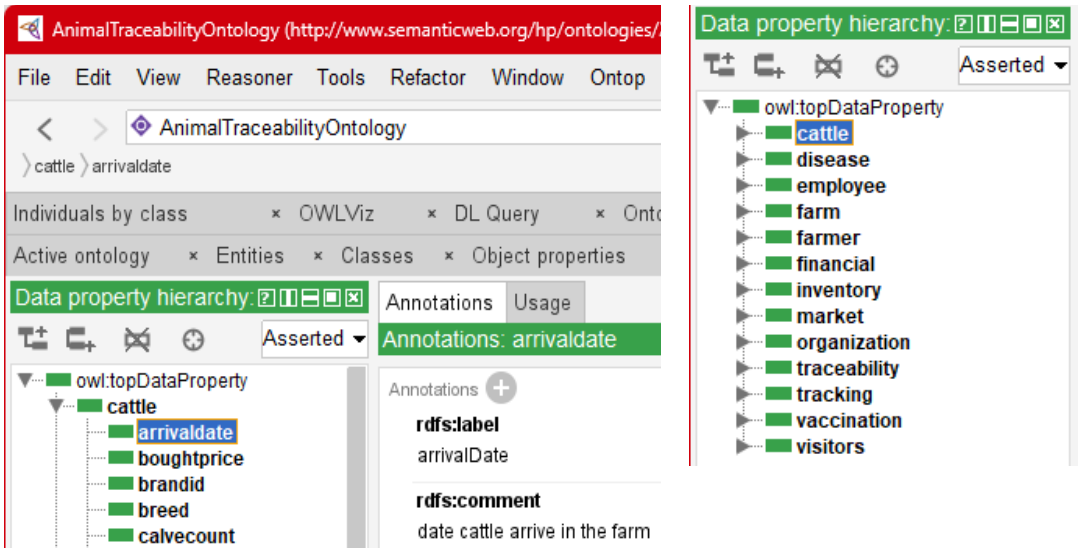


Figure 48 The domains and ranges for object property "taggedwith".

C. Creation of Data Properties in Protégé

The third step of ontology development is the creation of data properties. Figure 49.a illustrates the creation of data property “arrivadate” under the main class “cattle”. The data properties will be created under each class they belong. This was done for ease of use and understanding, rather than creating all data properties without their classes. All the classes that will have data properties under them were created for the ontology, and a full list is shown in Figure 49.b.



(a) Creation of data property ”arrivaldate” under class “cattle”

(b) All the classes that will have data properties created under.

Figure 49 Creation of Data Properties

Figure 50 shows all the data properties that were created for the ontology. Each of the data properties is under the class they belong to. For data properties like “firstname”, “lastname”, “age”, and “gender” that belong to more than one class, they are created under one class “farmer”. However, when instances are created, they will be used for the class “employee” also.

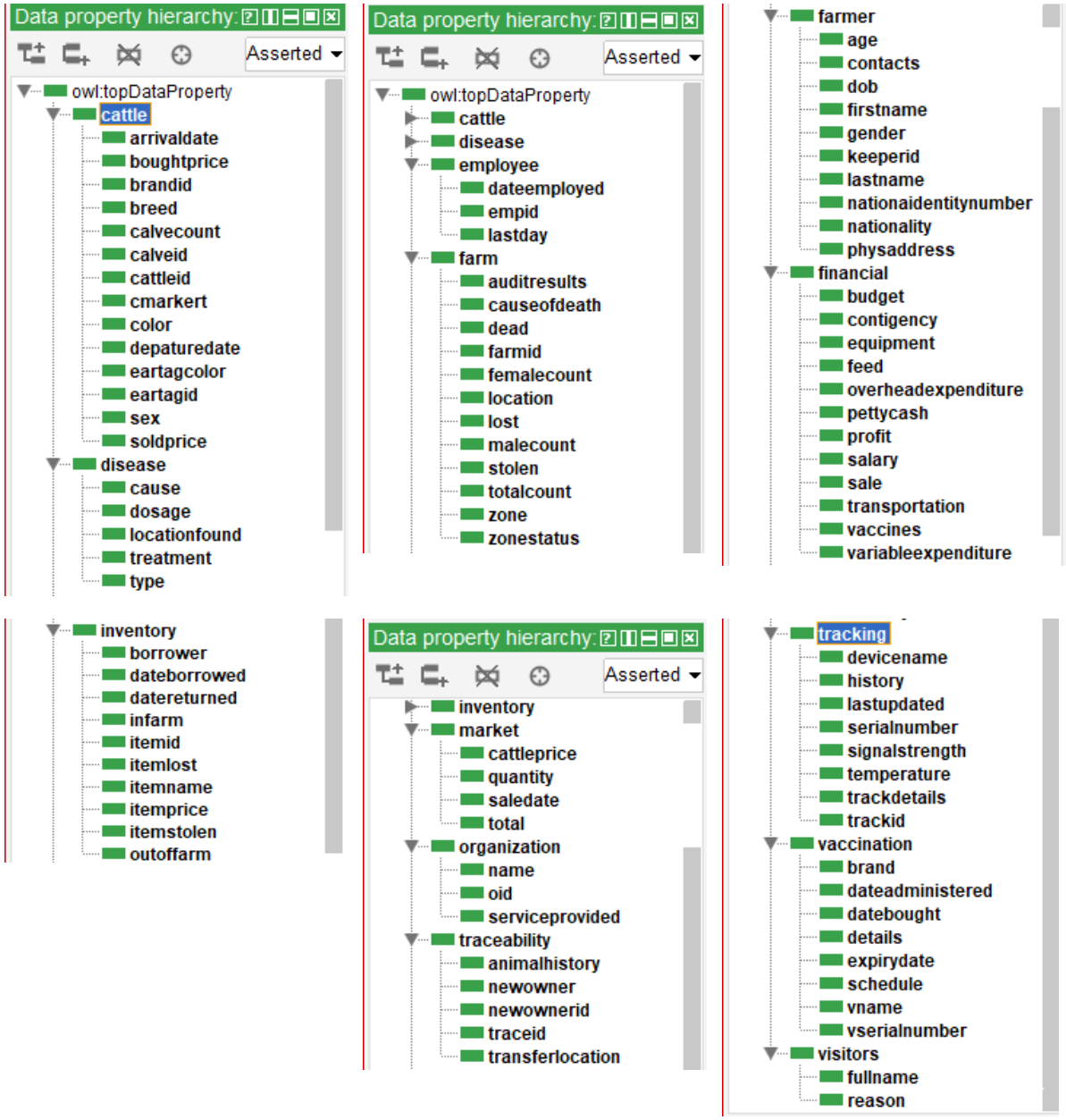


Figure 50 Created Data Properties of ATO using Protege

After the creation of the data types, they are assigned to classes. Figure 51 illustrates how each property had its domain defined, for the data property “arrivaldate” the domain was defined as “cattle”. The data type for the data property “arrivaldate” was also assigned a data type “dateTime” as shown in Figure 51.

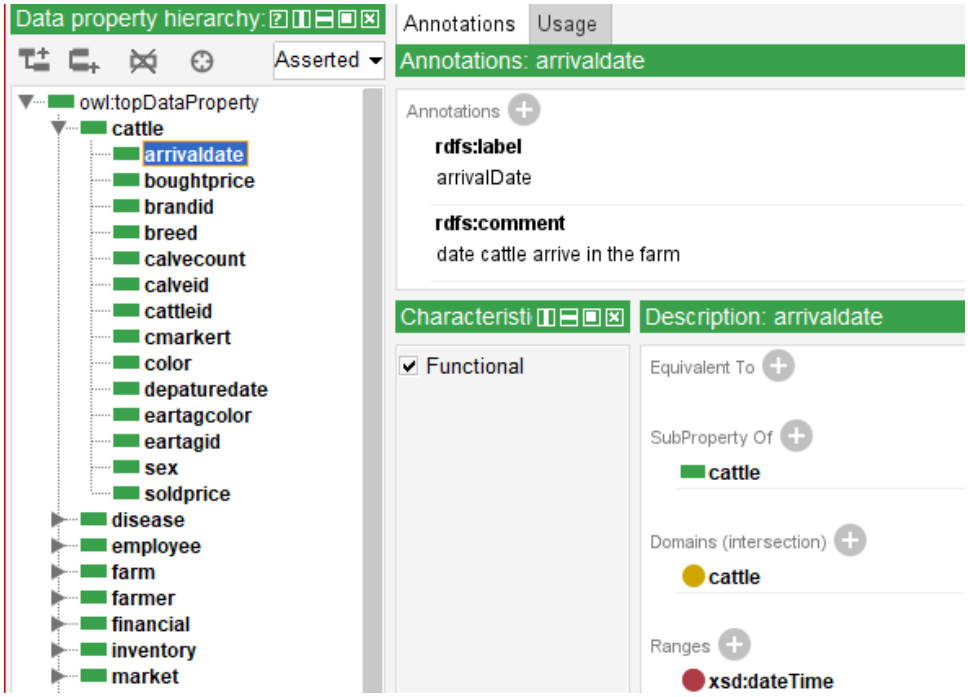


Figure 51 Adding features to data types

D. Creation of Individuals by Class in Protégé

Individuals in the ontology are the members of a defined class that have set constraints. The next step in the development of the animal traceability ontology is creating the individuals. To effectively describe the representation of knowledge, the constraints domain and range of the properties are used. The range defines the individual type that is the object in the RDF triple, and the domain defines the individual type that is the subject. The individuals in the ontology were added by manual entry. The instances were created directly in Protégé. Some of the data was sourced from the data that was collected in the field, for example, instance “Sandveld” of class Location, instance “Central” for class Zone, and instance “Foot and Mouth” of class Disease are real-world instances. Figure 52 illustrates the created individuals for the class “disease”. The individual “foot and mouth” is shown along with its description and property assertions. The property assertions are created by aligning object properties and individuals from other classes, for example object property “preventedby” is associated with the individual “aftovax” from the class “vaccination”.

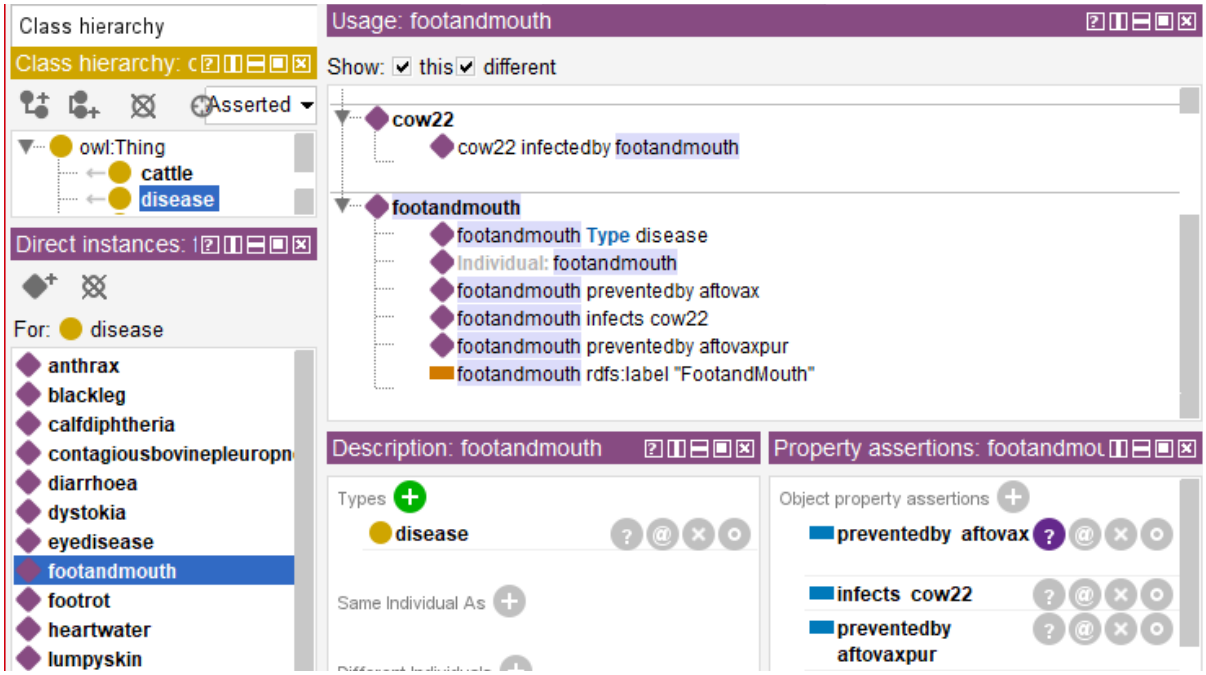
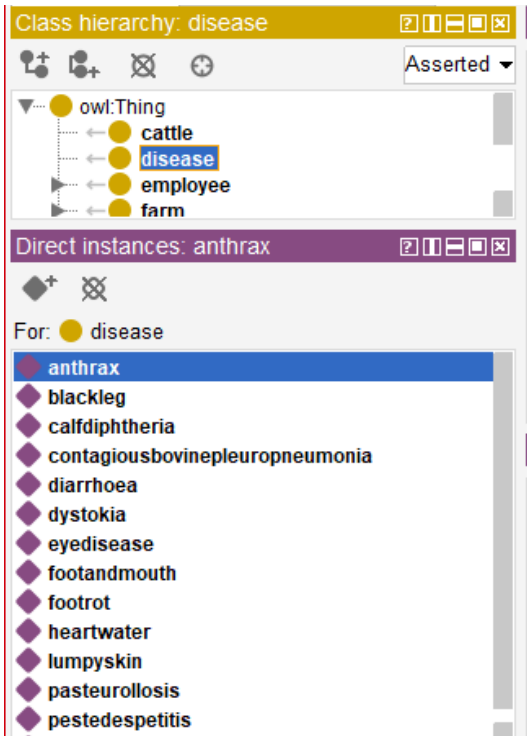
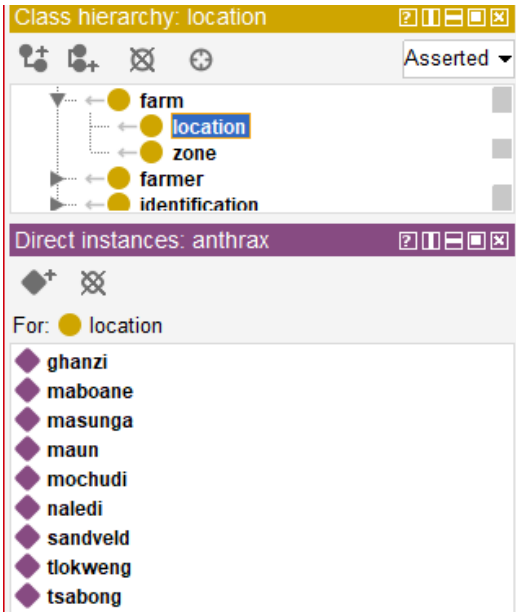


Figure 52 Creation of individual "foot and mouth" for class disease.

A complete list of individuals for class disease created is shown in Figure 53.a, listing the various diseases that infect cattle. More individuals were created for all the ontology classes and sub-classes, ans example of subclass location instanes is shown in Figure 53.b.



(a) Depiction of individuals created for class disease.



(b) List of the individuals that belong to subclass location of class farm.

Figure 53 Individuals for animal traceability

After the individuals are created, they are allocated to the different classes giving them meaning and details. The data properties and object properties created are used in this stage to assign properties to the individuals and to create all the relations for the classes of the ontology. For example, the “Farmer 1” individual will be assigned specific properties using data properties “firstname”, “lastname”, “age”, “gender”, “contacts”, “address”, and others. The object properties of “Farmer 1” that relate it to other individuals, for example, will be “employs” relating it to “employee”, object property “owns” related to “cow1”, object property "buys” related to the vaccine, and many more. The limitation of space restricts presenting all the data, relations that were created, and all ontology developments. The animal traceability ontology is published and all details can be found online at Figshare.com.

5.4.2.7 Results of the Animal Traceability Ontology

A. Model of the Animal traceability ontology - OWLViz



Figure 54 Pictorial from Protégé OWLViz of the animal traceability ontology.

OWLVi a plugin from the Protégé tool that creates a pictorial of the animal traceability ontology classes as seen in Figure 54. Animal Traceability ontology acts as a base for developers and researchers to perform reason and query in the knowledge-base environment. Figure 54 depicts the superclass, subclasses, sub sub-classes on the ontology. Class “Market” is linked with sub-classes “local” and “international”, they have further linked sub-classes “public”, “EU”, “feedlot”, “regulations”, and “others”. According to Kassani and Kassani [204] owl:Thing class is a representation of the concept of a user-defined class for reasoning facilitation. This visualization demonstrates how all classes and subclasses are organized under the root class owl:Thing, emphasizing the structured and hierarchical nature of ontology. By visually representing these relationships, OWLViz facilitates understanding how various entities within the animal traceability domain are interrelated.

B. OntoGraf

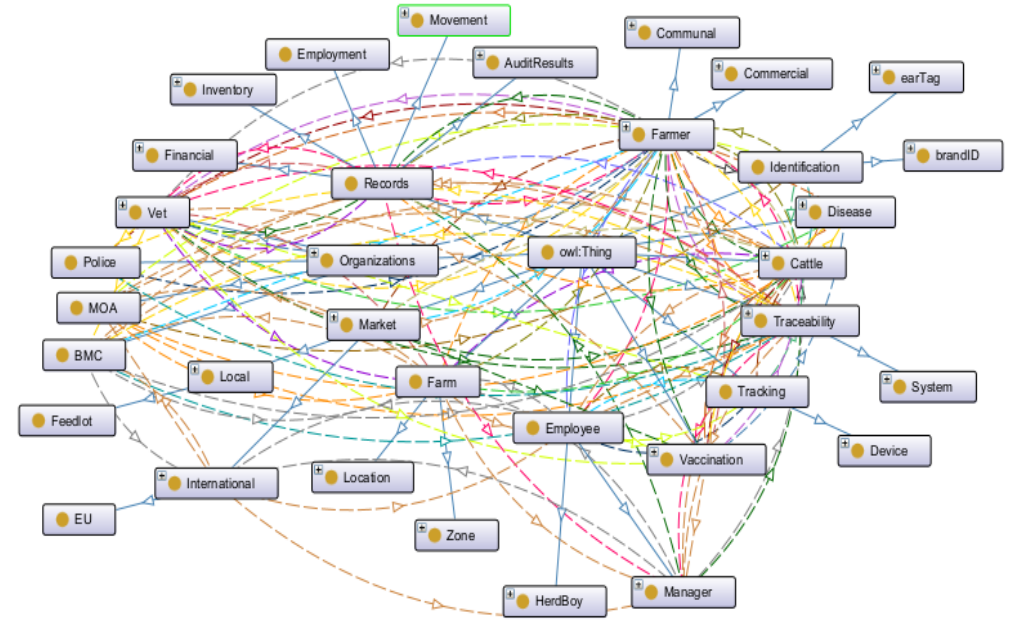


Figure 55 Pictorial from Protégé OntoGraf of the animal traceability ontology.

Visualization with Protégé OntoGraf, illustrated in Figure 55, provided additional insights by map-ping out the interrelationships among different classes using dotted arrows to show connections. These visual indicators of relationships demonstrated how individual components, such as "cattle", "farmer", "veterinarian", and "traceability", interacted within the overall system to achieve a unified approach to animal traceability. The OntoGraf tool highlighted the interconnectedness of these components, showing that although each class was developed individually, they were designed to work together seamlessly to support the comprehensive tracking and management of animal-related data.

The use of both OWLViz and OntoGraf ensured a robust visualization of the ontology's structure and interrelations, making it easier for developers, researchers, and stakeholders to comprehend and utilize the ontology. These visual tools not only demonstrate hierarchy and relationships but also validate the ontology's consistency and completeness, ensuring its reliability for practical appli-cations.

5.5 Chapter Summary

Building the ontology for Animal Traceability determining concepts, object properties, data properties, and individuals will provide Animal Traceability stakeholders to efficiently infer and extract knowledge that is valuable from the data. The ontology development tool Protégé was used to develop the ontology in stages as illustrated in the chapter. The ontology is expected to satisfy the identified requirements (Objective 1), and formal representation of knowledge requirements from the proposed Ontology-driven Animal Traceability Framework (Chapter 5). After building the ontology the next step is to use Apache Jena Fuseki server to execute SPARQL queries and extract triple information via the web. As well as perform ontology evaluation through an ontology evaluation tool called OntoMetrics. The following Chapter 6 will provide illustrations and descriptions of executing SPARQL queries and computing matrices through OntoMetrics. Evaluating and validating ontologies is important in avoiding excessive concepts, subclass relationships that are incompatible, and ambiguity of terminology.

6 CHAPTER 6 ONTOLOGY EVALUATION

6.1 INTRODUCTION

The purpose of this chapter is to present a critical and detailed evaluation of the animal traceability ontology that was detailed in Chapter 5. Evaluating the ATO is important to ensure that it is effective and can be applied in the real-world context of agriculture and supply chain. The aim is to determine how the ATO can capture the animal traceability domain, simplify main operations (identification, animal movements, regulatory complaince), and support cross system interoperability. The importance of ontology evaluation persists because ontologies are in constant development thus standardization and improvements are required [205]. According to Amith, et al. [206], the process of determining the ontology quality based on the set of evaluation criteria, considering the type of ontology and its purpose is known as ontology evaluation. McDaniel and Storey, [207] agree that a set of specified criteria or metric computations is important in the quality assessment of an ontology. Furthermore, using assessment metrics provides information that is quantifiable about the ontology attributes, which contributes to the interoperability among the systems.

This study conducted animal traceability ontology evaluation using various metrics from the OntoMetric web-based tool (https://ontometrics.informatik.uni-rostock.de/ontologymetrics). As stated by Achim [208] using metrics for ontology evaluation offers a means of quickly grasping the set of attributes. He further states the metrics provide objective methods of ontology assessments offering information that is condensed, for example, classes to relation ratio and breadth and depth of graph. Together with metrics, Lantow [50] suggests using solutions that are web-based because they are provided publicly and independent from the ontology development environments. The advantage of techniques that are metric-based for ontology evaluation is the results that give a quantitative perspective of the quality of the ontology [209]. Moreover, the ontology is thoroughly scanned gathering various statistics of the knowledge contained in the ontology [209].

It is important to validate ontologies to determine the quality and correctness based on ontology quality criteria [50], [210]. The quality and correctness of an ontology are determined by metrics that are formally defined for the evaluation process. The main measure types used are structural measurement, functional measurement, and usability-profiling measurement [211]. Each one of the measurements has metrics that are used to calculate the quality of the ontology using defined ontology attributes. The quality and correctness of the ontology determines interoperability and formal standardization of the ontology. All the metrics and application results will help in determining if the animal traceability ontology is interoperable and has standardized definitions for organization integration. Functional evaluation was also conducted through SPARQL queries (Apache Jena Fuseki). OntOlogy Pitfall Scanner! (OOPS!) was used to evaluate pitfalls for ontology consistency and correctness.

This research will be also conducting the evaluation by comparing the animal traceability ontology and a similar ontology called animal health ontology from (http://datadrivensurveillance.org/animal-health-ontology-aho/) the National Center for Biomedical Ontology [212]. The process of comparing the ontologies is called ontology alignment, one ontology is called the reference, and the other candidate. The animal health ontology is called the reference and the animal traceability ontology is the candidate. Animal health ontology has been trialed, tested, and managed by the Swedish National Veterinary Institute. It is available online for use at GitHub and Bioportal. The comparison will help validate the animal traceability ontology by identifying issues of inconsistencies, errors, and redundancies. Ensuring semantic interoperability and compatibility. Overall identification of parts that need optimization and improvement to ensure the developed animal traceability ontology performs effectively and efficiently. To the best of our knowledge, the animal health ontology was the only ontology this research found with published ontologies closer to this research. The comparison will also validate the reasons for

developing the animal traceability ontology from scratch and not adopting an already existing ontology. Table 44 shows the 3 types of ontology evaluation used against the metric used to calculate them.

A combination of these evaluation approaches will assist in assessing the alignment and structural integrity of the ATO with domain's best practices and eisting standards. At the end of this through evaluation, the objective is to establish the ATO's suitability for deployment in the traceability systems.

Evaluation technique	Metric
1.Structural	Base
	Schema
	Graph
2.Functional	Knowledgebass
	Class
3.Usability	Object Properties
	Data Properties
	Class Axioms

Table 44 Evalution Techniques and Metrics

6.2 Structural Evaluation

According to Rebekka et al. [211] the structural measure evaluation process is focused on the formal semantics and syntax of the ontology. Dividino et al. [213] concur that the structural measure employs metrics that assess the topological dimension and logical adequacy of the ontology, thus the ontology graph structure syntax. Furthermore, the topological dimensions include breadth, connectivity, leaf, depth, tangledness, sibling distribution, and modularity. The schema metrics also provide attributes that measure the structure of an ontology. This would include depth, richness, inheritance, and width [213], [214]. In addition, Lourdusamy and John [215] says that the structure is also measured using the total number of properties, classes, individuals, and children maximum number. Formal semantics are found through the logical adequacy of the ontology [214]. Logical adequacy is the measurement of the ontology consistency and complexity. The logical adequacy is explained as the lack of contraction of facts deduced from the ontology model. That means an agreement between entities based on the semantic ontology language is a consistent ontology [213].

6.2.1 OntoMetrics Evaluation Tool

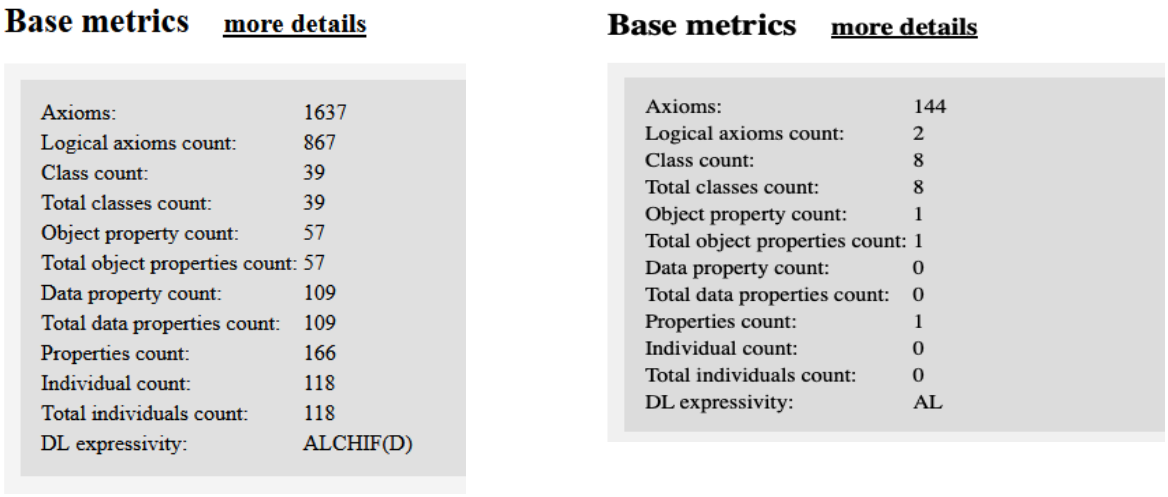
The main metrics used to evaluate ontologies on schema and instance metrics using five criteria: Base, Schema, Graph, knowledge, and Class. The quantity of the ontology elements has an effect on the ontology quality because it affects the ontology processes of reuse, alignment, and merging. The metrics used for evaluating the animal traceability ontology are adopted [209], [216], [217], [218]. The metrics adopted from researchers are computed automatically using the evaluation tool OntoMetrics developed by Michael Poppe and Martin Lichwark.

Lantow [219] emphasizes the need for ontology automatic calculations for the assessment of ontology quality. Furthermore, promoting knowledge development based on the ontology metrics application. The knowledge base and ontologies complexities require automated evaluation [50]. Manual evaluations will not suffice because of the major resources Lantow [50] and major time Lantow [219], Pizzutiet al. [96] that will have to be used for the metric computations. Various approaches have been suggested to ontology metric calculations for practical use and empirical evaluations but the web-based tool (OntoMetric) has been the best [216], [220], [221]. OntoMetric is the best because it is firstly independent and not attached to any ontology development environments like OntoCheck in Protégé which may cause bias to the ontology results. Secondly, OntoMetric is publicly available and easily and freely accessible to use at any time. Lantow [50] states that OntoMetric supports standardized ontology format OWL2 and RDF, collects a theory of the metrics and provides the validation, further analysis can be done from the XML machine-readable output that is provided. The first process using OntoMetrics to evaluate the Animal Traceability ontology is the OWL ontology file is uploaded to the

web-based tool, and the metrics Base, Schema, Knowledgebase, Class, and Graph are selected to be computed in the file.

6.2.1.1 Base Metric

The base metric is described by Gobin-Rahimbux [222] as a metric with basic computations that present the ontology element's quantity. The elements include counting the axioms, classes, data properties, individuals, and objects that are used in the ontology. According to Khan [217], the quantitative overview is important because the ontology size affects the alignment, merging, and reuse of the ontology process. Figure 56.a presents the base metric results from the OntoMetric tool for the animal traceability ontology and the reference ontology results in Figure 56.b. Overall base metric results are included in the computation of the ontology structure, to evaluate the quality of the ontology.



Base metrics more details

Axioms:	1637
Logical axioms count:	867
Class count:	39
Total classes count:	39
Object property count:	57
Total object properties count:	57
Data property count:	109
Total data properties count:	109
Properties count:	166
Individual count:	118
Total individuals count:	118
DL expressivity:	ALCHIF(D)

Base metrics more details

Axioms:	144
Logical axioms count:	2
Class count:	8
Total classes count:	8
Object property count:	1
Total object properties count:	1
Data property count:	0
Total data properties count:	0
Properties count:	1
Individual count:	0
Total individuals count:	0
DL expressivity:	AL

(a) Base Metric results from computing AnimalTraceabilityOntology.owl file.

(b) Base metric results from computing AnimalHealthOntology.owl file.

Figure 56 Base Metric

6.2.1.2 Axioms

The axioms are the main component in the ontology that presents the truth in the domain. The computed results from Figure 56.a indicate that the developed ATO encompasses 1637 axioms. This illustrates that a logically rich and comprehensive framework is designed to model the complexities of cattle identification, health, movement, regulatory compliance, and cattle ownership. These axioms spread across multiple categories of the ATO, including data and property characteristics, class hierarchies, logical constraints, and individual assertions. Class hierarchies were established for the representation of animal traceability entities. For example, subclass :calf :cattle axiom means that all the instances of "calf" are also instances of "cattle". The object property features define the relationships among individuals, with the axioms defining their features, like functional properties and domain and range. Functional property ensures the relationships are unique, for example, "FunctionalObjectProperty (:isLocated)" axiom means that each cow has one location. For data property axiom "DataPropertyDomain (:hasEarTag :Cow), DataPropertyRange (:hasEarTag xsd:string)" defines "hasEarTag" applies to "Cow" and the value is string. Data type defines the classes and their data types. This axiom indicates that Bessie has FarmerJohn as an owner.

Individual assertions are defined by axioms such as object property assertions, for example, "ObjectPropertyAssertion (:isEmployed :Lesego :FarmerNayang)". The axiom indicates that Lesego is employed by FarmerNayang. Logical axioms of ATO are used to enforce constraints and for enabling reasoning. Such as the axiom "DisjointClasses (:MovementRecord :FinancialRecord)" ensures that a record cannot be both a MovementRecord and a FinancialRecord. This prevents overlapping of class

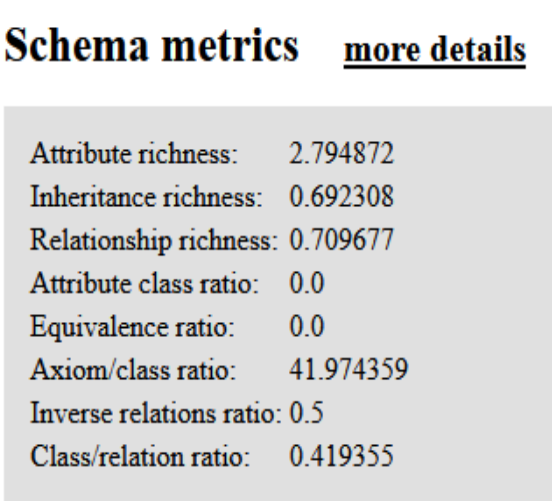
memberships. My results are consistent with researcher Gobin-Rahimbux [222], that axioms are for datatype definitions, classes, assertions, and properties. Axioms are important because they make sure that the ontologies developed are interoperable with the other ontologies and systems [217]. It is achieved by standardized definitions of constraints and relationships in an ontology. The axiom values also determine the ontology's ability to give structured and formal domain knowledge representation. Furthermore, the representations are used for decision-making, knowledge management, and reasoning.

The logical constraints axioms improve the data quality by ensuring that the data is accurate and consistent, which is critical for animal traceability effectiveness. The standardized relationships and terminology enhance interoperability by enabling data exchange that is seamless between the stakeholders (farmers, veterinarians, matimela officer, MOA, and BMC).

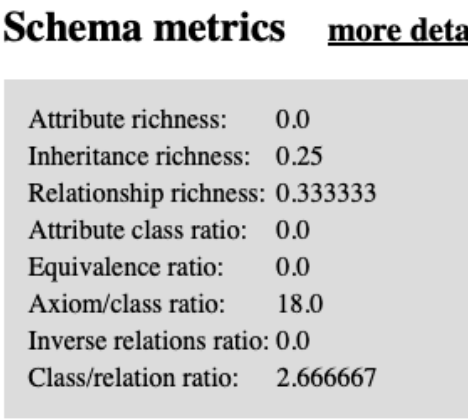
Axioms can be used to represent the size of the ontology. Figure 56.a indicates that the developed ATO contains 1637 axioms. The comparison Animal Health ontology (Figure 56.b) contains 144 axioms. Animal traceability ontology is not a small ontology but rather a medium-sized ontology, which means 1000 axioms is considered the optimum size; the recorded number is 1178. The animal health ontology (Figure 56.b) is small and thus recorded 144 axioms, which is smaller than the recommendation. Khan [217] interprets axioms based on the size of the ontology, he recommends that for small ontologies, the axiom value should be a minimum of 250, and 500 for an optimum size. As the size increases, the value will double.

6.2.1.3 Schema Metrics

The ATO was evaluated to assess the effectiveness of its design and structural quality using schema metrics, computed through the Ontometrics evaluation tool. The use of schema metrics to analyse the ontology design is aligned with research by Gobin-Rahimbux [224]. The schema metrics give insights into the ATO's relational complexity, depth, and breadth. This is important to ensure interoperability, knowledge management that is effective, and detailed domain representation. Figure 57 presents the results from the Ontometrics tool of the ATO schema metrics, showing inheritance richness, relationship richness, and class/relation ratio. Tartir [211] with this study's choice of schema metric to use, he lists inheritance richness, attribute class ratio, axioms ratio, attribute richness, and relation ratio to evaluate in an ontology.



(a) Animal Traceability Ontology Schema metrics results from OntoMetrics.



(b) Animal Health Ontology Schema metrics results from OntoMetrics.

Figure 57 Schema Metrics

A. Inheritance Richness (IR)

IR is used to measure the average number of subclasses per class in the ontology [218]. It stipulates how the distribution of knowledge is across hierarchy of the ontology [219]. The IR for ATO was computed through Ontometrics tool, using Equation 4.

$$IR = \frac{\sum_{C_i \in C} |H^C(C_1, C_i)|}{|C|}$$

Equation 4 Inheritance Richness (IR)

Where:

- Subclasses (C1) for class (Ci) - are defined as $|H^C(C_1, C_i)|$, total number of subclass (inheritance) relationships.
- (C) - defined as the ontology total number of classes.

The ATO has IR results of 0.692308 (Figure 57.a), indicating an ontology structure that is moderately detailed. The balance facilitates precision in the representation of animal traceability, interoperability and data with external stakeholder systems. Detailed knowledge encourages seamless data collaboration and integration, facilitating easy interoperability with external ontologies and systems.

The results interpretation is consistent with the IR results definition, which states that an ontology with low IR (closer to 0) is considered to be a deep ontology called vertical; it has very detailed knowledge [218]. High IR (approaching or exceeding 1) is a shallow ontology called horizontal; it has a wide range of knowledge that is general [218].

B. Relationship richness (RR)

The RR of the schema metric evaluates the diversity of relationships in the ATO by measuring the non-inheritance relationships to the total number of relationships. Tartir explains that an ontology containing inheritance relationships only typically conveys minimal information, unlike an ontology containing a set of diverse relationships [211]. RR is a representation of the non-inheritance relationship percentage among classes in comparison to the possible connections, including the non-inheritance and inheritance relationship. Equation 5 is used to calculate RR.

$$RR = \frac{|P|}{|H| + |P|}$$

Equation 5 Relationship richness (RR)

Where:

(P) – defined as the number of non-inheritance relationships, e.g., disjoint classes, object properties, and equivalent classes.

(H) - divided by the total number of relationships defined in the schema (the sum of the number of inheritance relationships). E.g., subclass relations.

The RR result of ATO is 0.709677 as presented in Figure 57.a, showing that a significant presence of diverse relationships in the ATO. The ontology will support queries that are complex and integrate with other systems and ontologies. ATO RR (0.709677) being closer to 1 than 0 suggests that ATO classes have strong connections, making it interoperable. As compared to the animal health ontology RR (0.333333), closer to 0 (Figure 57.b), which indicates that it only has sub-sub-class relations, it is limited. The results are following the interpretation that if the value RR is close to 1, it indicates a rich taxonomy that contains more relationships than class-subclass relations; if the value RR is closer to zero, the taxonomy has class-subclass relations only [211].

C. Class/Relation ratio (CRR)

The schema metric CRR is used to evaluate the balance between the number of classes and the number in the ATO. Equation 6 is used to calculate the CRR of ATO through the division of classes by relationships [224].

$$ClassRelationRatio = \frac{Classes}{Relationships}$$

Equation 6 Class/Relation ratio (CRR)

Where:

- (C) is the total number of classes.
- (R) is the total number of relationships, including inheritance and non-inheritance relationships.

The CRR for the ATO is calculated as 0.419355 (Figure 57.a). indicating a moderate level of descriptive richness per class. This balance ensures that classes are sufficiently described without overwhelming complexity. This interpretation of the ATO CRR results is as per definition that if CRR is equal to 1 coverage of elements with names is good, if CRR is greater than 1 it is an indication of problems that some descriptive names have collapsed, and if CRR is less than 1 there is medium coverage of the descriptive names [225]. The lower complexity of the ATO will make it easier to adopt in the real world, as compared to the animal health ontology with 2.66 CRR, which indicates problems with descriptive names.

6.2.1.4 Graph Metrics

An ontology could be envisioned as a graph structure that allows the implementation of structural analysis with the use of graph metrics [220]. Furthermore, the graph metrics are used to quantify the ontology configuration based on its hierarchical relationships, complexity, and structure. The main components used to compute the graph metrics for the ATO are depth, breadth, path ratio, structural properties, and cardinality. They were computed using the Ontometrics evaluation tool and provided the results presented in Figure 58.a.

Graph metrics more details

Absolute root cardinality:	12
Absolute leaf cardinality:	28
Absolute sibling cardinality:	39
Absolute depth:	71
Average depth:	1.820513
Maximal depth:	3
Absolute breadth:	39
Average breadth:	3.25
Maximal breadth:	12
Ratio of leaf fan-outness:	0.717949
Ratio of sibling fan-outness:	1.0
Tangledness:	0.0
Total number of paths:	39
Average number of paths:	13.0

(a) Animal Traceabilty Ontology Graph metrics results from OntoMetrics.

Graph metrics more details

Absolute root cardinality:	6
Absolute leaf cardinality:	6
Absolute sibling cardinality:	8
Absolute depth:	11
Average depth:	1.375
Maximal depth:	3
Absolute breadth:	8
Average breadth:	2.666667
Maximal breadth:	6
Ratio of leaf fan-outness:	0.75
Ratio of sibling fan-outness:	1.0
Tangledness:	0.0
Total number of paths:	8
Average number of paths:	2.666667

(b) Animal Health Ontology Graph metrics results from OntoMetrics.

Figure 58 Graph Metrics

The depth in the graph metrics results reflects the vertical complexity of the ATO, that is, how deeply the concepts are nested in the class hierarchies. The absolute depth of ATO is measured at 71 (Figure 58.a), indicating the maximum number of hierarchy levels. The average breadth of ATO is 1.820513 (Figure 58.a), computed by dividing the total sum of the hierarchical path lengths by the number of paths. Lastly, the maximal depth of 3 (Figure 58.a) is the length of the ATO's longest hierarchy path. These results indicate that the ATO has well-built vertical structures for supporting reliable hierarchy and classification reasoning, crucial for reuse and interoperability. The results follow the interpretation that ontologies that have an average depth of 0.9 are said to be structurally reliable [121].

The breadth metrics are used to evaluate the ontology horizontal spread, representing the subclasses number per class and showing how the concepts branch out at every level. The ATO absolute breadth is calculated at 36 as seen in Figure 58.a, as the total generational levels. 3.25 is the average breadth of the ATO computed by the division of the total number of child classes by the number of parent classes. The last breadth metric is the maximal breadth, representing the widest point of the ATO at 12. The total breadth of ATO (sum of absolute, generational levels, and average breadth) is 51.25. The breadth metric results of the ATO suggest a degree of horizontal complexity that is moderate. This interpretation is derived from research that indicates that high breadth values, total breadth >100 and average >20, mean the semantic relationships are rich [121]. As compared to the animal health ontology with a total of 16.666667 (Figure 58.b), the ATO performs better. The semantic richness of ATO is important for precision and expressiveness in the representation of a domain. Semantic relations also define unambiguous and clear relationships in ontologies by the use of synonyms and hyponyms to enhance interoperability.

The computed graph metric results of the ATO demonstrate an ontology design that is moderately complex and structurally reliable, in its horizontal breadth and vertical depth. The features are crucial to support knowledge sharing, facilitate integration of data, and achieve semantic interoperability across heterogeneous animal traceability systems.

6.2.1.1 Comparative analysis with reference ontology Animal Health Ontology (AHO)

The main objective of animal health ontology is to model the structure of animal populations and record the health data of animals. Selecting AHO as a reference for the evaluation of ATO was guided by using the criteria of interoperability potential, domain relevance, and functional similarity. The main aim of the ATO development is that it is interoperable with other animal traceability-based information systems. The AHO is a good reference because it complies with the semantic web standards, consistent use of properties and classes. The evaluation gives an understanding of whether the ATO achieves an approximate level of relational structure and semantic granularity, the main aspects of cross-system interoperability.

Focusing on the functional perspective, the ATO and AHO are both designed for knowledge representation of animal-level granularity (e.g., conditions and individuals), decision support of traceability systems and public health, and data exchange that is interoperable across agricultural and veterinary systems. This functional similarity justifies the ATO and AHO comparison, providing a meaningful reference to evaluate if ATO meets the semantic and structural standards that are similar.

AHO has a strong conceptual alignment with the ATO domain. The ontologies are based on various animal management aspects. The ATO is focused on animal tracking, identification, and history of movement. While the AHO is focused on animal history of disease and monitoring health. They intersect on health monitoring and animal identification. The comparison of ATO with AHO within the same conceptual domain will enhance the validity of the structural comparison. The criteria make sure that the evaluation is methodologically sound and domain appropriate, ensuring insights that are meaningful to the quality of the ATO developed. The results of the comparison of the two ontologies are presented in Table 45.

CATEGORY	ATO	AHO
Inheritance Richness (IR)	0.692308, indicating a vertically detailed knowledge structure for enhanced data integration and interoperability	0.25, reflecting a general knowledge structure suitable for broader applicability but less specialized use
Relationship Richness (RR)	0.678571, demonstrating a rich network of diverse relationships beyond simple class-subclass hierarchies	Lower RR value, suggesting limited use of non-inheritance relationships and less semantic connection diversity
Class/Relation Ratio (CRR)	Fair coverage of descriptive names, aligning well with stakeholder requirements for animal traceability	Higher CRR, indicating potential issues with the descriptive accuracy of class names, less suited for traceability
Usability Profiling	Provides precise and relevant information for animal traceability through fair coverage of descriptive names and class richness	Broader scope but less precise for traceability applications due to generality and potential descriptive issues
SPARQL Query Performance	Maintains consistency and high specificity in querying detailed information about farmers, employees, and farms	Reliable for general SPARQL queries but lacks the specificity required for detailed animal traceability tasks
Domain Focus	Tailored design ensures focused domain relevance, making it effective for precise animal traceability applications	Broader applicability across domains but not optimized for specialized animal traceability purposes
Alignment with Standards	Ensures better data integration, semantic clarity, and functionality in animal traceability systems	Offers general support for interoperability but may require customization for specialized applications

Table 45 Comparison of the ATO and AHO

The decision not to reuse or extend the AHO directly into the ATO development was based on the differences and design goals, interoperable focus, and domain scope. The ATO design objectives are different from those of AHO; the ATO model requires a modeled concept for movement events (export and transport), location hierarchies (farm and region), and tagging schemes. These concepts were not defined in the AHO, which meant significant modifications would be needed for developing ATO. Therefore, ATO was designed from scratch to specify the semantics and needs of the animal traceability domain.

Even though they share a conceptual intersection in the animal management domain. The AHO is mainly designed to model concepts related to health (disease, diagnosis, and treatment), whereas the ATO is focused on traceability (identification, ownership, and movement tracking) and regulatory compliance. The direct adoption would introduce an irrelevant concept in ATO that might dilute the focus on traceability.

The comparative analysis, therefore, revealed that the ATO, with its detailed structure, rich relationship network, and focused domain relevance, outperforms the AHO in specialized applications like animal traceability. While AHO offers a broader scope, the specific and consistent design of the ATO ensures better alignment with traceability requirements, providing more effective data integration, semantic clarity, and functionality in animal traceability systems. The comparison of the ATO with other ontologies is provided in Table 10 (Chapter 2).

6.3 Usability Profiling Evaluation

Usability profiling is the type of ontology evaluation that measures the type of context the specific ontology is used [211]. Usability profiling can be seen as the profile of the ontology based on the development and use of the ontology. Ontology annotations hold metadata and elements about the ontology [211]. Rebekka et al. [211] categorizes them into three named recognition annotations,

interface annotations, and efficiency annotations. Dividino et al. [213] concurs annotation evaluations are the best to quantify ontology quality. Moreover, the evaluation is used to emphasize significant ontology metadata using pragmatic dimensions. Pragmatic dimensions mean the usefulness of the ontology to users and developers [213]. Relevance and completeness are the main attributes that are used in usability measures. The object property metrics and class metrics from the Ontometrics tool will be used to measure usability measurements.

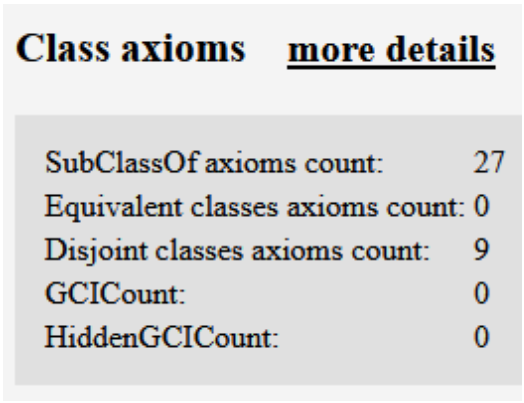
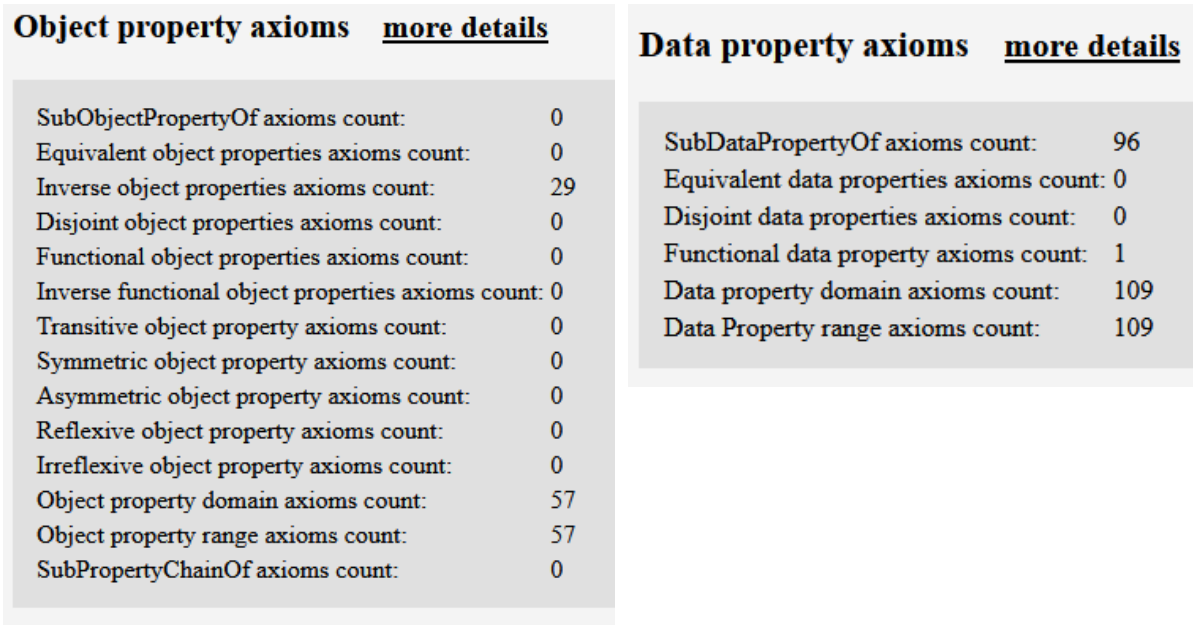


Figure 59 Class axioms Metric results from computing AnimalTraceabilityOntology.owl file

In ontologies, classes mean concepts that can contain individuals or other classes [222]. That is a class is simply a set of individuals. The Animal Traceability ontology was developed in OWL which has a thing class, and all user-defined classes are subclasses of the thing class. To showcase the quantity of classes the metric count class includes the thing class. Figure 59 presents the results of the calculated classes in the Animal Traceability ontology, resulting in a count of 27 subclass axioms, and 9 disjoint classes.

There are 2 types of properties that exist in OWL, the object property and the data property. The object property was used in the Animal Traceability ontology to link individuals to other individuals. The data property was used to link individuals to the data values. Figure 60 show the results from the ontology computation of the object and data property. Figure 60.a depicting object property domain axioms count at 57 and object property range axioms count at 57.

The result 96 subDataPropertyOfaxioms count (Figure 60.b) represents the instances of the sub-property which are subsets of the property extension of the second property. The functional data property can only have one value [222], in the results, it is 1. 109 data property domain axioms count the class expression of the domain. The data range 109 is the range of data property expression as shown in Figure 60.b.



Object property axioms more details

SubObjectPropertyOf axioms count:	0
Equivalent object properties axioms count:	0
Inverse object properties axioms count:	29
Disjoint object properties axioms count:	0
Functional object properties axioms count:	0
Inverse functional object properties axioms count:	0
Transitive object property axioms count:	0
Symmetric object property axioms count:	0
Asymmetric object property axioms count:	0
Reflexive object property axioms count:	0
Irreflexive object property axioms count:	0
Object property domain axioms count:	57
Object property range axioms count:	57
SubPropertyChainOf axioms count:	0

Data property axioms more details

SubDataPropertyOf axioms count:	96
Equivalent data properties axioms count:	0
Disjoint data properties axioms count:	0
Functional data property axioms count:	1
Data property domain axioms count:	109
Data Property range axioms count:	109

(a) Object axioms Metric results from computing AnimalTraceabilityOntology.owl file.

(b) Data axioms Metric results from computing AnimalTraceabilityOntology.owl file.

Figure 60 Property axioms

The information about animal traceability ontology presented in Figure 60 is significant to fulfilling the usability profile quality evaluation. The usability profile focuses on relevance and completeness which is illustrated in the Figures. Burton-Jones et al. [223] state that relevance is fulfilled when the ontology satisfies the stakeholder's requirements that were specific, requirements acquired before ontology development. Completeness is determined by the coverage of the domain of interest [206]. The class, object, and data values are all connected to indicate the rate of interoperability of the ontology. No specific values have been defined to indicate the perfect values that determine good interoperability. Moreover, what is important is the values corresponding to the requirements. This research believes the relevance and completeness of a good quality interoperable ontology was achieved. This is because the requirements from the stakeholders acquired during data collection and discussed in the findings (Chapter 4) correspond with the ontology classes, objects, data, and individuals created in the ontology (Chapter 5).

6.4 Functional Evaluation

The final evaluation measure is the functional measure, which performs the ontology evaluation by measuring the ontology quality based on its performance, given a task. Dividino et al. [213] state that the quality of an ontology can be seen by running a specific task, and the capability of the ontology will be observed. Moreover, the ontology degradation or improvement is observed during the task being performed. Functional evaluation is further explained as a measure that focuses on relations among the ontology graph and the meaning that is intended [211].

One of the attributes that focuses on the function is the accuracy of the ontology. The ATO will be evaluated for completeness and correct knowledge in the form of accuracy, when it is compared to animal traceability asserted knowledge. The asserted knowledge is the concepts, relationships, rules, and facts about animal traceability that have been defined in ATO. The accuracy criterion is used to check if the ATO captures true domain knowledge without any error or missed critical information. This is in line with the question posed by Amith et al. [206] that should be answered for accuracy. Is there an agreement between the ontology-asserted knowledge and knowledge from the experts? The real-world domain should be reflected in the ATO through the encoded knowledge. These functional

evaluation techniques should be accurate by specifically providing trustworthy and correct content. This follows the statement that the accuracy criterion should be based on the outcome of the properties, definitions, individuals, and class descriptions of the ontology [212]. Dividino et al. [213] is in accordance that an ontology that is good should have a close approximation with the description of its conceptualization. The functional measure for the animal traceability ontology was evaluated using Class metric, Knowledgebase metric, Apache Jena Fuseki (SQL Queries), and Pellet Reasoner from Protégé. For example, the property earTagID should represent how the cattle are tracked correctly, shows correct accuracy. However, if property isEmployed misused as a databased and not an object property, will indicate compromised accuracy. Section 6.4.6.2 Running SPARQL queries in Fuseki server is another example of expert validation for ontology accuracy, by running SPARQL queries to check if they match the competency questions set based on stakeholders.

6.4.1 Pellet Reasoner

Interoperability of the animal traceability ontology can be verified by checking for consistency. Dividino et al. [213] argue that an ontology model should have no contradictions, meaning that the deducible facts should not contradict each other. Therefore, consistency is considered an agreement between ontology entities based on the semantics of the ontology language. This project achieved consistency evaluations through the use of the Pellet reasoner plugin from protégé. Throughout the development synchronization of the project was conducted with the Pellet reasoner. The inconsistencies that the reasoner picked were fixed before the continuation of the development. The reasoner is useful in assessing the ontology consistency.

6.4.2 Knowledge Metrics

From the point of view of Lourdusamy [218] the provision of knowledge regarding the real world is an important goal of ontology. He further states that the ontology quality is described by the extent of the knowledge. Tartir [209] concurs with the importance of ontology quality, He says that it can be achieved by the measure of the placement of data in the ontology. Furthermore, an indication of the ontology design effectiveness can be indicated as the amount of knowledge of the real world represented. The knowledge metrics are used for checking the placement of data within the ontology [218]. The animal traceability ontology was evaluated using knowledge metrics to ensure that the ontology is efficient and there is an indication of the amount of real-world knowledge represented in the ontology. The knowledge metrics that will be evaluated will be the class richness and average population. Figure 61 presents the results of the knowledge metrics automatically calculated by OntoMetrics for the ATO.



Figure 61 Knowledge metrics results from OntoMetrics

6.4.3 Average Population (AP)

The metric average population is not used to evaluate the quality of the ontology but used to evaluate the structure of the ontology. Average metric assesses the degree at which the ATO is populated with individual instances relative to the ATO classes. Furthermore, indicating the level of the ATO utilization in the real-world data through the specific entities. Sicilia et al. [224] describes the AP metric of an ontology as a measurement for the average distribution across all the classes of instances. [216] further

states that the result of AP is a number that presents how well the data extraction process performed to populate KB. Furthermore, it is important to note that some schema classes may consist of a low number or a high number depending on the representation. AP for ATO is calculated as 3.025641 (Figure 61) using equation 7. This means that the ATO is moderately populated having a few instances for each class.

$$AP = \frac{|I|}{|C|}$$

Equation 7 Average Population (AP)

where,

- (I) - is the total number of instances in the ontology.
- (C) – is the total number of classes defined in the ontology.

6.4.4 Class Richness (CR) metric

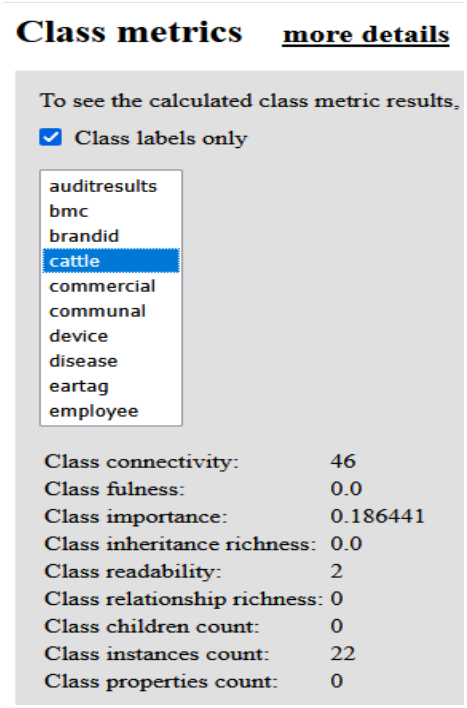
The CR metric represents the distribution of instances across classes [216], [224]. This means that a comparison of the amount of classes with instances in the knowledge base and the total number of classes is calculated. This will confirm how well the knowledge developed by the schema classes is utilized by the knowledge base. Equation 8 is used to calculate CR, where, C' – number of nonempty classes (classes with instances), C – number of total classes defined in ontology schema). The CR for the Animal Traceability Ontology is calculated as 0.487179 (Figure 61). The percentage in Figure 61 is low indicating that the ratio of the number of classes has no data to exemplify all the knowledge in the schema. The functionality of the ontology quality evaluated is fair, not the best that is should be but can be improved. The current CR value combined with other good results achieved will still ensure that the animal traceability ontology is interoperable.

$$CR = \frac{|C'|}{|C|}$$

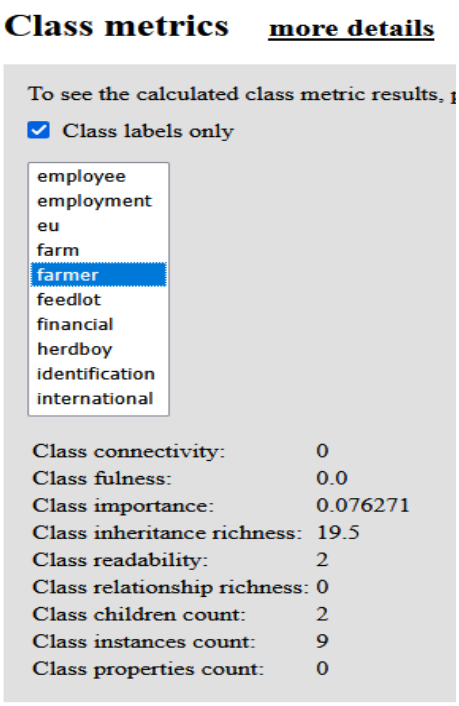
Equation 8 Class Richness (CR) metric

6.4.5 Class Metric

The class metric is used for the exploration of the ATO classes and relationships. The metric mainly represents the distribution of the ontology instances over the ontology classes. According to Tartir et al. [216] it measures the instances percentage connected to classes belonging to the subtree rooted where the current class is to the instance total number. Tartir et al. [216] assert that it is important for assisting to identifying the areas in the schema that the focus during the extraction of instances. Furthermore, informing the user/extractor how suitable the ontology is for their intended use. It may not be exact but a clear idea of parts of the ontology that are said to be focal and the parts on the edge can be observed [216].



(a) Class Cattle metric results for class importance and class instance count.



(b) Class metric results for class inheritance richness.

Figure 62 Class Metric

Figure 62 presents the results of the class metric after computation of the ATO on the OntoMetric tool. The main features of the classes were calculated, and the results given, not all the features of the class metric were calculated. Class connectivity is used to indicate the classes that are central in the Animal Traceability ontology in line with the instance relationship graph. That is, where edges define the relationship among them and nodes define instances [209]. Figure 62.a shows the class connectivity of class cattle being 46, this measure is used to help in understanding the ontology nature by depicting the classes that play a significant role as compared to other classes. In Figure 62.a, the class cattle represents the class importance metric at 0.186441. This percentage indicates the instances belonging to classes at the rooted current class of the inheritance subtree to the instance total number [209]. Class importance metric is important for the identification of areas in the schema that are focal at the time instances will be added to the knowledgebase.

Gobin-Rahimbux [222] state that class inheritance richness is important for indicating the efficient grouping of knowledge into differing categories and also subcategories for that particular class. Figure 62.b indicates the class inheritance richness from the class farmer at 19.5 as the average number of classes per schema level. Tartir et al. [216] say the interpretation of the results depends on the ontology nature either having a specific domain or a wide domain. The Animal Traceability ontology represents a specific domain thus the results being a low 19.5, for wide domain the inheritance richness is usually higher. The last measure in the class metric was the class instances indicated in Figure 62.a by class cattle resulting in 22 instances. OWL groups resources that have similar characteristics, each class has an association with a class extension (a set of individuals). Those individuals contained in the class extension are the class instances.

6.4.6 Evaluation using Apache Jena Fuseki

This research used Apache Jena Fuseki as an evaluation technique for the functional evaluation of the Animal Traceability ontology. As stated by Swenja et al. [225] , Apache Jena Fuseki serve is widely used and among the most famous RDF frameworks. Apache Jena Fuseki server is defined as an HTTP-

based query engine that is used for the execution of SPARQL queries [204]. Furthermore, executing the queries is run on the web of linked data, also extracting triple information for the querying on the Apache server web page. According to Chatterjee [226] Apache includes storage of triples in memory, a SPARQL API query engine used to update and query RDF models, command line tools [227], an ontology API to handle RDF and OWL ontologies, and GUI-based Apache Jena Fuseki SPARQL serve to support queries [228]. Apache Jena is a website that is open source found at http://jena.apache.org [227]. Apache Jena Fuseki 4.9.0 was chosen for this research.. Apache is used to evaluate the performance and accuracy of the ontology, as a pre-requisite to determining the quality of the ontology. The task of running SPARQL queries verified that the ontology can perform tasks according to requirements in a short period of time. Therefore, indicating that it is interoperable as required.

6.4.6.1 Triples created in Knowledge Base

After an ontology has been developed it should be made accessible to end users by publishing it on the web for further evaluations and reusability [229]. The ATO developed for this research has been published on the web on Figshare, and users are able to freely access and download the file. To evaluate the Animal Traceability ontology the file on the local machine was uploaded to the Fuseki server. To get started with the Fuseki server, after the application is downloaded the Windows command line is used to begin running the Fuseki server on Apache Tomcat. When the server was successfully started, the Fuseki server was accessed through the browser at http://localhost:3030. The ontology file AnimalTraceability.rdf was uploaded from the local drive. The persistent dataset was created next called Animal Traceability and it is used as the triple store database or knowledge base. The next step was to upload the AnimalTraceability.rdf file to the triple store database, apache showed that it contained 906 triples. SPARQL endpoint created from the triple store database output is significant because it allows users to use SPARQL language to query the knowledge base.

6.4.6.2 Running SPARQL queries in Fuseki server

The generated SPARQL endpoint is used on the Animal Traceability triple store database for accessing and querying the ontology. The defined set of competency questions in Chapter 5 (Section 5.3.2.4 Specification) of this study were used to run the SPARQL queries for querying the ontology. A uniform resource identifier is added to the triples before running the queries to make the farmer and resource unique from others. The queries in Figure 63, Figure 65, and Figure 67 were run for accessing the AnimalTraceability.rdf data. The query in Figure 63 is executed on the Animal Traceability system when a stakeholder (MOA, Veterinary, BMC) wants to audit the farmer's employees. This query searched the triple store for the “employee of the first farmer?” and that was “herdboy1”, The results of the query are seen in Figure 64. The second query was run to search for the “name of the farm the farmer owns?” (Figure 65), output as seen in Figure 66 was “Sandveld”. The last query was searching for the “location the farm belongs?” (Figure 67) to and the output as shown in Figure 68 was “Central”. The results currently in Fuseki for the SPARQL query are in four formats namely: CST, JSON, TSV, and XML. The output results from Figure 64, Figure 66, and Figure 68 are displayed in JSON format. The future direction is to present the results using a graphical user interface where stakeholders type farmer/cattle name/ID in the search option and results are presented in an easy-to-read and understandable form like a mobile application or software application.

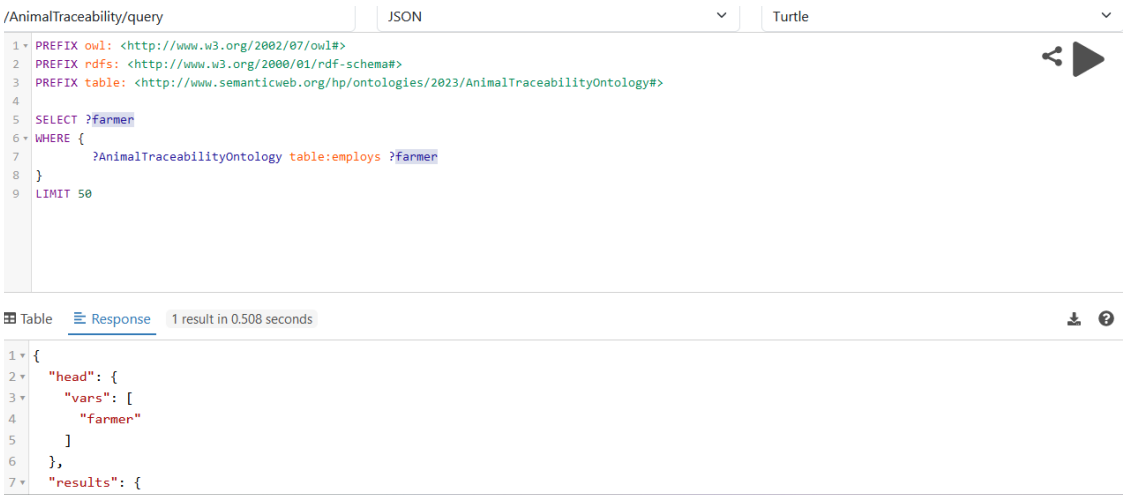


Figure 63 Running SPARQL query to search for farmer employee.

Uniform resource identifier <http://www.semanticweb.org/hp/ontologies/2023/AnimalTraceabilityOntology#> is shown in Figure 63, it makes the farmer resource unique.

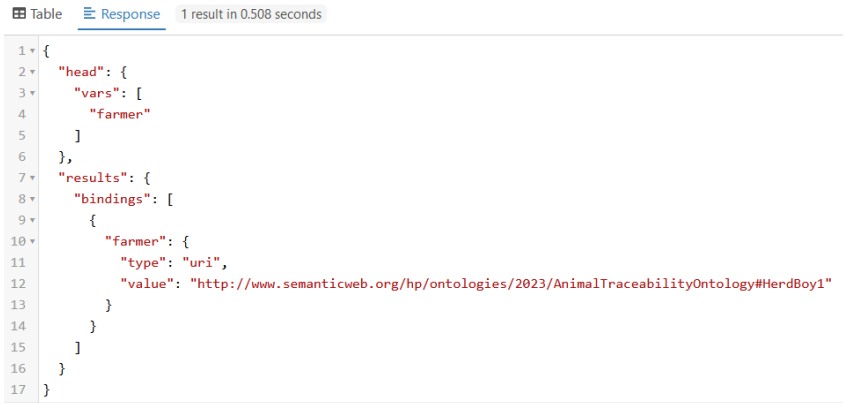


Figure 64 Output from SPARQL query to searching for farmer employee.

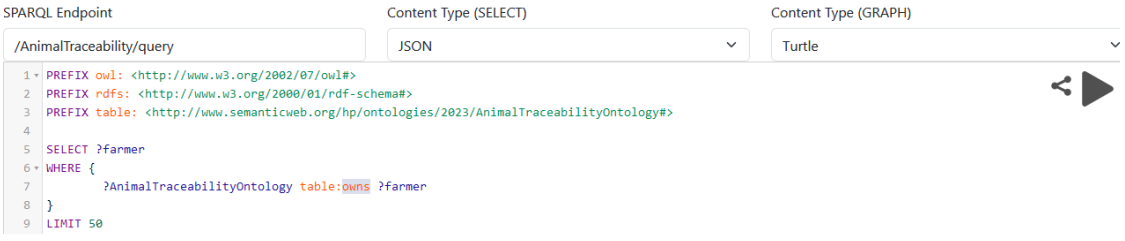


Figure 65 Running SPARQL query to search for name of farm the farmer owns.

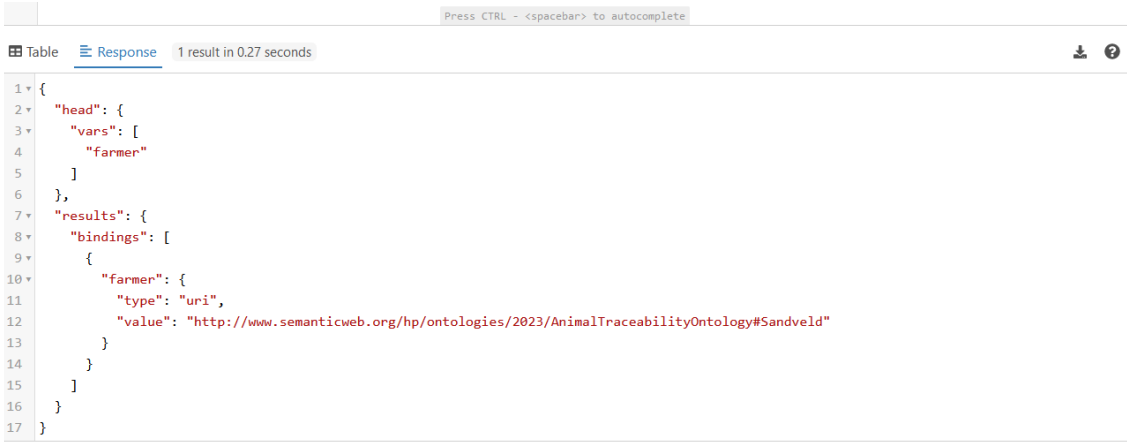


Figure 66 Output from SPARQL query to searching for farm name.

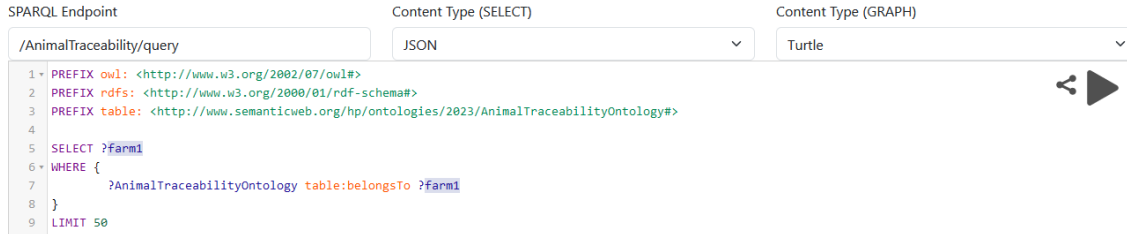


Figure 67 Running SPARQL query to search for the farm location.

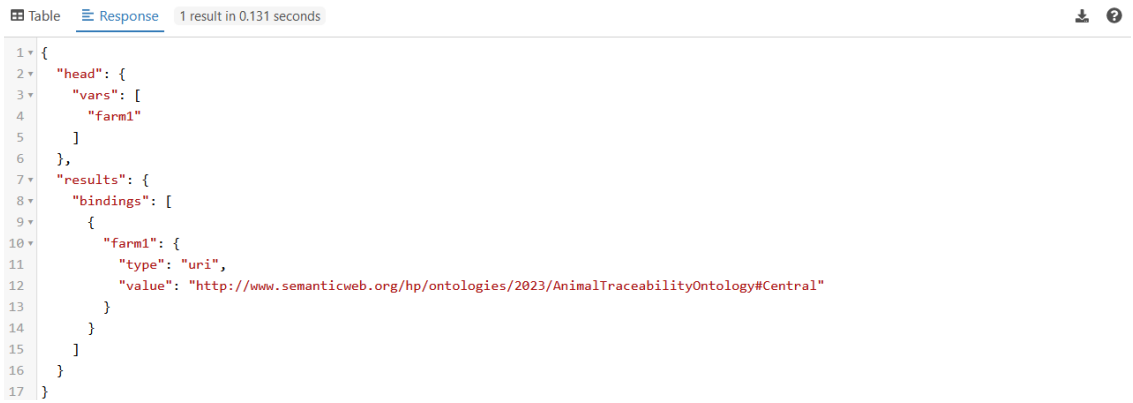


Figure 68 Output from SPARQL query to search for farm location

6.5 OntOlogy Pitfall Scanner! (OOPS!) Evaluation

OOPS! is an ontology evaluation tool used to detect ontology pitfalls and was used to detect the pitfalls in the ATO. The tool provides a list of pitfalls to detect through available and recent approaches. The list of pitfalls and what they represent can be found at https://oops.linkeddata.es/catalogue.jsp . An indicator for each pitfall is also extended based on the pitfall possible negative impact, Figure 69 shows the OOPS! Indicator impacts are either critical, important, or minor.

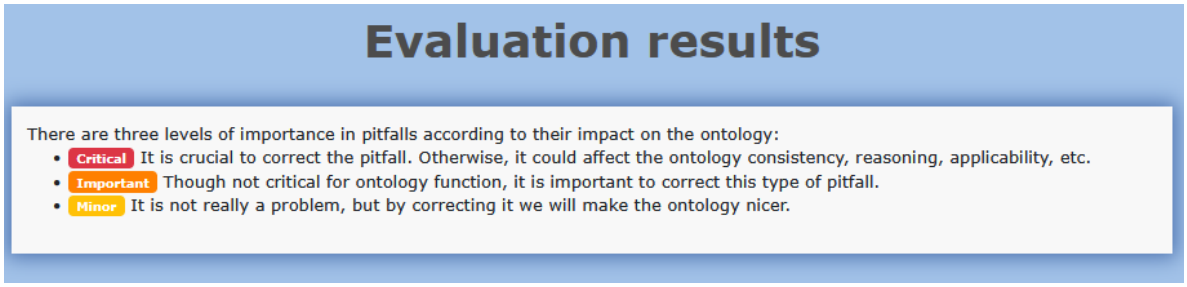


Figure 69 OOPS! Pitfall indicator

The ATO was inserted in the pitfall scanner and the results are presented in Figure 70. Each pitfall detected was addressed in the ATO in protégé. ATO was scanned for all pitfalls and specifically for consistency, completeness, and conscieseness.

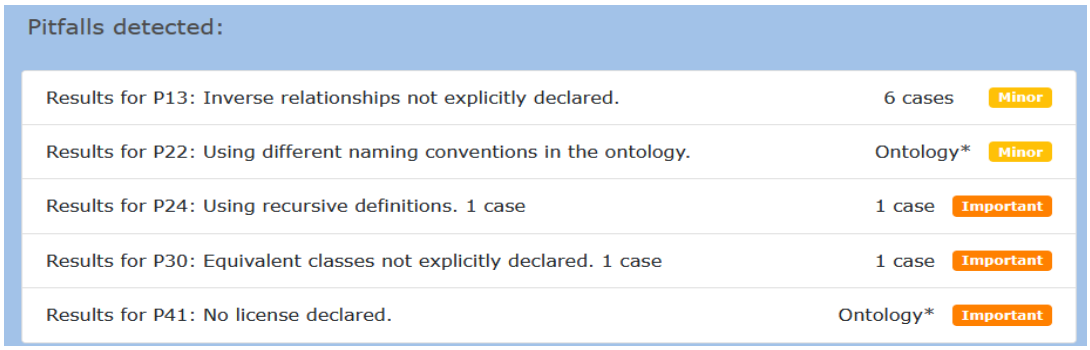


Figure 70 ATO results from OOPS! pitfall scanner

The first pitfall detected is “P13: Inverse relationships not explicitly declared”, and there are six cases in the ATO. Figure 71.a shows the exact affected properties, which are object properties in the ATO. The pitfall means that the object property inverse relationships are not defined in the ATO. The object property relationships were correctly defined in the ontology under “Inverse Of” in protégé. Figure 71.b presents one of the properties “given” being defined in protégé with its Inverse Of relation “own”.

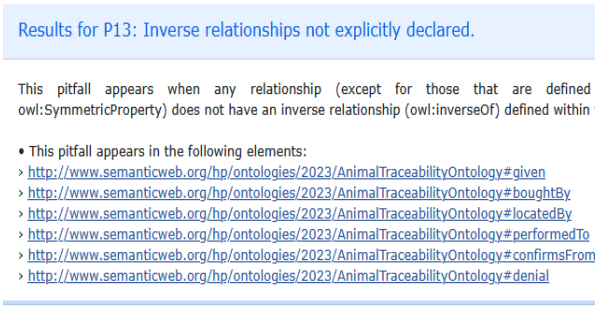


Figure 71.a OOPS! Results P13

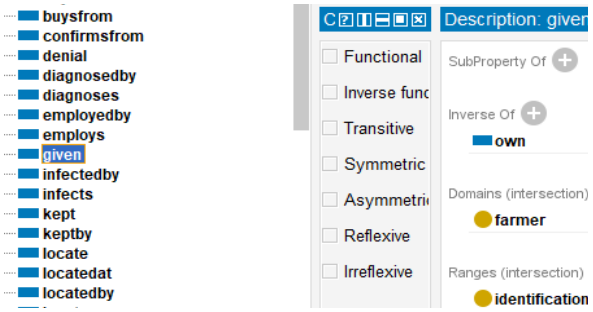


Figure 71.b Pitfall fixed in Protégé (object property “given”)

Figure 71 ATO pitfall detected P13: Inverse relationships not explicitly declared

The second pitfall detected was “P22: Using different naming conventions” in the ontology as shown in Figure 72.a. This means that the class, data properties, object properties, and individual instances do not have the same naming conventions. For consistency, all the names defined in the ontology should use the same naming convention. Figure 72.b presents the data properties “transferLocation” that was mistakenly defined using lowercase and uppercase conventions; this was corrected by going to refracting in protégé and selecting to rename IRI from “transferLocation” to “transferlocation”. The ATO used a lowercase naming convention.

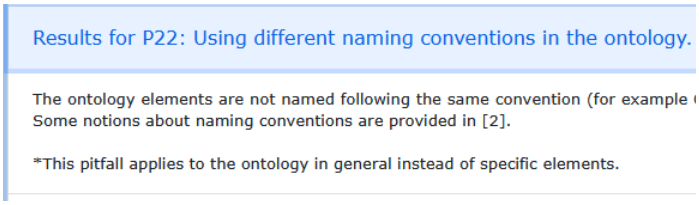


Figure 72.a OOPS! Results P22

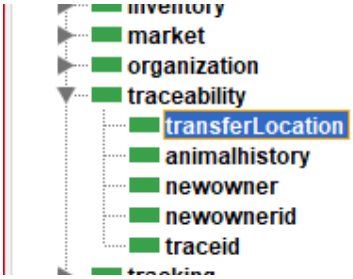


Figure 72.b Pitfall fixed in Protégé (data property transferLocation")

Figure 72 ATO pitfall detected P22: Using different naming conventions

The next pitfall, as shown in Figure 70, was "P24: Using recursive definition", and it was one case. This meant that in ATO, a class or property is used in its own definition. It was identified as the class "cattle" defined in data properties as seen in Figure 73; the pitfall was corrected by removing the domain "cattle" in the data property definition.

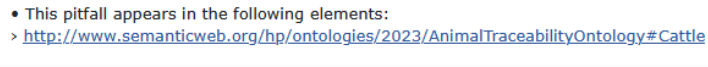


Figure 73 ATO pitfall detected P24: Using recursive definition

OOPS! Detected "P30 Equivalent classes not explicitly declared", meaning that one of the classes was missing the definition of equivalent classes as shown in Figure 70. The pitfall detected class "system" as being equivalent to "organization" presented in Figure 74. In the ATO definition, these classes are not equivalent; the pitfall was fixed by renaming the "system" class to "tracesystem". This pitfall can happen when reusing terms from other ontologies that have the same meaning. The term "system" has been used in other ontologies and may have the same definition as organization, causing this pitfall. However, in the ATO, they do not have the same meaning.

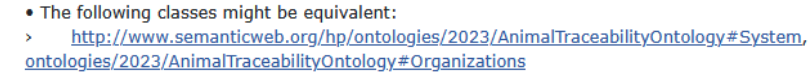


Figure 74 ATO pitfall detected P30 Equivalent classes not explicitly declared

The last pitfall detected by OOPS! Was "P41: No license declared" as shown in Figure 75.a. This means that the ATO was missing the license definition. The pitfall was fixed by adding the license definition to the ATO ontology as presented in Figure 75.b.

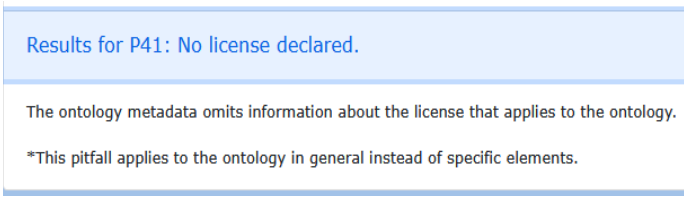


Figure 75.a OOPS! Results P41

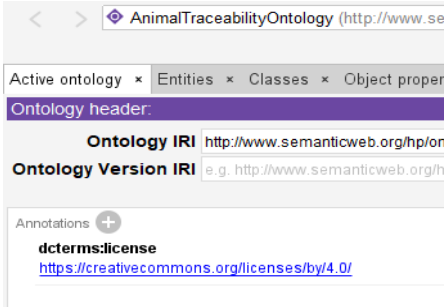


Figure 75.b Pitfall fixed in Protégé

Figure 75 ATO pitfall detected P41: No license declared

All the pitfalls were fixed and the ATO was re-evaluated in OOPS! The results are presented in Figure 76. Showing that the ATO had no pitfalls. This means that ATO has met the criteria of being consistent, complete, and conscience.

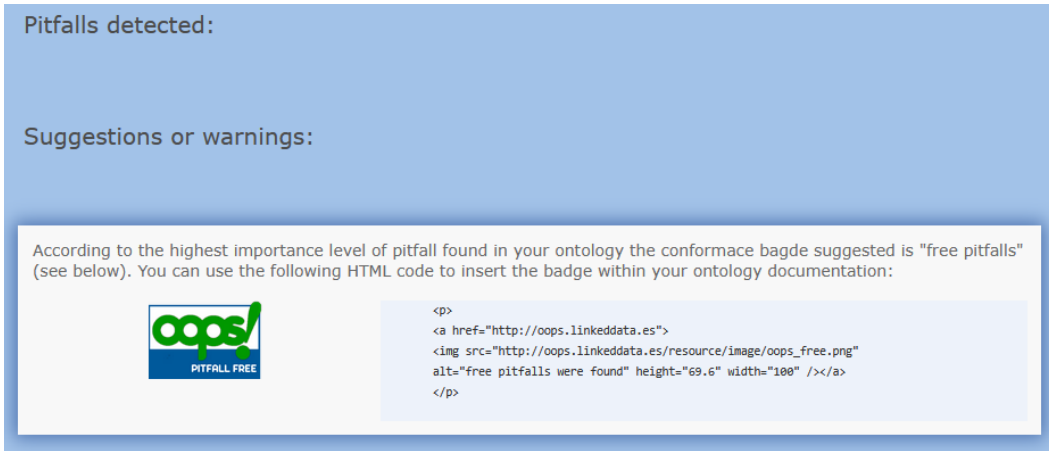


Figure 76 ATO results presenting no pitfalls detected in the ontology

6.6 Chapter Conclusion

The main objective of this research was met in Chapter 5 with the developed interoperable knowledge-driven animal traceability framework. The framework was further developed into an ontology to further demonstrate and guide in achieving interoperability and standardization of protocols. This chapter focused on evaluating the animal traceability ontology and producing a report of the evaluation. The animal traceability ontology was evaluated for quality and correctness. The types of ontology evaluation measures used were structural, functional, and usability profiling. Within these evaluation methods, two techniques were used; the first technique used is computing metrics using the OntoMetric tool, and the second is running SPARQL queries in Apache Jena Fuseki. Each step of the evaluation process was carried out and results were produced, as seen in the figures displayed. The figures are screenshots depicting multiple stages of the evaluation process. Starting with the environment of the evaluation tool, the AnimalTraceability.rdf file uploaded, the evaluation process, and the results from the computation of the file. Animal health ontology was used in comparison with the animal traceability ontology in the structural evaluation process.

Using Apache Jena Fuseki for evaluating the AnimalTraceability.rdf ontology was important for multiple reasons. SPARQL endpoint provided by Fuseki allowed queries to be executed on the Animal Traceability ontology, resulting in the evaluation of the ontology relationships, structure, and data consistency. The ontology data was checked for inconsistencies and the data was validated against the ontology axioms. It is mostly important in providing deployment and integration of ontology into systems and applications, which is significant for usage in the real world. Users can keep track of changes in the data and ontology over time using the version and historical analysis feature provided by Fuseki. Running SPARQL queries in Fuseki helped verify the ontology performance, correctness, and consistency ensuring that the expected requirements are met to be used in applications effectively.

OntoMetric calculated various aspects of the Animal Traceability ontology to determine the effectiveness and quality. The evaluation process involved analyzing and calculating multiple dimensions of the ontology through base metrics, schema metrics, knowledge metrics, class metrics, and graph metrics features from the OntoMetric web-based tool. Results from calculations are presented in statistical format giving a percentage of the ontology dimension. OntoMetric evaluation process was to find out if the ontology meets these requirements: 1) The organization of the structure based on its relationships, classes, and properties. There is clarity in the structure and it adheres to best modelling practices by using standard ontology constructs. 2) There is logical consistency and coherence, looking

at the soundness of the content and structure of the ontology. No inconsistencies and contradictions. 3) How well the domain is represented for the model intended. Evaluating if all relevant constraints, concepts, and relationships are captured. 4) Performance of query execution times, resource usage, and large dataset handling to help assess the scalability of the ontology. 5) The usability and reusability of the ontology focus on modularity, system integration, and documentation quality. The ease of ontology use for its purpose and reuse for different contexts. 6) The ease of ontology interoperability with other systems and ontologies and integration. Considering external vocabulary alignments, standard data format compatibility, and semantic interoperability support. 7) The ability of the ontology to evolve, long-term sustainability, and management of changes.

Overall, the Proposed Ontology-Driven Animal Traceability Framework supported by the Animal Traceability ontology artefact evaluation results shows that it meets the ontology quality and correctness requirements to be used in systems and applications for animal traceability. Based on ontology alignment of comparing an existing ontology (Animal health ontology) and a developed ontology, results show that the animal traceability ontology developed by this research is more interoperable than the animal health ontology. The evaluation metrics present animal health ontology with 33.33% and animal traceability with 83.33% from the 6 metric measures calculated against each other. Interoperability of the animal traceability systems will be seamless for the stakeholders, as shown by the inheritance richness 0.692308 results indicating very detailed knowledge. The relationship richness of the ontology at 0.678571 means that the ontology contains strong connections between the classes. Fair results were presented by results from class richness with average results of 0.487179. Animal traceability ontology will achieve flawless data integration and exchange based on the reliable results of 0.9 average depth from graph metrics. Usability measures relevance and completeness of the ontology, were achieved because the requirements essential to stakeholders and gap in literature correspond with the developed features of the ontology. Correctness and quality were further measured based on functionality and performance using SPARQL queries, pellet reasoner, class metric, and knowledge metric. The tasks given as seen in the results were successful providing a logically consistent ontology. Overall, the evaluation results present an animal traceability ontology that meets the quality and correctness ontology requirements.

7 CONCLUSION, LIMITATIONS AND RECOMMENDATION

7.1 Introduction

In this final Chapter 7 the conclusions derived from the findings of this study on the research problem: "The problem this research addresses is the lack of interoperable knowledge-based frameworks for the design and development of animal traceability systems and platforms" is described. The conclusions stem from the research's main objective, research questions, synopsis of the literature, results of the study, and the proposed framework. A summary of the conclusion including a brief discussion of each objective is presented along with the outcomes of the study. The limitations of the study are explained stating the weaknesses. The chapter concludes with a list of recommendations that could be addressed in the future, by the study researcher or other researchers.

7.2 Summary

This research focused on developing an ontology-driven framework for animal traceability. The research is premised on the animal traceability scenario in Botswana. It was discovered in the study that a large portion of the resources and budget in Botswana is allocated to the safety and preservation of animals. Yet, the main stakeholders farmers, veterinary services, Ministry of Agriculture, Matimela officers BMC, and police officers are still facing challenges in managing cattle, animal movements, feeding, diseases, stray animals, meeting livestock regulations, and vaccinations. There is no existence of interoperability, knowledge management, and standardization of animal traceability in Botswana. Implementation of interoperable traceability systems with standardized protocols will improve animal traceability by reducing animal diseases, increasing exports from better management of cattle, reducing animal theft, a solution to cross-border theft, and increasing the economy – economy that is not only agricultural-driven but also knowledge-driven.

Data collection from the districts of Botswana was carried out through interviews and questionnaires along with the literature reviewed, to answer the research questions posed in this study. The first objective is:

1. To identify limitations of the current Animal Traceability frameworks and systems.

This objective was addressed by conducting data collection, interviewing, and giving questionnaires to the main stakeholders of animal traceability in Botswana and carrying out a detailed extensive literature review. The study found that most of the focus in animal traceability has been on devices to monitor the animals and less on the management of the animal data. Zhao et al. [1] talks of another system that only focused on the device used for animal traceability with an ineffective information management system that relies on how the animal will be slaughtered to get information. Only a few traceability frameworks were found in research, one being the general framework for animal food safety traceability using GSI and RFID [8]. The focus of the framework is tracing the health of cattle that are going to be sold. The system in Botswana started as LITS and then later upgraded to BAITS. However, the technical challenges were still not addressed, and managing the animal data was not effective. There is still no interoperability and integration among the stakeholders as well as the lack of semantic knowledge management.

The second objective addressed:

2. To generate components for the design and development of the proposed interoperable framework for animal traceability.

This research chose Botswana as the scenario for animal traceability, thus data collection from the stakeholders was conducted about the challenges of animal traceability they are facing. Literature carried out about animal traceability systems, frameworks, and their challenges were reviewed. The collected data was analyzed using statistical analysis to get quantitative results and Nvivo software for

content and thematic analysis for qualitative results. The findings and reviewed literature were used together to get results that were used to generate the necessary components for the design and development of the framework.

From the generated components the third objective:

3. To use the generated components in the design and development of the proposed interoperable framework for animal traceability.

The third objective is answered. Chapter 5 of the study presents Figure 40 which illustrates the proposed Ontology-driven framework for animal traceability to be used as a guiding principle for the development and management of animal traceability platforms. The framework can also be adopted to develop systems in multiple fields like healthcare, cyber-crime, and education.
This research aims to infuse ontologies to solve and improve the animal traceability scenario in the world. The beginning of Chapter 5 presents the proposed framework that serves as a guiding principle for developers and researchers. To further simplify and demonstrate the benefits of using ontologies this research developed an animal traceability ontology.

Building the animal traceability ontology addresses the objective:

4. To identify existing ontologies that can be re-used, alternatively – in the absence of existing ontologies identify ontology methodology to develop a new ontology for Animal Traceability.

The research did not find an ontology to reuse, an animal traceability ontology was developed from scratch. [96] state that currently ontology technologies are mature enough to develop ontologies from scratch using the available methodologies, languages, and tools. Animal traceability ontology was built using the methodology called METHONTOLOGY in the Protégé environment using OWL and RDF language. Building the ontology for Animal Traceability determining concepts, object properties, data properties, and individuals will provide Animal Traceability stakeholders to efficiently infer and extract knowledge that is valuable from the data. The ontology is a freely available open-source resource that can be downloaded and used for further research. Chapter 5 illustrates the development of interoperable knowledge-driven systems with standardized protocols for easy communication among the stakeholders.

To conclude, the study answered the objective:

5. To evaluate the Animal Traceability ontology.

Chapter 6 addresses the last objective using metrics and SPARQL queries to evaluate the built animal traceability ontology. The evaluation is carried out using 2 ontologies the animal health ontology called reference ontology and the animal traceability ontology called the candidate ontology. The quality of the ontology was evaluated in three ways structural evaluation, functional evaluation, and usability-profile evaluation. The technique of computing metrics using a tool called OntoMetric which calculates the quality of the ontology based on schema, base, graph, class, and knowledge metrics was implemented. The results of the knowledge metric specifically class richness resulted in a percentage of 0.487179 indicating that the ratio of the number of classes has no data to exemplify all the knowledge in the schema. The percentage is determined by how big the ontology is and if it is domain-specific. The study domain was made low since it was a proof of concept, and it is domain-specific (animal traceability), thus the low results. The other evaluation technique used was running SPARQL queries in Apache Jena Fuseki. The figures in Chapter 6 display each step of the evaluation process from uploading the AnimalTraceability.rdf file to running queries and getting JSON results. SPARQL queries illustrated the ability of the ontology to carry out the requirements that were set at the conceptual development stage of the ontology.

7.3 Limitations

This study has several limitations. The first limitation was identified during the interview process with stakeholders. Setting for interviews had to happen in a sequence at certain dates and times to get as much data from relevant sources as possible given the limited budget and time. Some of the volunteers were missed because they were not available at the time the research team was in their area. Getting assistance from some of the veterinary offices for contacts to commercial farmers was a challenge and the research team failed in other districts to interview commercial farmers. Getting access to communal farmers was easier but commercial farmers lock their gates and you cannot randomly show up at their farm to interview them.

1. Resource and logistical constraints in data collection.

The data in this study only included 66 participants from the nine districts of Botswana. The majority of the participants were randomly sampled while moving around the farm areas. This ensured unbiased information was collected from the participants. However, due to resource and logistical constraints, the participation engagement of the domain experts (veterinarians, MOA, farmers, BMC, Matimela officers) was limited. The ATO may not fully reflect the reality of the operations of all the stakeholders in the animal traceability value chain. The ATO also depends on the availability of data that is structured and standardized data from the various stakeholders. Even so, the ecosystem of data in the different districts is poorly digitized and fragmented. The lack of data entries that are inconsistent, standardized data schemes, and limited information sharing across the organizations boundaries will notably affect the ontology validation and population. Having considered the definition of tacit knowledge, it was deemed out of scope of the current and can be considered as future work to extend work done in this thesis.

2. Technical challenges that might cause limited system adoption.

The ontology-driven systems implemented will require maintenance with specialised knowledge in semantic technologies, and will not be available easily in an agricultural environment. It will require technical skills in using ontology management tools, such as OWL API and Protégé. A potential learning challenge for non-technical stakeholders. The technical challenge may limit the adoption of the ontology. Stakeholders with low resources may experience socio-economic and technical limitations that will be a barrier to the adoption of the ontology-driven animal traceability system. The necessary semantic technologies, like SPARQL and OWL, are often not readily accessible to users with no technical training. Furthermore, in developing countries like Botswana and Kenya, there is a digital divide that is a concern about equity of ontology-driven animal traceability systems that will inadvertently marginalize the same stakeholders the systems aspire to equip.

3. Semantic limitations due to data collection challenges

The evaluation of the ontologies presented a limitation in getting good semantic results because the research developed a medium-sized ontology. Numbers higher than 100 after computing total values for the breadth, and an average of 20 is said to indicate ontologies with good semantic relations [102]. Animal traceability ontology shows a total breadth of 54 and an average breadth of 3.25, indicating fair semantic relations that are not too good. Limited time and resources restricted the research to developing an ontology big enough to get good semantic relations.

4. Potential epistemic misalignment in diverse cultures

The development and testing of the ATO were primarily conducted within the context of Botswana's livestock industry, potentially limiting its applicability in other geographical or cultural settings without appropriate customization. There is a potential for an epistemic misalignment when the ATO is implemented in diverse cultural settings. Data-driven systems sometimes intersect with issues that are sensitive, like national sovereignty, data privacy, legal jurisdictions, and farmer autonomy over agricultural data. The framework is sound technically, but the ethical aspect was not directly addressed.

When applying the systems across borders, there will be a need to incorporate cross-border exchange of data, ownership of data, and cattle movement to fit the system's governance.

7.4 Recommendations for further research

Based on the identified limitations in this study, there are several recommendations for future research to improve the applicability and robustness of the ontology-driven animal traceability framework. It is important, while enhancing the ontology-driven animal traceability framework for future research, to start by introducing Open data to Animal traceability using FAIR Data Principles for transparency and trust among stakeholders.

1. Application of computer literacy strategies

Developing strategies to accelerate computer literacy among the animal traceability community, especially the elderly. The ontology-driven animal traceability systems should have user-friendly tools and interfaces as a design priority for farmers and officers who are non-technical. It will encourage the adoption and use of the systems, further enhancing effective farming practices. For example, the systems should not be in English only, but also include the local languages. Using pictures, simpler interfaces, and bigger fonts will help with accessibility. Regular training programs for stakeholders will help in overcoming the technical barriers. Limited resources may be a hindrance in regular training programs for stakeholders with limited funds to hire technical experts. Inclusion of tacit knowledge to get advance the data collected.

2. Collaboration with international organizations and stakeholders

Collaboration with international organizations and stakeholders could further support the standardization and global adoption of the ATO, fostering more robust and effective traceability systems. Additionally, leveraging Botswana's leadership in livestock traceability offers an opportunity to drive regional innovation, making the ATO a scalable and adaptable model for broader applications across Southern Africa and beyond. Research in the future should also put effort into extending the ontology by capturing more district specific animal practices around the world, including the various ethical and legal requirements in each country. This will require a wider array of collaborative input from domain experts to refine the ontology concepts and relationships. The more the stakeholders are involved it will enhance the ontology's usability, acceptance, and accuracy of the domain. Including the legal and ethical requirements will support the policies in place for better decision-making. The main challenges could be having enough resources and time for the research to expand wider. Research may have to find multiple research grants to carry out the research across multiple countries.

3. Integration of emerging technologies

Technological integration into the agricultural programs' strategy is crucial. With emerging technologies being released continuously, the ontology interoperability can be enhanced in the future by exploring an integration with advanced ontology alignment tools. Tools to automate mapping of data standards that are external with less intervention from stakeholders, for example, machine learning could be leveraged to align diverse ontologies automatically. IOT sensors to automatically capture data with the use of GPS and RFID. Incorporating real-time data processing capabilities would improve the ontology's responsiveness to urgent health and safety concerns, through integrating advanced technologies such as blockchain and IoT could bolster data security, traceability, and transparency. Animal movement tracking based on their location using geospatial systems. Moreover, implementing mobile platforms for small for smallholder farmers to access and enter their data at an affordable rate. Integrating with these technologies will reduce the manual input errors and improve the data authenticity. For developing countries, this could be a challenge because the internet is necessary to carry out these functions. However, most of the farms do not have internet, and some have no network at all. Until the cellphone network and internet reach these areas, it will slow down the progress.

8 REFERENCES

[1] J. Zhao, A. Li, X. Jin, and L. Pan, "Technologies in individual animal identification and meat products traceability," *Biotechnol. Biotechnol. Equip.*, vol. 34, no. 1, pp. 48–57, Jan. 2020, doi: 10.1080/13102818.2019.1711185.

[2] C. Dalvit, M. De Marchi, and M. Cassandro, "Genetic traceability of livestock products: A review," *Meat Sci.*, vol. 77, no. 4, pp. 437–449, Dec. 2007, doi: 10.1016/j.meatsci.2007.05.027.

[3] V. Caporale, A. Giovannini, C. Di Francesco, and P. Calistri, "Importance of the traceability of animals and animal products in epidemiology: -EN- -FR- -ES-," *Rev. Sci. Tech. OIE*, vol. 20, no. 2, pp. 372–378, Aug. 2001, doi: 10.20506/rst.20.2.1279.

[4] M. B. Bowling *et al.*, "REVIEW: Identification and Traceability of Cattle in Selected Countries Outside of North America," *Prof. Anim. Sci.*, vol. 24, no. 4, pp. 287–294, Aug. 2008, doi: 10.15232/S1080-7446(15)30858-5.

[5] H. Bai *et al.*, "Traceability technologies for farm animals and their products in China," *Food Control*, vol. 79, pp. 35–43, Sep. 2017, doi: 10.1016/j.foodcont.2017.02.040.

[6] B. Vallat, "Animal Identification and Traceability (From farm to fork),"," in *World Organization for Animal Health*, Buenos Aires, Argentina, Mar. 2019, pp. 1–35. [Online]. Available: https://www.woah.org/fileadmin/Home/eng/Conferences_Events/sites/traceability-2009/documents.html

[7] G. Lavoie and J.-F. Forest, "Implementation of a Traceability System From Constraints to Opportunities for the Industry: A Case Study of Quebec, Canada," 2009, doi: 10.22004/AG.ECON.53753.

[8] W. Cao, L. Zheng, H. Zhu, and P. Wu, "General Framework for Animal Food Safety Traceability Using GS1 and RFID," in *Computer and Computing Technologies in Agriculture III*, vol. 317, D. Li and C. Zhao, Eds., in IFIP Advances in Information and Communication Technology, vol. 317. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 297–304. doi: 10.1007/978-3-642-12220-0_43.

[9] C. J. C. Phillips, "Zoonotic Disease Risks of Live Export of Cattle and Sheep, with a Focus on Australian Shipments to Asia and the Middle East," *Animals*, vol. 12, no. 23, p. 3425, Dec. 2022, doi: 10.3390/ani12233425.

[10] S. Bahta *et al.*, "Productive efficiency of beef cattle production in Botswana: a latent class stochastic meta-frontier analysis," *Front. Sustain. Food Syst.*, vol. 7, p. 1098642, Jun. 2023, doi: 10.3389/fsufs.2023.1098642.

[11] G. Mwanga, E. Mbega, Z. Yonah, and M. G. G. Chagunda, "How Information Communication Technology Can Enhance Evidence-Based Decisions and Farm-to-Fork Animal Traceability for Livestock Farmers," *Sci. World J.*, vol. 2020, pp. 1–12, Dec. 2020, doi: 10.1155/2020/1279569.

[12] Department of Agriculture, Fisheries, and Forestry, "Sanitary and phytosanitary measures," Australian Government, Department of Agriculture, Fisheries, and Forestry, Genève, Switzerland, 2012. [Online]. Available: https://www.agriculture.gov.au/sites/default/files/sitecollectiondocuments/animal-plant/plant-health/publications/taxanomy/wto_sps_agreement_booklet.pdf

[13] The European Commission, "Commission Delegated Regulation(EU) 2019/2035," Official Journal of the European Union L314/115, Jun. 2019. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019R2035

[14] B. Nkatekho, "Animal Welfare Policies and Their Effect on Livestock Productivity and Trade," *Int. J. Livest. Policy*, vol. 3, no. 2, pp. 27–40, Jun. 2024, doi: 10.47941/ijlp.1967.

[15] F. Mutua, J. Lindahl, and D. Randolph, "Possibilities of establishing a smallholder pig identification and traceability system in Kenya," *Trop. Anim. Health Prod.*, vol. 52, no. 2, pp. 859–870, Mar. 2020, doi: 10.1007/s11250-019-02077-9.

[16] F. Mutua *et al.*, "Piloting a livestock identification and traceability system in the northern Tanzania–Narok–Nairobi trade route," *Trop. Anim. Health Prod.*, vol. 50, no. 2, pp. 299–308, Feb. 2018, doi: 10.1007/s11250-017-1431-4.

[17] Y. Resti *et al.*, "A review of on-farm recording tools for smallholder dairy farming in developing countries," *Trop. Anim. Health Prod.*, vol. 56, no. 5, p. 168, Jun. 2024, doi: 10.1007/s11250-024-04024-9.

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

[19] L. Magliulo, L. Genovese, V. Peretti, and N. Murru, "Application of ontologies to traceability in the dairy supply chain," *Agric. Sci.*, vol. 04, no. 05, pp. 41–45, 2013, doi: 10.4236/as.2013.45B008.

[20] N. Power, J. Alcock, R. Philpot, and M. Levine, "The psychology of interoperability: A systematic review of joint working between the UK emergency services," *J. Occup. Organ. Psychol.*, vol. 97, no. 1, pp. 233–252, Mar. 2024, doi: 10.1111/joop.12469.

[21] D. Koncová and I. Kremeňová, "Interoperability and Its Significance," *Transp. Commun.*, vol. 10, no. 1, pp. 6–9, May 2022, doi: 10.26552/tac.C.2022.1.2.

[22] O. I. Oladele and M. Jood, "Factors affecting adoption of livestock identification and trace-back system among cattle farmers in Kgalagadi district, Botswana," *Livest. Res. Rural Dev.*, vol. 22, no. 8, 2010, [Online]. Available: http://www.lrrd.org/lrrd22/8/olad22147.htm

[23] Ministry of Agricultural Development and Food Security, "Botswana Animal Information and Traceability System (BAITS)," *Government of Botswana*, Botswana, 2018. [Online]. Available: , https://www.gov.bw/sites/default/files/2020-03/(BAITS)%20Farmers%20Guidelines%20Scope.pdf

[24] P. Ntokwane and K. Dibeela, "Utilisation of ICT for improved livestock traceability: a case for Botswana, a developing country perspective," in *Session 110 - Agricultural Libraries*, IFLA Library, 2016. Accessed: Jan. 27, 2025. [Online]. Available: https://library.ifla.org/id/eprint/1491/

[25] N. Derah and M. Mokopasetso, "The control of Foot and Mouth Disease in Botswana and Zimbabwe," *Tropicultura*, vol. 23, pp. 3–7, 2005, [Online]. Available: http://www.tropicultura.org/eng/content/v23ns.html

[26] N. S. Ndubo and J. C. Moreki, "Evaluation of Livestock Identification and Trace-Back System (LITS) in Three Villages of Kweneng District, Botswana," *Afr. J. Livest. Ext.*, vol. 10, pp. 43–50, 2012, [Online]. Available: https://www.ajol.info/index.php/ajlex/article/view/78791

[27] K. Dizyee, D. Baker, and K. M. Rich, "A quantitative value chain analysis of policy options for the beef sector in Botswana," *Agric. Syst.*, vol. 156, pp. 13–24, Sep. 2017, doi: 10.1016/j.agsy.2017.05.007.

[28] Statistics Botswana, "Annual Agricultural Survey Report 2019 – Traditional Sector," Statistics Botswana, Gaborone, 2020. [Online]. Available: https://www.statsbots.org.bw/sites/default/files/publications/ANNUAL%20AGRICULTURAL%20SURVEY%20REPORT%202019%20TRADITIONAL%20SECTOR.pdf.

[29] B. E. Mooketsi, "Optimisation of Livestock Identification and Trace-back System LITS Database to meet Local Needs: Case Study of Botswana," *J. Community Inform.*, vol. 9, no. 4, Dec. 2013, doi: 10.15353/joci.v9i4.3143.

[30] World Organisation for Animal Health, "Botswana PVS Gap Analysis Report: Strengthening Veterinary Services for International Compliance." Paris: World Organisation for Animal Health, 2011. [Online]. Available: https://www.woah.org/fileadmin/Home/eng/Support_to_OIE_Members/pdf/PVSGapAnalysis-Report-Botswana.pdf.

[31] H. J. Du Plessis and G. E. Du Rand, "The significance of traceability in consumer decision making towards Karoo lamb," *Food Res. Int.*, vol. 47, no. 2, pp. 210–217, Jul. 2012, doi: 10.1016/j.foodres.2011.05.029.

[32] T. R. Gruber, "Toward principles for the design of ontologies used for knowledge sharing?," *Int. J. Hum.-Comput. Stud.*, vol. 43, no. 5–6, pp. 907–928, Nov. 1995, doi: 10.1006/ijhc.1995.1081.

[33] N. F. Noy and D. L. McGuinness, “Ontology Development 101: A Guide to Creating Your First Ontology,” *Protege Stanf.*, pp. 1–25, 2001, [Online]. Available: http://protege.stanford.edu/publications/ontology_development/ontology101.pdf

[34] N. F. Noy and M. A. Musen, “Evaluating Ontology-Mapping Tools: Requirements and Experience,” in *CEUR Workshop Proceedings 62*, Siguenza (Spain), Sep. 2002. [Online]. Available: https://ceur-ws.org/Vol-62/EON2002_Noy.pdf

[35] V. Nguyen, “Ontologies and Information Systems: A Literature Survey,” Defence Science and Technology Organization, Australian Government, DSTO-TN-102, Jun. 2011. [Online]. Available: https://apps.dtic.mil/sti/tr/pdf/ADA546186.pdf

[36] S. Yasunaga, M. Nakatsuka, and K. Kuwabara, “Web Ontology Building System for Novice Users: A Step-by-Step Approach,” in *Intelligent Information and Database Systems*, vol. 5991, N. T. Nguyen, M. T. Le, and J. Świątek, Eds., in Lecture Notes in Computer Science, vol. 5991. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 134–143. doi: 10.1007/978-3-642-12101-2_15.

[37] G. Brusa, M. L. Caliusco, and O. Chiotti, “A process for building a domain ontology: an experience in developing a government budgetary ontology,” in *Proceedings of the Second Australasian Ontology Workshop (AOW 2006)*, Hobart, Australia: Australian Computer Society Inc, Dec. 2006, pp. 7–13. [Online]. Available: https://crpit.scem.westernsydney.edu.au/Vol72.html

[38] K. Saravanan, G. S. Mahalakshmi, and T. V. Geetha, “An Ontology Based Approach To Automated Domain-Specific SRS Development,” in *Proceedings of National Conference*, Chennai, Tamilnadu, INDIA: Anna University, 2006, pp. 1–11. [Online]. Available: https://www.researchgate.net/publication/251237637_An_Ontology_Based_Approach_To_Automated_Domain-Specific_SRS_Development

[39] J. Hou, C. Su, Y. Su, and W. Wang, “A Methodology of Knowledge Management Based on Ontology in Collaborative Design,” in *2008 Second International Symposium on Intelligent Information Technology Application*, Shanghai, China: IEEE, Dec. 2008, pp. 409–413. doi: 10.1109/IITA.2008.503.

[40] S. Brinkmann and S. Kvale, “The Qualitative Research Interview as Context,” in *InterViews Learning the Craft of Qualitative Research Interviewing*, 3RD ED., in Chapter 5. , Thousand Oaks, CA: Sage Publications, 2015. [Online]. Available: https://us.sagepub.com/en-us/nam/interviews/book239402#features

[41] J. F. Sequeda and D. P. Miranker, “A Pay-As-You-Go Methodology for Ontology-Based Data Access,” *IEEE Internet Comput.*, vol. 21, no. 2, pp. 92–96, Mar. 2017, doi: 10.1109/MIC.2017.46.

[42] C. Berg, “Interoperability,” *Internet Policy Rev.*, vol. 13, no. 2, Apr. 2024, doi: 10.14763/2024.2.1749.

[43] N. Kaewboonma and K. Tuamsuk, “Application of Ontologies for Knowledge Management,” *TLA Res. J.*, vol. 10, no. 2, Dec. 2017, [Online]. Available: https://so06.tci-thaijo.org/index.php/tla_research/article/view/99757

[44] E. Widajanti and T. Ratnawati, “Information Technology Infrastructure Optimization for Achievinh Innovation-based Company Perfomance,” *Int. J. Bus. Econ.*, vol. 22, no. 1, pp. 180–185, Aug. 2020, [Online]. Available: https://www.ijbel.com/wp-content/uploads/2020/11/IJBEL22_247.pdf

[45] I. Ahmed Al-Baltah, A. Azim Abdul Ghani, W. Nurhayati Wan Ab. Rahman, and R. Atan, “A Comparative Study on Ontology Development Methodologies towards Building Semantic Conflicts Detection Ontology for Heterogeneous Web Services,” *Res. J. Appl. Sci. Eng. Technol.*, vol. 7, no. 13, pp. 2674–2679, Apr. 2014, doi: 10.19026/rjaset.7.584.

[46] A. Gómez-Pérez, M. Fernández-López, and O. Corcho, “Ontology Tools: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web,” in *Ontological Engineering*, 1st ed., in Advanced Information and Knowledge Processing. , London: Springer London, 2004, pp. 293–362. [Online]. Available: https://link.springer.com/book/10.1007/b97353

[47] Rahamatullah Khondoker and Mueller, “Comparing Ontology Development Tools Based on an Online Survey,” in *Proceedings of the World Congress on Engineering 2010*, London, U.K: International Association of Engineers (IAENG), Jul. 2010, pp. 1–5. [Online]. Available: https://www.iaeng.org/publication/WCE2010/WCE2010_pp188-192.pdf

[48] E. Alatrish, “Comparison Some of Ontology Editors,” *Manag. Inf. Syst.*, vol. 8, no. 2, pp. 1–7, 2013, [Online]. Available: https://ef.uns.ac.rs/mis/archive-pdf/2013%20-%20No2/MIS2013-2-4.pdf

[49] J. Heflin and Z. Pan, “A Model Theoretic Semantics for Ontology Versioning,” in *The Semantic Web – ISWC 2004*, vol. 3298, S. A. McIlraith, D. Plexousakis, and F. Van Harmelen, Eds., in Lecture Notes in Computer Science, vol. 3298. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 62–76. doi: 10.1007/978-3-540-30475-3_6.

[50] B. Lantow, “OntoMetrics: Putting Metrics into Use for Ontology Evaluation:,” in *Proceedings of the 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, Porto, Portugal: SCITEPRESS - Science and and Technology Publications, 2016, pp. 186–191. doi: 10.5220/0006084601860191.

[51] J. Zhou, M. N. Shafique, A. Adeel, S. Nawaz, and P. Kumar, “What is Theoretical Contribution? A Narrative Review,” *Sarhad J. Manag. Sci.*, vol. 3, no. 2, pp. 261–271, Dec. 2017, doi: 10.31529/sjms.2017.3.2.6.

[52] D. A. Whetten, “What Constitutes a Theoretical Contribution?,” *Acad. Manage. Rev.*, vol. 14, no. 4, p. 490, Oct. 1989, doi: 10.2307/258554.

[53] H. Kugita, “Implementing OIE guidelines to ensure the excellence of the veterinary profession,” in *Office International des Epizooties (OIE)*, Bangkok, Thailand: World Organisation for Animal Health, Mar. 2018, pp. 1–22. [Online]. Available: https://rr-asia.woah.org/app/uploads/2020/01/3-1_oie_activities_to_support_good_governance_1.pdf

[54] T. R. Levine and D. M. Markowitz, “The role of theory in researching and understanding human communication,” *Hum. Commun. Res.*, vol. 50, no. 2, pp. 154–161, Apr. 2024, doi: 10.1093/hcr/hqad037.

[55] B. Denkena, M.-A. Dittrich, and G. V. Münch, “Ontology-based production planning under the consideration of system robustness,” *Procedia CIRP*, vol. 104, pp. 547–552, 2021, doi: 10.1016/j.procir.2021.11.092.

[56] C. Zhang, D. Garlan, and E. Kang, “A behavioral notion of robustness for software systems,” in *Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, Virtual Event USA: ACM, Nov. 2020, pp. 1–12. doi: 10.1145/3368089.3409753.

[57] J. K. Siror, S. Huanye, D. Wang, and W. Jie, “Use of RFID Technologies to Combat Cattle Rustling in the East Africa,” in *2009 Fifth International Joint Conference on INC, IMS and IDC*, Seoul: IEEE, Aug. 2009, pp. 1556–1562. doi: 10.1109/NCM.2009.146.

[58] R. Addo-Tenkorang, N. Gwangwava, E. N. Ogunmuyiwa, and A. U. Ude, “Advanced Animal Track-&-Trace Supply-Chain Conceptual Framework: An Internet of Things Approach,” *Procedia Manuf.*, vol. 30, pp. 56–63, 2019, doi: 10.1016/j.promfg.2019.02.009.

[59] E. Pereira *et al.*, “RFID Technology for Animal Tracking: A Survey,” *IEEE J. Radio Freq. Identif.*, vol. 7, pp. 609–620, 2023, doi: 10.1109/JRFID.2023.3334952.

[60] H. H. R. Zanetoni *et al.*, “Blockchain applied to the traceability of animal products: a systematic literature review,” *Rev. Ciênc. AGRONÔMICA*, vol. 55, 2024, doi: 10.5935/1806-6690.20240033.

[61] L. Modise, “Botswana Animal Identification Traceability System,” presented at the ICAR Presentation Session, Pretoria, South Africa, 2015. [Online]. Available: : https://www.icar.org/wp-content/uploads/2015/09/Modisa.pdf

[62] L. Yang, X.-Y. Liu, and J. S. Kim, “Cloud-based Livestock Monitoring System Using RFID and Blockchain Technology,” in *2020 7th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2020 6th IEEE International Conference on Edge Computing and Scalable*

Cloud (EdgeCom), New York, NY, USA: IEEE, Aug. 2020, pp. 240–245. doi: 10.1109/CSCloud-EdgeCom49738.2020.00049.

[63] C. Aquilani, A. Confessore, R. Bozzi, F. Sirtori, and C. Pugliese, "Review: Precision Livestock Farming technologies in pasture-based livestock systems," *Animal*, vol. 16, no. 1, p. 100429, Jan. 2022, doi: 10.1016/j.animal.2021.100429.

[64] D. W. Bailey, M. G. Trotter, C. Tobin, and M. G. Thomas, "Opportunities to Apply Precision Livestock Management on Rangelands," *Front. Sustain. Food Syst.*, vol. 5, p. 611915, Apr. 2021, doi: 10.3389/fsufs.2021.611915.

[65] E. A. De Nadai Fernandes, G. A. Sarriés, M. A. Bacchi, Y. T. Mazola, C. L. Gonzaga, and S. R. V. Sarriés, "Trace elements and machine learning for Brazilian beef traceability," *Food Chem.*, vol. 333, p. 127462, Dec. 2020, doi: 10.1016/j.foodchem.2020.127462.

[66] B. Xu *et al.*, "Evaluation of Deep Learning for Automatic Multi-View Face Detection in Cattle," *Agriculture*, vol. 11, no. 11, p. 1062, Oct. 2021, doi: 10.3390/agriculture11111062.

[67] L. Elbakidze, "Economic Benefits of Animal Tracing in the Cattle Production Sector," *J. Agric. Resour. Econ. West. Agric. Econ. Assoc.*, vol. 32, no. 1, pp. 169–180, 2007, [Online]. Available: https://www.jstor.org/stable/40987357

[68] L. A. González, G. Bishop-Hurley, D. Henry, and E. Charmley, "Wireless sensor networks to study, monitor and manage cattle in grazing systems," *Anim. Prod. Sci.*, vol. 54, no. 10, p. 1687, 2014, doi: 10.1071/AN14368.

[69] S. Kumar, S. Tiwari, and S. K. Singh, "Face Recognition of Cattle: Can it be Done?," *Proc. Natl. Acad. Sci. India Sect. Phys. Sci.*, vol. 86, no. 2, pp. 137–148, Jun. 2016, doi: 10.1007/s40010-016-0264-2.

[70] I. Hernandez San Juan and L. González-Vaqué, "The blockchain technology and the regulation of traceability: the digitization of food quality and safety," *Eur. Food Feed Law Rev.*, vol. 15, no. 6, pp. 563–570, 2020.

[71] M. García-Infante, P. Castro-Valdecantos, M. Delgado-Pertíñez, A. Teixeira, J. L. Guzmán, and A. Horcada, "Effectiveness of machine learning algorithms as a tool to meat traceability system. A case study to classify Spanish Mediterranean lamb carcasses," *Food Control*, vol. 164, p. 110604, Oct. 2024, doi: 10.1016/j.foodcont.2024.110604.

[72] S. Gbashi and P. B. Njobeh, "Enhancing Food Integrity through Artificial Intelligence and Machine Learning: A Comprehensive Review," *Appl. Sci.*, vol. 14, no. 8, p. 3421, Apr. 2024, doi: 10.3390/app14083421.

[73] M. Navia, L. Chancay-García, C. Cedeño-Sarmiento, and D. Mendoza-Briones, "A Systematic Literature Review about Sensors Systems and Applications for Livestock Vital Signs Monitoring," *Rev. Politécnica*, vol. 53, no. 2, pp. 69–78, May 2024, doi: 10.33333/rp.vol53n2.07.

[74] L. León, A. Ortiz, M. Freire, F. J. Mesías, and D. Tejerina, "Effectiveness of handheld near infrared spectrometer for traceability of Angus steaks," *Food Chem.*, vol. 455, p. 139958, Oct. 2024, doi: 10.1016/j.foodchem.2024.139958.

[75] I. Fuzesi, M. Herdon, and A. Pentek, "Food Tracing and Interoperability of Information Systems in the Hungarian Meat Industry," 2010, doi: 10.22004/AG.ECON.96866.

[76] F. Marinello, M. Atzori, L. Lisi, and A. Pezzuolo, "Development of a Traceability System for the Animal Product Supply Chain Based on Blockchain Technology," presented at the 8th European Conference on Precision Livestock Farming, 2017, pp. 256–268. [Online]. Available: https://www.researchgate.net/publication/319702558_Development_of_a_traceability_system_for_the_animal_product_supply_chain_based_on_Blockchain_Technology

[77] T. Chen, K. Ding, S. Hao, G. Li, and J. Qu, "Batch-based traceability for pork: A mobile solution with 2D barcode technology," *Food Control*, vol. 107, p. 106770, Jan. 2020, doi: 10.1016/j.foodcont.2019.106770.

[78] P. S. Girish, S. K, K. K, P. Palekar, H. C. H, and S. Rathod, "Artificial intelligence based muzzle recognition technology for individual identification of animals," *Indian J. Anim. Sci.*, vol. 90, no. 7, pp. 1070–1073, Oct. 2020, doi: 10.56093/ijans.v90i7.106684.

[79] M. Tripoli and J. Schmidhuber, “Optimising traceability in trade for live animals and animal products with digital technologies: -EN- -FR- Optimiser la traçabilité des animaux vivants et des produits d’origine animale faisant l’objet d’échanges internationaux grâce aux technologies numériques -ES- Uso de la tecnología digital para optimizar la trazabilidad en el comercio de animales vivos y productos de origen animal,” *Rev. Sci. Tech. OIE*, vol. 39, no. 1, pp. 235–244, Apr. 2020, doi: 10.20506/rst.39.1.3076.

[80] M. Arvana, A. D. Rocha, and J. Barata, “Agri-Food Value Chain Traceability Using Blockchain Technology: Portuguese Hams’ Production Scenario,” *Foods*, vol. 12, no. 23, p. 4246, Nov. 2023, doi: 10.3390/foods12234246.

[81] V. Medennikov, “RETRACTED: Digital twin of livestock production in unified digital platform of Russian agriculture management,” *BIO Web Conf.*, vol. 116, p. 02012, 2024, doi: 10.1051/bioconf/202411602012.

[82] A. Walunj and R. Gourkar, “Blockchain in Animal Health Care and Their Records Maintenance,” *Int. J. Multidiscip. Res. Sci. Eng. Technol.*, vol. 7, no. 6, pp. 11044–11047, 2024, [Online]. Available: https://www.researchgate.net/publication/381739368_INTERNATIONAL_JOURNAL_OF_MULTIDISCIPLINARY_RESEARCH_Blockchain_in_Animal_Health_Care_and_their_Records_Maintenance

[83] J. Witt *et al.*, “Relationship between animal-based on-farm indicators and meat inspection data in pigs,” *Porc. Health Manag.*, vol. 10, no. 1, p. 8, Jan. 2024, doi: 10.1186/s40813-024-00359-9.

[84] K. Mann, “The Largest Agriculture Sector in Botswana: Cattle and Livestock.” ProAgri Media, Oct. 2024. [Online]. Available: https://proagri.co.za/the-largest-agriculture-sector-in-botswana-cattle-and-livestock

[85] D. M. Chatibura, “Beef as intangible cultural heritage in Botswana: a documentary review,” *J. Ethn. Foods*, vol. 10, no. 1, p. 41, Nov. 2023, doi: 10.1186/s42779-023-00199-y.

[86] Botswana Meat Commission (BMC), “Botswana Beef.” Botswana Meat Commission (BMC), 2024. [Online]. Available: https://bmc.bw/botswana-beef

[87] Government of Botswana, “Ministry of Agricultural Development and Food Security.” 2022. [Online]. Available: https://www.gov.bw/ministry-agriculture

[88] R. Gofamodimo, “Botswana Cattle Population Declines Drastically.” Farmers’ Review, 2022. [Online]. Available: https://farmersreview.co.bw

[89] Food and Agriculture Organization of the United Nations (FAO), *The State of Food and Agriculture 2023 - Revealing the true cost of food to transform agrifood systems*. Rome: FAO, 2023. doi: 10.4060/cc7724en.

[90] O. Temoso, R. Villano, and D. Hadley, “Evaluating the productivity gap between commercial and traditional beef production systems in Botswana,” *Agric. Syst.*, vol. 149, pp. 30–39, Nov. 2016, doi: 10.1016/j.agsy.2016.07.014.

[91] M. Fanikiso, “Animal Identification and Traceability: Public sector perspective and experience from Botswana,” in *Public sector perspectives and experiences*, Buenos Aires, Argentina: World Organisation for Animal Healthcare, Mar. 2009, pp. 1–15. [Online]. Available: https://www.woah.org/fileadmin/Home/eng/Conferences_Events/sites/traceability-2009/documents.html

[92] K. Braunschweig, J. Eberius, M. Thiele, and W. Lehner, “The State of Open Data Limits of Current Open Data Platforms,” in *Proceedings of the 21st international conference on World Wide Web*, Lyon France: Association for Computing Machinery, Apr. 2012. [Online]. Available: https://wwwdb.inf.tu-dresden.de/opendatasurvey/www2012_short.pdf

[93] R. L. Boy, “Cattle Traceability - A Threat to Sustainable Supply of Beef to EU: A Botswana Meat Commission,” *Eur. J. Logist. Purch. Supply Chain Manag.*, vol. 1, no. 1, pp. 1–9, Jun. 2013, [Online]. Available: https://www.eajournals.org/wp-content/uploads/Cattle-Traceability-A-Threat-to-Sustainable-Supply-Of-Beef-To-EU-A-Botswana-Meat-Commission.pdf

[94] C. Monyatsiwa, D. Ferguson, and N. Phumaphi, “Bridging the IT divide — Case study on the Botswana livestock industry,” presented at the 2013 IST-Africa Conference & Exhibition, Nairobi, Kenya: IEEE, Jan. 2014, pp. 1–11. [Online]. Available: https://ieeexplore.ieee.org/document/6701774

[95] M. Verdonck, F. Gailly, S. De Cesare, and G. Poels, “Ontology-driven conceptual modeling: A systematic literature mapping and review,” *Appl. Ontol.*, vol. 10, no. 3–4, pp. 197–227, Dec. 2015, doi: 10.3233/AO-150154.

[96] O. Corcho, M. Fernández-López, and A. Gómez-Pérez, “Methodologies, tools and languages for building ontologies. Where is their meeting point?,” *Data Knowl. Eng.*, vol. 46, no. 1, pp. 41–64, Jul. 2003, doi: 10.1016/S0169-023X(02)00195-7.

[97] M. Hristozova and L. Sterling, “Experiences with Ontology Development for Value-Added Publishing,” in *Proceedings of the Workshop on Ontologies in Agent Systems (OAS 2003)*, Melbourne, Australia: DBLP Computer Science Bibliography, 17-24, pp. 17–24. [Online]. Available: https://dblp.org/db/conf/oas/oas2003.html

[98] M. Grüninger and M. Fox, “Methodology for the Design and Evaluation of Ontologies,” in *Workshop on Basic Ontological Issues in Knowledge Sharing*, Montreal, Canada, 1995, pp. 1–10. [Online]. Available: http://www.eil.utoronto.ca/wp-content/uploads/enterprise-modelling/papers/gruninger-ijcai95.pdf

[99] M. C. Bravo Contreras, L. F. Hoyos Reyes, and J. A. Reyes Ortiz, “Methodology for ontology design and construction,” *Contad. Adm.*, vol. 64, no. 4, p. 134, Mar. 2019, doi: 10.22201/fca.24488410e.2020.2368.

[100] M. Uschold and M. King, “Towards a Methodology for Building Ontologies,” in *Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing*, Montreal, Canada: IJCAI Inc, Jul. 1995, pp. 1–13. [Online]. Available: https://www.aiai.ed.ac.uk/publications/documents/1995/95-ont-ijcai95-ont-method.pdf

[101] M. F. Lopez, A. Gomez-Perez, J. P. Sierra, and A. P. Sierra, “Building a chemical ontology using Methontology and the Ontology Design Environment,” *IEEE Intell. Syst.*, vol. 14, no. 1, pp. 37–46, Jan. 1999, doi: 10.1109/5254.747904.

[102] M. Fernández, C. Overbeeke, M. Sabou, and E. Motta, “What Makes a Good Ontology? A Case-Study in Fine-Grained Knowledge Reuse,” in *The Semantic Web*, vol. 5926, A. Gómez-Pérez, Y. Yu, and Y. Ding, Eds., in Lecture Notes in Computer Science, vol. 5926. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 61–75. doi: 10.1007/978-3-642-10871-6_5.

[103] M. R. Bautista-Zambrana, “Methodologies to Build Ontologies for Terminological Purposes,” *Procedia - Soc. Behav. Sci.*, vol. 173, pp. 264–269, Feb. 2015, doi: 10.1016/j.sbspro.2015.02.063.

[104] Y. Alfaifi, “Ontology Development Methodology: A Systematic Review and Case Study,” in *2022 2nd International Conference on Computing and Information Technology (ICCIT)*, Tabuk, Saudi Arabia: IEEE, Jan. 2022, pp. 446–450. doi: 10.1109/ICCIT52419.2022.9711664.

[105] N. Siti Husnah Ab Rahim Periamalai, A. Mustapha, and A. Alqurneh, “An ontology for Juz’ Amma based on expert knowledge,” in *2016 7th International Conference on Computer Science and Information Technology (CSIT)*, Amman: IEEE, Jul. 2016, pp. 1–5. doi: 10.1109/CSIT.2016.7549480.

[106] T. Slimani, “Ontology Development: A Comparing Study on Tools, Languages and Formalisms,” *Indian J. Sci. Technol.*, vol. 8, no. 1, pp. 1–12, Jan. 2015, doi: 10.17485/ijst/2015/v8i34/54249.

[107] P. Gokhale, S. Deokattey, and K. Bhanumurthy, “Ontology Development Methods,” *DESIDOC J. Libr. Inf. Technol.*, vol. 31, no. 2, pp. 77–83, Mar. 2011, doi: 10.14429/djlit.31.2.860.

[108] O. Corcho, M. Fernández-López, and A. Gómez-Pérez, “Ontological Engineering: What are Ontologies and How Can We Build Them?,” in *Semantic Web Services: Theory, Tools and Applications*, IGI Global, 2007, pp. 44–62. [Online]. Available: https://www.igi-global.com/chapter/ontological-engineering-ontologies-can-build/28879

[109] A. Stellato *et al.*, “VocBench 3: A collaborative Semantic Web editor for ontologies, thesauri and lexicons,” *Semantic Web*, vol. 11, no. 5, pp. 855–881, Aug. 2020, doi: 10.3233/SW-200370.

[110] P. Stolic, D. Milosevic, Z. Stevic, and I. Radovanovic, “Ontology Development for Creating Identical Software Environments to Improve Learning Outcomes in Higher Education Institutions,” *Electronics*, vol. 12, no. 14, p. 3057, Jul. 2023, doi: 10.3390/electronics12143057.
[111] A. Stadnicki, F. Filip Pietroń, and P. Burek, “Towards a Modern Ontology Development Environment,” *Procedia Comput. Sci.*, vol. 176, pp. 753–762, 2020, doi: 10.1016/j.procs.2020.09.070.
[112] B. Kapoor and S. Sharma, “A Comparative Study of Ontology building Tools in Semantic Web Applications,” *Int. J. Web Semantic Technol.*, vol. 1, no. 3, pp. 1–13, Jul. 2010, doi: 10.5121/ijwest.2010.1301.
[113] R. Garcia-Castro and A. Gomez-Perez, “Perspectives in Semantic Interoperability:,” in *Proceedings of the International Workshop on Semantic Interoperability*, Rome, Italy: SciTePress - Science and and Technology Publications, 2011, pp. 13–22. doi: 10.5220/0003346700130022.
[114] D. Kalibatiene and O. Vasilecas, “Survey on Ontology Languages,” in *Perspectives in Business Informatics Research*, vol. 90, J. Grabis and M. Kirikova, Eds., in Lecture Notes in Business Information Processing, vol. 90. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 124–141. doi: 10.1007/978-3-642-24511-4_10.
[115] X. Su and L. Ilebrekke, “A Comparative Study of Ontology Languages and Tools,” in *Active Flow and Combustion Control 2018*, vol. 141, R. King, Ed., in Notes on Numerical Fluid Mechanics and Multidisciplinary Design, vol. 141. , Cham: Springer International Publishing, 2002, pp. 761–765. doi: 10.1007/3-540-47961-9_62.
[116] P. Hitzler, M. Krotzsch, B. Parsia, P. F. Patel-Schneider, and S. Rudolph, “OWL 2 Web Ontology Language Primer,” Wright State University, Ohio, America, Sep. 2012. [Online]. Available: https://corescholar.libraries.wright.edu/cse/17/
[117] V. Maniraj and R. Sivakumar, “Ontology Languages – A Review,” *Int. J. Comput. Theory Eng.*, pp. 887–891, 2010, doi: 10.7763/IJCTE.2010.V2.257.
[118] V. Fuentes, T. Martin, P. Valtchev, A. B. Diallo, R. Lacroix, and M. Leduc, “DCPO: The Dairy Cattle Performance Ontology, a Tool for Domain Modelling and Data Analytics,” in *Proceedings of the ICAR Conference*, The Global Standard for Livestock Data, 2022, pp. 57–65. [Online]. Available: https://www.icar.org/Documents/technical_series/ICAR-Technical-Series-no-26-Montreal/11%20DCPO%20Dairy%20cattle%20performance%20ontology.pdf
[119] T. Pizzuti, G. Mirabelli, M. A. Sanz-Bobi, and F. Goméz-Gonzaléz, “Food Track & Trace ontology for helping the food traceability control,” *J. Food Eng.*, vol. 120, pp. 17–30, Jan. 2014, doi: 10.1016/j.jfoodeng.2013.07.017.
[120] S. F. Mohamad Hashim, J. Salim, S. A. M. N. Mohd Noah, and W. A. Wan Mustapha, “Ontology-Based Traceability System for Halal Status of Flavour: A Conceptual Framework,” *Malays. J. Inf. Commun. Technol. MyJICT*, pp. 65–77, Dec. 2023, doi: 10.53840/myjict8-2-97.
[121] J. Storøy, M. Thakur, and P. Olsen, “The TraceFood Framework – Principles and guidelines for implementing traceability in food value chains,” *J. Food Eng.*, vol. 115, no. 1, pp. 41–48, Mar. 2013, doi: 10.1016/j.jfoodeng.2012.09.018.
[122] T. Pizzuti, G. Mirabelli, G. Grasso, and G. Paldino, “MESCO (MEat Supply Chain Ontology): An ontology for supporting traceability in the meat supply chain,” *Food Control*, vol. 72, pp. 123–133, Feb. 2017, doi: 10.1016/j.foodcont.2016.07.038.
[123] B. V. Batlajery, M. Weal, A. Chapman, and L. Moreau, “prFood: Ontology Principles for Provenance and Risk in the Food Domain,” in *2018 IEEE 12th International Conference on Semantic Computing (ICSC)*, Laguna Hills, CA, USA: IEEE, Jan. 2018, pp. 17–24. doi: 10.1109/ICSC.2018.00012.
[124] S. Ouf, “Ontology-Based Knowledge Modelling for Food Supply Chain Data Representation:,” *Int. J. E-Collab.*, vol. 18, no. 1, pp. 1–15, May 2022, doi: 10.4018/IJeC.299009.
[125] saloua Allouch, K. Amechnoue, and I. Achatbi, “An ontological approach to model outbound logistics based on Internet of things (OLP-IoT),” *Int. J. Supply Oper. Manag.*, vol. 10, no. 4, pp. 456–464, Aug. 2023, doi: 10.22034/ijsom.2023.109230.2270.

[126] P. Hemnil, E. Boonchieng, and S. Osiriphun, “Food Safety Block Chain System using Ontology in UHT Milk Factory,” in *2023 17th International Conference on Ubiquitous Information Management and Communication (IMCOM)*, Seoul, Korea, Republic of: IEEE, Jan. 2023, pp. 1–5. doi: 10.1109/IMCOM56909.2023.10035606.

[127] M. Frendi, B. Nachet, and A. Adla, “Physical Internet Based Ontology for Supporting Traceabilityin Logistic IoT,” *Int. J. Comput. Digit. Syst.*, vol. 15, no. 1, pp. 427–440, Feb. 2024, doi: 10.12785/ijcds/150133.

[128] B. M. Asenahabi, “Basics of Research Design: A Guide to selecting appropriate research design,” *Int. J. Contemp. Appl. Res.*, vol. 6, no. 5, pp. 76–89, May 2019, [Online]. Available: http://www.ijcar.net/assets/pdf/Vol6-No5-May2019/07.-Basics-of-Research-Design-A-Guide-to-selecting-appropriate-research-design.pdf

[129] M. R. Ansari, K. Rahim, R. Bhoje, and S. Bhosale, “A Study on Research Design and its Types,” *Int. Res. J. Eng. Technol. IRJET*, vol. 9, no. 7, pp. 1132–1134, Jul. 2022, [Online]. Available: https://www.irjet.net/archives/V9/i7/IRJET-V9I7216.pdf

[130] G. Ditsa, “A Research Design and a Methodological Approach to an Explanatory User Behavior Testing: Lessons Learnt,” in *Innovations Through Information Technology*, Hershey, PA 17033: IDEA Publishing, 2004, pp. 763–766. [Online]. Available: https://bahrami.tums.ac.ir/uploads/112/old/A-Research-Design-and-a-Methodological-Approach-to-an-Explanatory-User-Behavior-Testing.pdf

[131] M. Patel and N. Patel, “Exploring Research Methodology: Review Article,” *Int. J. Res. Rev.*, vol. 6, no. 3, pp. 48–54, Mar. 2019, [Online]. Available: https://www.ijrrjournal.com/IJRR_Vol.6_Issue.3_March2019/IJRR0011.pdf

[132] S. Almalki, “Integrating Quantitative and Qualitative Data in Mixed Methods Research—Challenges and Benefits,” *J. Educ. Learn.*, vol. 5, no. 3, p. 288, Jul. 2016, doi: 10.5539/jel.v5n3p288.

[133] J. W. Creswell, “Five Qualitative Approaches to Inquiry: Choosing Among Five Approached,” in *Qualitative Inquiry and Research Design*, 2nd ed., Thousand Oaks, California: Sage Publications, 2007, pp. 53–84. [Online]. Available: https://revistapsicologia.org/public/formato/cuali2.pdf

[134] Creswell and J. D. Creswell, *Research Design Qualitative, Quantitative, and Mixed Methods Approaches*, 4th ed. Thousand Oaks, California: SAGE Publications, Inc, 2019. [Online]. Available: https://us.sagepub.com/en-us/nam/research-design/book270550

[135] L. A. Palinkas, S. M. Horwitz, C. A. Green, J. P. Wisdom, N. Duan, and K. Hoagwood, “Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research,” *Adm. Policy Ment. Health Ment. Health Serv. Res.*, vol. 42, no. 5, pp. 533–544, Sep. 2015, doi: 10.1007/s10488-013-0528-y.

[136] J. M. Morse, “Principles of mixed methods and multi-method re search design,” in *Handbook of mixed methods in social and behavioral research*, Thousand Oaks, California: Sage Publication, 2003, pp. 189–208. [Online]. Available: https://www.researchgate.net/publication/224927643_Principles_of_Mixed_Methods_and_Multimethod_Research_Design

[137] G. Radhakrishnan, “Sampling in Mixed Methods Research,” *Int. J. Adv. Nurs. Manag.*, vol. 2, no. 1, pp. 24–27, Mar. 2014, [Online]. Available: https://ijanm.com/HTMLPaper.aspx?Journal=International%20Journal%20of%20Advances%20in%20Nursing%20Management;PID=2014-2-1-7

[138] W. E. Hanson, J. W. Creswell, V. L. P. Clark, K. S. Petska, and J. D. Creswell, “Mixed methods research designs in counseling psychology.,” *J. Couns. Psychol.*, vol. 52, no. 2, pp. 224–235, Apr. 2005, doi: 10.1037/0022-0167.52.2.224.

[139] S. Dawadi, S. Shrestha, and R. A. Giri, “Mixed-Methods Research: A Discussion on its Types, Challenges, and Criticisms,” *J. Pract. Stud. Educ.*, vol. 2, no. 2, pp. 25–36, Feb. 2021, doi: 10.46809/jpse.v2i2.20.

[140] J. W. Creswell and T. C. Guetterman, *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*, 7th ed. Boston:MA, United States: Pearson, 2024. [Online]. Available: https://www.pearson.com/en-us/subject-catalog/p/educational-research-planning-conducting-and-evaluating-quantitative-and-qualitative-research/P200000010566/9780138173784?tab=table-of-contents

[141] O. D. Apuke, "Quantitative Research Methods : A Synopsis Approach," *Kuwait Chapter Arab. J. Bus. Manag. Rev.*, vol. 6, no. 11, pp. 40–47, Sep. 2017, doi: 10.12816/0040336.

[142] E. Daniel, "The Usefulness of Qualitative and Quantitative Approaches and Methods in Researching Problem-Solving Ability in Science Education Curriculum," *J. Educ. Pract.*, vol. 7, no. 15, pp. 91–100, 2016, [Online]. Available: https://eric.ed.gov/?id=EJ1103224

[143] H. K. Mohajan, "QUALITATIVE RESEARCH METHODOLOGY IN SOCIAL SCIENCES AND RELATED SUBJECTS," *J. Econ. Dev. Environ. People*, vol. 7, no. 1, p. 23, Mar. 2018, doi: 10.26458/jedep.v7i1.571.

[144] U. Chinyere and E. Val, "Qualitative Research," *IDOSR J. Comput. Appl. Sci.*, vol. 8, no. 1, pp. 20–35, 2023, [Online]. Available: https://www.idosr.org/qualitative-research/

[145] H. Taherdoost, "Data Collection Methods and Tools for Research; A Step-by-Step Guide to Choose Data Collection Technique for Academic and Business Research Projects," *Int. J. Acad. Res. Manag.*, vol. 10, no. 1, pp. 10–38, 2021, [Online]. Available: https://ideas.repec.org/p/hal/journl/hal-03741834.html

[146] R. Sarmento and V. Costa, *Comparative Approaches to Using R and Python for Statistical Data Analysis:* in Advances in Systems Analysis, Software Engineering, and High Performance Computing. IGI Global, 2017. doi: 10.4018/978-1-68318-016-6.

[147] S. M. Ospina, M. Esteve, and S. Lee, "Assessing Qualitative Studies in Public Administration Research," *Public Adm. Rev.*, vol. 78, no. 4, pp. 593–605, Jul. 2018, doi: 10.1111/puar.12837.

[148] B. Hancock, E. Windridge, and K. Ockleford, "An Introduction to Qualitative," National Institute for Health Research, United Kingdom, 2009. [Online]. Available: https://dl.icdst.org/pdfs/files3/315d0c3a18c9426593c9f5019506a335.pdf

[149] L. Busetto, W. Wick, and C. Gumbinger, "How to use and assess qualitative research methods," *Neurol. Res. Pract.*, vol. 2, no. 1, p. 14, Dec. 2020, doi: 10.1186/s42466-020-00059-z.

[150] S.-M. Kabir, "Introduction to Research," in *Basic Guidelines for Research: An Introductory Approach for All Disciplines*, 1st ed., Chittagong-4203, Bangladesh: Book Zone Publication, 2016. [Online]. Available: https://www.researchgate.net/profile/Syed-Muhammad-Kabir/publication/325390597_BASIC_GUIDELINES_FOR_RESEARCH_An_Introductory_Approach_for_All_Disciplines/links/5b0a89094585157f8719626c/BASIC-GUIDELINES-FOR-RESEARCH-An-Introductory-Approach-for-All-Disciplines.pdf

[151] R. Roberts, "Qualitative Interview Questions: Guidance for Novice Researchers," *Qual. Rep.*, Sep. 2020, doi: 10.46743/2160-3715/2020.4640.

[152] S. Bhome *et al.*, *Research Methodology (Tools and Analysis)*, 1st ed. Mumbai, India: Himalaya Publishing House, 2015. [Online]. Available: https://www.drnishikantjha.com/papersCollection/Research%20Methodology%20.pdf

[153] M. Castillo-Montoya, "Preparing for Interview Research: The Interview Protocol Refinement Framework," *Qual. Rep.*, May 2016, doi: 10.46743/2160-3715/2016.2337.

[154] M. Peredaryenko and S. Krauss, "Calibrating the Human Instrument: Understanding the Interviewing Experience of Novice Qualitative Researchers," *Qual. Rep.*, Jan. 2015, doi: 10.46743/2160-3715/2013.1449.

[155] R. Kumar, *Research Methodology: A Step by Step Guide for Beginners*, 4th ed. New Delhi, India: SAGE Publications, Inc, 2015. [Online]. Available: http://www.sociology.kpi.ua/wp-content/uploads/2014/06/Ranjit_Kumar-Research_Methodology_A_Step-by-Step_G.pdf

[156] S. Roopa and M. Rani, "Questionnaire Designing for a Survey," *J. Indian Orthod. Soc.*, vol. 46, pp. 273–277, Oct. 2012, doi: 10.5005/jp-journals-10021-1104.

[157] N. Parsons, “Botswana | Capital, Population, People, Currency, & Language | Britannica,” Britannica. [Online]. Available: https://www.britannica.com/place/Botswana
[158] S. Campbell *et al.*, “Purposive sampling: complex or simple? Research case examples,” *J. Res. Nurs.*, vol. 25, no. 8, pp. 652–661, Dec. 2020, doi: 10.1177/1744987120927206.
[159] I. Etikan, “Comparison of Convenience Sampling and Purposive Sampling,” *Am. J. Theor. Appl. Stat.*, vol. 5, no. 1, p. 1, 2016, doi: 10.11648/j.ajtas.20160501.11.
[160] Ma. D. C. Tongco, “Purposive Sampling as a Tool for Informant Selection,” *Ethnobot. Res. Appl.*, vol. 5, p. 147, Dec. 2007, doi: 10.17348/era.5.0.147-158.
[161] O. Isaac, “Convenience and Purposive Sampling Techniques: Are they the Same?,” *Int. J. Innov. Soc. Sci. Educ. Res.*, vol. 11, no. 1, pp. 1–7, 2023, [Online]. Available: https://www.studocu.com/row/document/university-of-the-punjab/quantitative-research-methods/ijisser-m-1-2023/83845929
[162] J. Omona, “Sampling in Qualitative Research: Improving the Quality of Research Outcomes in Higher Education,” *Makerere J. High. Educ.*, vol. 4, no. 2, pp. 169–185, Jul. 2013, doi: 10.4314/majohe.v4i2.4.
[163] E. Badu, A. P. O’Brien, and R. Mitchell, “An integrative review on methodological considerations in mental health research – design, sampling, data collection procedure and quality assurance,” *Arch. Public Health*, vol. 77, no. 1, p. 37, Dec. 2019, doi: 10.1186/s13690-019-0363-z.
[164] H. Ames, C. Glenton, and S. Lewin, “Purposive sampling in a qualitative evidence synthesis: a worked example from a synthesis on parental perceptions of vaccination communication,” *BMC Med. Res. Methodol.*, vol. 19, no. 1, p. 26, Dec. 2019, doi: 10.1186/s12874-019-0665-4.
[165] T. C. Guetterman, “Descriptions of Sampling Practices Within Five Approaches to Qualitative Research in Education and the Health Sciences,” *Forum Qual. Sozialforschung Forum Qual. Soc. Res.*, vol. Vol 16, p. No 2 (2015), May 2015, doi: 10.17169/FQS-16.2.2290.
[166] T. Prinsloo and C. De Villiers, “A Framework to Define the Impact of Sustainable ICT for Agriculture Projects: The Namibian Livestock Traceability System,” *Electron. J. Inf. Syst. Dev. Ctries.*, vol. 82, no. 1, pp. 1–22, Sep. 2017, doi: 10.1002/j.1681-4835.2017.tb00606.x.
[167] L. Sürücü and A. Maslakçi, “VALIDITY AND RELIABILITY IN QUANTITATIVE RESEARCH,” *Bus. Manag. Stud. Int. J.*, vol. 8, no. 3, pp. 2694–2726, Sep. 2020, doi: 10.15295/bmij.v8i3.1540.
[168] B. B. Kawulich, “Participant Observation as a Data Collection Method,” *Forum Qual. Sozialforschung Forum Qual. Soc. Res.*, vol. Vol 6, p. Reuse, May 2005, doi: 10.17169/FQS-6.2.466.
[169] M. E. Amin, *Social Science Research: Conception, Methodology and Analysis*. Uganda: Makerere University, 2005. [Online]. Available: https://books.google.co.bw/books/about/Social_Science_Research.html?id=qundMQAACAAJ&redir_esc=y
[170] H. Taherdoost, “Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research,” *SSRN Electron. J.*, 2016, doi: 10.2139/ssrn.3205040.
[171] A. R. Kunselman, “A brief overview of pilot studies and their sample size justification,” *Fertil. Steril.*, vol. 121, no. 6, pp. 899–901, Jun. 2024, doi: 10.1016/j.fertnstert.2024.01.040.
[172] M. Ibrahim, “The Art of data analysis,” *J. Allied Health Sci. Pak.*, vol. 1, no. 1, pp. 98–104, 2015, [Online]. Available: https://www.researchgate.net/publication/283269432_The_art_of_Data_Analysis
[173] G. Marshall and L. Jonker, “An introduction to descriptive statistics: A review and practical guide,” *Radiography*, vol. 16, no. 4, pp. e1–e7, Nov. 2010, doi: 10.1016/j.radi.2010.01.001.
[174] C. R. Kothari, *Research Methodology: Methods and Techniques*, 2nd ed. 4835/24, Ansari Road, Daryaganj, New Delhi - 110002: New Age International Publishers, 2004. [Online]. Available: http://dl.saintgits.org/jspui/bitstream/123456789/1133/1/Research%20Methodology%20C%20R%20Kothari%20%28Eng%29%201.81%20MB.pdf
[175] P. Kaur, J. Stoltzfus, and V. Yellapu, “Descriptive statistics,” *Int. J. Acad. Med.*, vol. 4, no. 1, p. 60, 2018, doi: 10.4103/IJAM.IJAM_7_18.

[176] K. Jackson and P. Bazeley, *Qualitative Data Analysis with NVivo*, 3rd ed. London: SAGE Publications Ltd, 2019. [Online]. Available: https://us.sagepub.com/sites/default/files/upm-assets/114916_book_item_114916.pdf
[177] S. Dollah, A. Abduh, and Ms. Rosmaladewi, "Benefits and Drawbacks of NVivo QSR Application," in *Proceedings of the 2nd International Conference on Education, Science, and Technology (ICEST 2017)*, Makassar, Indonesia: Atlantis Press, 2017. doi: 10.2991/icest-17.2017.21.
[178] H. H. AlYahmady and S. S. Al Abri, "Using Nvivo for Data Analysis in Qualitative Research," *Int. Interdiscip. J. Educ.*, vol. 2, no. 2, pp. 181–186, Feb. 2013, doi: 10.12816/0002914.
[179] E. Welsh, "Dealing with Data: Using NVivo in the Qualitative Data Analysis Process," *Forum Qual. Sozialforschung Forum Qual. Soc. Res.*, vol. Vol 3, p. No 2 (2002): Using Technology in the Qualitative Research Process, May 2002, doi: 10.17169/FQS-3.2.865.
[180] D. B. Allsop, J. M. Chelladurai, E. R. Kimball, L. D. Marks, and J. J. Hendricks, "Qualitative Methods with Nvivo Software: A Practical Guide for Analyzing Qualitative Data," *Psych*, vol. 4, no. 2, pp. 142–159, Mar. 2022, doi: 10.3390/psych4020013.
[181] B. Johnson and L. Christensen, *Educational Research: Quantitative, Qualitative, and Mixed Approaches*, 7th ed. Thousand Oaks, California: SAGE Publications, Inc, 2019. [Online]. Available: https://us.sagepub.com/en-us/nam/educational-research/book259335
[182] K. Caine, "Local Standards for Sample Size at CHI," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, San Jose California USA: ACM, May 2016, pp. 981–992. doi: 10.1145/2858036.2858498.
[183] M. K. Bhatia, "Data Analysis and its Importance," *Int. Res. J. Adv. Eng. Sci.*, vol. 2, no. 1, pp. 166–168, 2017, [Online]. Available: https://irjaes.com/wp-content/uploads/2020/10/IRJAES-V2N1P58Y17.pdf
[184] H.-F. Hsieh and S. E. Shannon, "Three Approaches to Qualitative Content Analysis," *Qual. Health Res.*, vol. 15, no. 9, pp. 1277–1288, Nov. 2005, doi: 10.1177/1049732305276687.
[185] O. Maureen, "The NVivo toolkit- how to apply NVivo in your PhD for research and publishing success Australia: QSR International," University of Sunshine Coast, QSR International. Australia, 2013. [Online]. Available: https://research.usc.edu.au/esploro/outputs/book/The-NVivo-Toolkit-How-to-apply/99448826502621
[186] J. Saldana, *The Coding Manual for Qualitative Researchers*, 4th ed. Thousand Oaks, California: Sage Publications, 2021. [Online]. Available: https://us.sagepub.com/en-us/nam/the-coding-manual-for-qualitative-researchers/book273583
[187] V. Scherman, L. Zimmerman, and B. Smit, "Mixed Method Data Analysis: An Exploratory Approach to Strengthening Inferences about Relationships and Affinities," *Int. J. Mult. Res. APPROACHES*, vol. 10, no. 1, pp. 57–76, Jul. 2018, doi: 10.29034/ijmra.v10n1a4.
[188] L. Tonin *et al.*, "USE OF NVIVO 10® SOFTWARE IN CONCEPT ANALYSIS STUDY," *Texto Contexto - Enferm.*, vol. 32, p. e20230033, 2023, doi: 10.1590/1980-265x-tce-2023-0033en.
[189] B. M. Edhlund and A. G. McDougall, *NVivo 10 essentials : your guide to the world's most powerful data analysis software*. Stallarholmen, Sweden: Form & Kunskap AB, 2012. [Online]. Available: https://books.google.co.in/books?id=DiWGDwAAQBAJ&printsec=frontcover#v=onepage&q&f=false
[190] A. J. Hutchison, L. H. Johnston, and J. D. Breckon, "Using QSR-NVivo to facilitate the development of a grounded theory project: an account of a worked example," *Int. J. Soc. Res. Methodol.*, vol. 13, no. 4, pp. 283–302, Oct. 2010, doi: 10.1080/13645570902996301.
[191] U. Flick, *The SAGE Handbook of Qualitative Data Analysis*. 1 Oliver's Yard, 55 City Road London EC1Y 1SP: SAGE Publications, Inc., 2014. doi: 10.4135/9781446282243.
[192] F. Shafique and K. Mahmood, "Model Development as a Research Tool: An Example of PAK-NISEA," *Libr. Philos. Pract.*, vol. 427, pp. 1–13, 2010, [Online]. Available: https://digitalcommons.unl.edu/libphilprac/427

[193] S. Loeb, S. Dynarski, D. McFarland, P. Morris, S. Reardon, and S. Reber, “Descriptive Analysis in Education: A Guide for Researchers,” Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Washington, DC: U.S. Department of Education, 4023, 2017. [Online]. Available: https://eric.ed.gov/?id=ED573325

[194] C. A. DePaolo and K. Wilkinson, “Get Your Head into the Clouds: Using Word Clouds for Analyzing Qualitative Assessment Data,” *TechTrends*, vol. 58, no. 3, pp. 38–44, May 2014, doi: 10.1007/s11528-014-0750-9.

[195] M. Jadeja and K. Shah, “TREE-MAP: A VISUALIZATION TOOL FOR LARGE DATA,” in *Proceedings of the First International Workshop on Graph Search and Beyond*, Santiago, Chile: CEUR-WS, Aug. 2015, pp. 1–5. [Online]. Available: https://ceur-ws.org/Vol-1393/paper-07.pdf

[196] J. Wilson and V. Bruni-Bossio, “Using Visual Mapping to Communicate Connections Between Learning Outcomes and Student Tasks,” *Int. J. Des. Learn.*, vol. 11, no. 3, pp. 13–22, Jul. 2020, doi: 10.14434/ijdl.v11i3.26101.

[197] H. J. Chokshi and R. Panchal, “Using Apache Jena Fuseki Server for Execution of SPARQL Queries in Job Search Ontology Using Semantic Technology,” *Int. J. Innov. Res. Comput. Sci. Technol.*, vol. 10, no. 2, pp. 497–504, 2022, doi: https://doi.org/10.55524/.

[198] Ari Fajar Santoso and Ahmad Suryan, “Modelling Ontology of Craft Convection as Small and Medium Enterprises,” *Bincang Sains Dan Teknol.*, vol. 1, no. 01, pp. 26–40, Aug. 2022, doi: 10.56741/bst.v1i01.47.

[199] O. Corcho, M. Fernández-López, A. Gómez-Pérez, and A. López-Cima, “Building Legal Ontologies with METHONTOLOGY and WebODE,” in *Law and the Semantic Web*, vol. 3369, V. R. Benjamins, P. Casanovas, J. Breuker, and A. Gangemi, Eds., in Lecture Notes in Computer Science, vol. 3369. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 142–157. doi: 10.1007/978-3-540-32253-5_9.

[200] J. Park, K. Sung, and S. Moon, “Developing Graduation Screen Ontology Based on the METHONTOLOGY Approach,” in *2008 Fourth International Conference on Networked Computing and Advanced Information Management*, Gyeongju, South Korea: IEEE, Sep. 2008, pp. 375–380. doi: 10.1109/NCM.2008.215.

[201] A. Sawsaa and J. Lu, “Developing a domain ontology of Information Science (OIS),” presented at the International Conference on Information Society (i-Society 2012), London, UK: IEEE, Jun. 2012, pp. 447–452. [Online]. Available: https://ieeexplore.ieee.org/document/6285013

[202] Z. Zakaria, S. Kasim, N. Husna Mohamad Hasbullah, A. Alyani Azadin, A. Saleh Ahmar, and R. Hidayat, “The Development of Personality Ontology Based on the Methontology Approach,” *Int. J. Eng. Technol.*, vol. 7, no. 2.5, p. 73, Mar. 2018, doi: 10.14419/ijet.v7i2.5.13955.

[203] B. Conrad *et al.*, “OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax (Second Edition),” W3C Recommendation. [Online]. Available: https://www.w3.org/TR/owl2-syntax/#Axioms

[204] S. H. Kassani and P. H. Kassani, “Building an Ontology for the Domain of Plant Science using Protégé,” 2018, *arXiv*. doi: 10.48550/ARXIV.1810.04606.

[205] D. Maynard, W. Peters, and Y. Li, “Evaluating Evaluation Metrics for Ontology-Based Applications: Infinite Reflection,” in *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC’08)*, Marrakech, Morocco: European Language Resources Association, May 2008. [Online]. Available: http://www.lrec-conf.org/proceedings/lrec2008/pdf/273_paper.pdf

[206] M. Amith, Z. He, J. Bian, J. A. Lossio-Ventura, and C. Tao, “Assessing the practice of biomedical ontology evaluation: Gaps and opportunities,” *J. Biomed. Inform.*, vol. 80, pp. 1–13, Apr. 2018, doi: 10.1016/j.jbi.2018.02.010.

[207] M. McDaniel and V. C. Storey, “Evaluating Domain Ontologies: Clarification, Classification, and Challenges,” *ACM Comput. Surv.*, vol. 52, no. 4, pp. 1–44, Jul. 2020, doi: 10.1145/3329124.

[208] R. Achim and K. Sandkuhl, “NEOntometrics: A Flexible and Scalable Software for Calculating Ontology Metrics,” in *Proceedings of Poster and Demo Track and Workshop Track*, Vienna,

Austria: CEUR-WS, Sep. 2022, pp. 1–5. [Online]. Available: https://ceur-ws.org/Vol-3235/paper16.pdf

[209] S. Tartir, I. B. Arpinar, and A. P. Sheth, "Ontological Evaluation and Validation," in *Theory and Applications of Ontology: Computer Applications*, R. Poli, M. Healy, and A. Kameas, Eds., Dordrecht: Springer Netherlands, 2010, pp. 115–130. doi: 10.1007/978-90-481-8847-5_5.

[210] J. Raad and C. Cruz, "A Survey on Ontology Evaluation Methods:," in *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, Lisbon, Portugal: SCITEPRESS - Science and and Technology Publications, 2015, pp. 179–186. doi: 10.5220/0005591001790186.

[211] A. Rebekka, K. Sven, Birger Lantow, and S. Kurt, "Applicability of Quality Metrics for Ontologies on Ontology Design Patterns:," in *Proceedings of the International Conference on Knowledge Engineering and Ontology Development*, Vilamoura, Algarve, Portugal: SCITEPRESS - Science and and Technology Publications, 2013, pp. 48–57. doi: 10.5220/0004541400480057.

[212] R. Crawford and D. Fernanda, "Animal Health Ontology (AHO)," Data-driven Surveillance. [Online]. Available: http://datadrivensurveillance.org/animal-health-ontology-aho/

[213] R. Dividino, M. Romanelli, and D. Sonntag, "Semiotic-based Ontology Evaluation Tool (S-OntoEval)," in *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco: European Language Resources Association (ELRA), May 2008. [Online]. Available: https://aclanthology.org/L08-1398/

[214] A. Gangemi, C. Catenacci, M. Ciaramita, and J. Lehmann, "A theoretical framework for ontology evaluation and validation," CEUR-WS, 2005, pp. 1–16. [Online]. Available: https://www.ceur-ws.org/Vol-166/9.pdf

[215] R. Lourdusamy* and A. John, "Metric Based Ontology Quality Evaluation," *Int. J. Eng. Adv. Technol.*, vol. 8, no. 6, pp. 3072–3077, Aug. 2019, doi: 10.35940/ijeat.F8679.088619.

[216] S. Tartir, I. Arpinar, M. Moore, A. Sheth, and B. Aleman-Meza, "OntoQA: Metric-Based Ontology Quality Analysis," in *Distributed, Autonomous, Semantically Heterogeneous Data and Knowledge Sources*, Houston, TX: IEEE, Nov. 2005. [Online]. Available: https://corescholar.libraries.wright.edu/knoesis/660

[217] Z. C. Khan, "Evaluation metrics in ontology modules," in *Proceedings of the 29th International Workshop on Description Logics*, Cape Town, South Africa: CEUR Workshop Proceedings, Apr. 2016, pp. 1–13. [Online]. Available: https://ceur-ws.org/Vol-1577/

[218] R. Lourdusamy and A. John, "A review on metrics for ontology evaluation," in *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, Coimbatore: IEEE, Jan. 2018, pp. 1415–1421. doi: 10.1109/ICISC.2018.8399041.

[219] B. Lantow, "OntoMetrics: Application of on-line Ontology Metric Calculation," in *Joint Proceedings of the BIR 2016 Workshops and Doctoral Consortium*, Prague, Czech Republic: CEUR Workshop Proceedings, Sep. 2012, pp. 1–12. [Online]. Available: https://ceur-ws.org/Vol-1684/

[220] W. Jarosław, "An Attempt to Knowledge Conceptualization of Methods and Tools Supporting Ontology Evaluation Process," *Procedia Comput. Sci.*, vol. 126, pp. 2238–2247, 2018, doi: 10.1016/j.procs.2018.07.225.

[221] M. Poppe and M. Lichtwark, "Ontometrics." Ontometrics, 2024. [Online]. Available: https://ontometrics.informatik.uni-rostock.de/ontologymetrics/

[222] B. Gobin-Rahimbux, "Evaluation Metrics for Ontology Modules," in *2022 IEEE International Conference on Data Science and Information System (ICDSIS)*, Hassan, India: IEEE, Jul. 2022, pp. 1–6. doi: 10.1109/ICDSIS55133.2022.9915950.

[223] A. Burton-Jones, V. C. Storey, V. Sugumaran, and P. Ahluwalia, "A semiotic metrics suite for assessing the quality of ontologies," *Data Knowl. Eng.*, vol. 55, no. 1, pp. 84–102, Oct. 2005, doi: 10.1016/j.datak.2004.11.010.

[224] M. A. Sicilia, D. Rodríguez, E. García-Barriocanal, and S. Sánchez-Alonso, “Empirical findings on ontology metrics,” *Expert Syst. Appl.*, vol. 39, no. 8, pp. 6706–6711, Jun. 2012, doi: 10.1016/j.eswa.2011.11.094.

[225] S. Swenja, N. Gerst, C. Keller, and S. Thomas, “Defining a Context Model for Smart Manufacturing,” *Procedia Comput. Sci.*, vol. 204, pp. 22–29, 2022, doi: 10.1016/j.procs.2022.08.003.

[226] A. Chatterjee, A. Prinz, M. Gerdes, and S. Martinez, “An Automatic Ontology-Based Approach to Support Logical Representation of Observable and Measurable Data for Healthy Lifestyle Management: Proof-of-Concept Study,” *J. Med. Internet Res.*, vol. 23, no. 4, p. e24656, Apr. 2021, doi: 10.2196/24656.

[227] X. Yang, M. Yang, D. Yang, and Y. Huang, “Research on Implementation of Knowledge Convergence Based on Apache Jena3,” in *Proceedings of the 2018 International Conference on Computer Science, Electronics and Communication Engineering (CSECE 2018)*, Sanya, China: Atlantis Press, 2018. doi: 10.2991/csece-18.2018.101.

[228] R. Bansal and S. Chawla, “Evaluation Metrics for Computer Science Domain Specific Ontology in Semantic Web Based IRSCSD System,” *Int. J. Comput. IJC*, vol. 19, no. 1, pp. 129–139, 2015, [Online]. Available: https://ijcjournal.org/index.php/InternationalJournalOfComputer/article/view/497

[229] T. Sigwele *et al.*, “Building a Semantic RESTFul API for Achieving Interoperability between a Pharmacist and a Doctor using JENA and FUSEKI,” *J. Phys. Conf. Ser.*, vol. 1376, no. 1, p. 012015, Nov. 2019, doi: 10.1088/1742-6596/1376/1/012015.