Development of a Domain Ontology to Support Information Retrieval on the South African Informal Sector Services

by

Glodi Mokombati Atoba

Thesis submitted in fulfilment of the requirements for the degree

Master of Technology: Information Technology

in the Faculty of Informatics and Design

at the Cape Peninsula University of Technology

Supervisor: Prof Justine Olawande Daramola

Cape Town
Date submitted: December 2022

DECLARATION

I, Glodi Mokombati Atoba, declare that the contents of this dissertation/thesis represent my own unaided work and that the dissertation/thesis has not previously been submitted for academic examination towards any qualification. Furthermore, it represents my own opinions and not necessarily those of the Cape Peninsula University of Technology.

16/02/2023

Signed **Date**

ABSTRACT

The high unemployment rate in South Africa has encouraged several types of informal occupational and business activities in the informal business sector. Like in many Sub-Saharan Africa (SSA) countries, the South African informal business sector (SAIBUS) has many challenges to contend with, such as a lack of organisation, poor regulation, poor service delivery, low capitalisation, and low productivity. Attempts have been made to address some of these challenges by applying ICTs. One proposed solution is using web portal technology to promote informal businesses and facilitate informal business transactions. However, to attain maximum impact, the platform must be enabled with Artificial Intelligence (AI) to create a system that can leverage the convergence of intelligent reasoning and semantic technology in tackling the challenges of the South African informal business sector. In essence, such a web portal must possess capabilities such as semantic-based reasoning, semantic analysis of user queries, and intelligent recommendations, which require the existence of a dedicated domain ontology. Therefore, this study aimed to develop a domain ontology that can support information retrieval and promote intelligent reasoning on services in the informal business sector in South Africa.

Based on the study's aim, the objectives were: (1) to determine the requirements for developing a domain ontology for the South African informal business sector; (2) To design and develop a domain ontology for the South African informal business sector; (3) To evaluate the quality attributes of a domain ontology for the South African informal business. Design Science Research (DSR) was employed as a research strategy to accomplish the objectives of this research. The research design mapped the phases of DSR to the Lean Ontology Development (LOD) method to realise the phases of problem awareness, design, development, and evaluation. While reviewing some use cases from the South African informal business sector, we used competency questions to identify the requirements for the South African informal business sector (SAIBUS). The LOD method was adopted as the ontology engineering methodology. A combination of ontology modularisation and design patterns were used as concepts to realise the design of the SAIBUS ontology, while the Protégé ontology editor was used to develop the ontology. The evaluation of the ontology was divided into two phases – validation and verification. The validation of the SAIBUS ontology (assessing its quality and correctness) was done using the web-based tool - OOPS! and the HermiT ontology reasoner. The SAIBUS ontology was verified by comparing it against the ontology requirements and competency questions to assess whether the ontology is built to fit its purpose. This was done by attempting to answer some of the competency questions using the SPARQL query to test its ability to retrieve relevant ontological instances (individuals). The results from the evaluation show that the SAIBUS ontology can support information retrieval on the South African informal sector services.

Keywords: Informal sector, Information retrieval, domain ontology, Knowledge representation, Modular ontology, Ontology engineering, Artificial intelligence, South African informal business.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to those who have supported me throughout the process of completing this study.

First and foremost, I would like to give thanks to God, in the name of Jesus Christ, for His grace and mercy, for giving me the strength to persevere and for being sufficient in my life to allow me to finish this study.

Next, I would like to extend my heartfelt gratitude to my supervisor, Professor Justine Olawande Daramola. He has been a constant source of support, encouragement, and mentorship throughout this journey. His patience, consideration, guidance, expertise, commitment, advice, and mentorship have been invaluable to me. I am deeply grateful to this exceptional professor with so much knowledge and experience.

I would also like to express my gratitude to my parents, Mr Atoba Ekomba Jacques and Mrs Christine Minimbu. Their love, generosity, prayers, and blessings have been a constant source of strength and motivation throughout my life and academic career.

My wife, Deborah Lupasa Mafika, deserves a special mention for her unwavering support and encouragement. She has been my partner in every sense of the word, studying alongside me for long hours and being there for me through late nights. I am incredibly grateful for her love and prayers.

I would also like to thank my extended family, including Papa Dominique Minimbu, Mama Dede Mbombo, and Mama Clotilde Minimbu for the prayers, support, and encouragement throughout this journey. I also extend my appreciation to Papa Georges Lupasa and Mama Claude Lupasa, for their prayers and life advice.

I am grateful to my family and friends, who have supported me in one way or another through this journey with their prayers, encouragement, and support. In particular, I would like to thank Nancy Minimbu, Nicolette Minimbu, Leaticia Minimbu, Scarlette Minimbu, Pamela Safari, Tresor Ngonzi, Medie Minimbu, Fortuna Minimbu, Reagan Kazadi, Florence Kazadi Mwamba, Shekinah Lupasa, Jordy Makosso, Francis Bosenge, Arnold Kabuiku, Felly Itota, Jean Luc Loutchia, and Malikah Osman. Lastly, I would like to express my gratitude to my family at church for their prayers and shared belief in a positive outcome.

DEDICATION

This thesis is dedicated to my dearest mother, Christine Minimbu, and my cherished wife, Deborah Lupasa Mafika, for their selfless love and steadfast support throughout my academic journey. I am eternally grateful for the constant encouragement and love that they have provided me. Their unwavering belief in me has been the driving force behind my success, and I cannot thank them enough. I dedicate this work to my mother and my wife with all my love and admiration.

PUBLICATION FROM THE THESIS

Published papers

Atoba, G., & Daramola, O. (2023). Design of a Modular Ontology for South African Informal Sector Services. In Proceedings of Seventh International Congress on Information and Communication Technology (pp. 85-98). Springer, Singapore.

Paper in Preparation

SABIBUS: A Domain Ontology for the South African Informal Sector

Table of Contents

List of Figures

List of Tables

GLOSSARY

Terms/Acronyms/Abbreviations	Definition/Explanation
Axioms	Information about classes, attributes, and individuals may be encoded using OWL axioms. These cover declaring the existence of certain kinds of facts as well as subclass hierarchy, property domains, and ranges.
CoModIDE	Comprehensive Modular Ontology IDE
H&B	Health and Beauty
ICLS	International Conference of Labour Statisticians
ICTs	Information Communication Technologies
IEEE	Institute of Electrical and Electronics Engineers
ILO	International Labour Organization
Individuals	Includes concrete objects such as animals, tables, cars, people, molecules and planets, as well as abstract things like words and numbers.
IR	Information Retrieval
LOD	Learn Ontology Development
NDP	National Development Plan
Object property	Links two individuals with a predicate
ODP	Ontology Design Pattern
OOPS!	OntOlogy Pitfall Scanner!
OWL	Web Ontology Language - a computational logic-based language that enables computer programs to use information provided in OWL
RDF	Resource Description Framework - initially intended to serve as a data model for metadata,
SAIBUS	South African Informal Business Sector
SPARQL	SPARQL Protocol and RDF Query Language - an RDF query language that can access and operate with data stored in Resource Description Framework format.
TESSISA	Technology Support for the Informal Sector of South Africa
VOWL	Visual Notation for OWL Ontologies
XML	Extensible Markup Language

Data property	Offers a relation to link an instance of an entity to a literal datatype value that measures or approximates the subject matter of the data property (for example, an RDF integer, text, or date).
Class	Concepts that are also called type, sort, category, and kind
Entity	A representation of an object or a thing

CHAPTER ONE
INTRODUCTION

1.1 Introduction

Many people in developing countries are involved in informal occupational activities (Mramba et al., 2015). The informal sector is the commercial operation partly or completely independent of policy supervision, taxes, and observation (Mramba et al., 2015; Daramola, 2018). These activities are often invisible, irregular, unstructured, parallel, underground, or survivalist (Williams and Horodnic, 2016a). The informal sector can help many developing countries grow and evolve (Rogan and Skinner, 2017). Informal business workers are commonly referred to as street vendors. They sell products of many kinds, such as fruits, foods, drinks, clothes, electronic devices, appliances, and other goods. Moreover, they also provide services of different kinds, such as haircuts, plumbing, and beauty services (Mramba et al., 2015). Thus far, the impact of the informal sector on the republic of South Africa's economy remains comparatively small compared with other African countries. The impact of the South African informal business sector is substantial for the creation of employment, job retention, and as a means of poverty alleviation (Rogan and Skinner, 2017). In keeping with Rogan and Skinner (2017), over 2.5 million operators contribute to the South African informal sector, yet these people require support for accessibility and information exchange (Jiyane *et al.*, 2013).

A study by Jiyane et al. (2013) recommended using Information and Communication Technologies (ICTs) for the informal sector to support information exchange. The authors observed that South Africa remains in the phase of being an information-knowledge country and that reaching the expectations of an information-and-knowledge country will help the nation develop those disadvantages in the informal area. They further mentioned that in many activities enabled by technology and ICTs, ladies within the informal sector might also play a significant role in the benefit of their informal enterprises to facilitate their contact with vendors and provide them with the ability to obtain certain business resources domestically as well as abroad. Also, it will improve the economic cycle, jobs, and commerce (Jiyane *et al.*, 2013).

A project named Technology Support for the Informal Sector of South Africa (TESISSA) was proposed to tackle some challenges experienced within the informal business (Daramola, 2018). The TESISSA project aimed at creating a web/mobile portal that combines a smart cloud-based system and semantic knowledge. The goal is to give more attention to the challenges of South Africa's informal sector to improve economic participation in the South African informal market (Daramola, 2018). However, to achieve this goal, there is a need to develop a knowledgebase infrastructure for the portal in the form of a domain ontology to support user search operations and facilitate intelligent reasoning.

Ontology is a concept native to philosophy, described as a structured interpretation of existence (Gruber, 1993; Keet, 2020). This is also known as defining the kinds of objects that may exist and their relationships with each other (Li and Larsen, 2017). The concept of ontology is broadly used nowadays for knowledge engineering, computing, information retrieval, engineering, and the semantic web (Keet, 2020). Hence, the definition of ontology is similar in software engineering and computer science but more precise and concrete. A domain ontology is used to describe terms, their meanings, and the associations between them as they are concerned with the dominant concept in a particular domain (Yun *et al.*, 2009). Saad and Shaharin (2016) explain knowledge representation as an important component in providing intelligent systems-supported knowledge as a key component for empowering the reasoning method and higher cognitive process. Thus, this study develops a knowledgebase e artefact that can aid an online portal that will enable intelligent reasoning and information retrieval on the South African Informal sector services.

1.2 Background

The informal sector contains various activities involving many informal business providers. These providers are frequently mobile and travel from place to place to offer their goods and services to potential customers. They usually carry their products by hand in bags, on the shoulders, on trolleys. However, other informal business workers are fixed based more on the setting and sell their goods or provide their services from the same place (Mramba et al., 2015). These informal activities are undertaken for different reasons. A study by Legodi and Kanjere (2015) stated that informal employment in South Africa has increased due to a lack of job opportunities. Moreover, they mentioned in their study that street vendors generally join informal businesses due to their needs or obligation rather than an opportunity. This may be linked to the poor level of competence needed to take up a meaningful job, but a small proportion still has some competence or skill (Legodi and Kanjere, 2015). Similarly, Ndabeni and Maharajh (2013) reported that both the government and companies could not provide adequate employment for the rising workforce. Hence the unofficial sector is increasingly recognised as a possible replacement, especially among the young and the underprivileged. Thus, there is a need to improve the performance of the informal sector potentially. Yet, informal business providers experience many challenges that impact the operation of their businesses. According to studies by Legodi and Kanjere (2015) and Nel and Rogerson (2016), these challenges include a lack of support from local authorities, difficulties in recording transactions and entering the national tax net, low quality of operations, inadequate facilities and access to electricity, absence of service quality, regulatory issues, limited financial support, and a lack of expertise and access to Information and Communication Technologies. These challenges pose a significant barrier to the growth and success of the informal sector in South Africa.

Even though some classified websites in South Africa (such as GUMTREE, LOCANTO, OLX, etc.) and social media platforms (Facebook, Instagram, etc.) are used to promote informal business activities between consumers and service providers among the population, they are not primarily designed to support informal businesses. This ensures that these challenges remain (Daramola, 2018). For example, interactions on these classified websites are done predominantly in the English language, which is unsuitable for most practitioners in the informal sector who are not educated. Also, many nuanced aspects and issues peculiar to the informal sector are not captured/addressed by these websites (Daramola, 2018). According to Etim and Daramola (2020), a domain ontology which is a model of knowledge representation for a specific domain can be a basis for generating knowledgebase information services that pertain to that domain. Thus, a dedicated web portal for the South African informal sector that is enabled by an underlying domain ontology for the South African informal sector will offer much more than any other conventional information service platform.

In recent years, ontology has been recognised as an important concept for studies in various fields, such as machine learning, Robotics, and Internet of things (IoT) technology, due to the ability to provide intelligent reasoning and information exchange between separate systems (McDaniel and Storey, 2019). Therefore, a domain ontology is needed to support the multimodal portals with intelligent reasoning to facilitate easy and fast retrieval of information on the platform. This will enable easy and fast retrieval of real-time information that can be shared and reused by different automated systems. Furthermore, the content of the domain ontology can also be referenced by different automated systems that need to analyse or process the information on the informal sector. In addition, many applications nowadays require a knowledgebase ontology to support intelligent reasoning due to the increasing number of systems requiring access to real-world information knowledge (McDaniel and Storey, 2019).

This study aimed to develop a domain ontology for the South African informal business Sector that can be shared and reused to support the multimodal framework to improve operators' profitability, enable wealth generation, and facilitate a quasi-formalisation of the informal business sector in South Africa.

1.3 Research Problem

The application of ICTs has been proposed to handle some of the challenges identified in the informal sector of South Africa (Nel and Rogerson, 2016; Daramola, 2018) by developing a viable web/mobile portal (information system platform) that can help to give visibility to the activities of the informal sector of South Africa. There is, however, a lack of knowledgebase to properly equip this web portal with what is needed to build a system that facilitates the convergence of intelligent thinking and semantic technology, which brings with it the need to establish a domain ontology capable of supporting intelligence resources such as service listing and semantic search.

There is currently no domain ontology or knowledgebase that can support search and information retrieval, intelligent reasoning, and semantic analysis of data on informal business services in the South African informal sector. Consequently, this results in a lack of adequate usable information on informal sector services and the quality of services. This affects informal business operators in the South African business sectors, their customers, and the government.

1.4 Aim, Objectives, and Research Questions

1.4.1 Research aim

The study aimed to create a domain ontology that can support information retrieval and promote intelligent reasoning on services in the informal business sector in South Africa.

1.4.2 Research Objectives

The following objectives helped to realise the aim of this study:

1. To determine the requirements for developing a domain ontology for the South African informal business sector.
2. To design and develop a domain ontology for the South African informal business sector.
3. To evaluate the quality attributes of a domain ontology for the South African informal business.

1.4.3 Research Questions

The main research question for this study is, how can a domain ontology that facilitates access to intelligent information services in the South African informal business sector be developed?

1. What are the requirements for developing a domain ontology for the South African informal business sector?
2. How can a domain ontology be developed to facilitate access to intelligent information services in the South African informal business sector?
3. What are the quality attributes of a domain ontology for the South African informal business sector?

1.5 Ethical Consideration

To address the ethical concerns that pertain to the proposed study, the following was considered:

Informed Consent Participant: All participants in this study could decide to stop participating at any time. The consent form was given to participants before the collection of data.

Data privacy and confidentiality: Throughout this research, no personal details such as date of birth, political affiliation, religious ideology, fitness, or sexual orientation were obtained. Data such as informal sector services, the business profile, location, and pricing were collected.

Dissemination of research: Since this research was intended to create a domain ontology that can be reused and shared, there is the potential for results in the popular and academic press.

1.6 Delineation of the Study

Given the large size of the informal sector in the different regions of South Africa and the limited timeframe for this study, an ontology for the informal sector services in South Africa could only be created for some of the informal sector services that are prominent in the Cape Town metropolis. The informal sector of Cape Town was used as the case study for this research, but the informal sector services and categories are quite similar to those found in other regions of South Africa. However, certain categories of services such as health and beauty, event services, construction services, automotive services, and transportation services may differ based on the peculiar characteristics of other regions.

1.7 Significance of the Study

The results of this study will have a significant impact, through the domain ontology that was developed in a knowledgebase form to support intellectual reasoning, on empowering web portal platforms that will improve the capabilities of the informal economy sector and organisations in this area. The operators in this sector, ontology users, developers, and the government will also benefit from this by facilitating information retrieval on informal sector services and helping to enhance the regulation and quality of the informal market.

1.8 Thesis Structure

This thesis is structured into seven comprehensive chapters, each of which covers a distinct aspect of the study. Chapter one introduces the research topic and sets the scene for the study. Chapter two provides a comprehensive literature review of the relevant research, highlighting the gaps in knowledge and identifying the study's unique contribution. Chapter three details the research methodology used in this study, explaining the research design, data collection, and analysis methods. In chapter four, the study's requirements and design are presented, providing a clear overview of the project's scope and objectives. Chapter five provides a detailed account of the ontology's development, discussing the design process and key features. Chapter six presents the study's evaluation, highlighting the findings and their implications. Finally, chapter seven concludes the thesis, summarizing the study's main findings, outlining the key recommendations, and identifying future research directions.

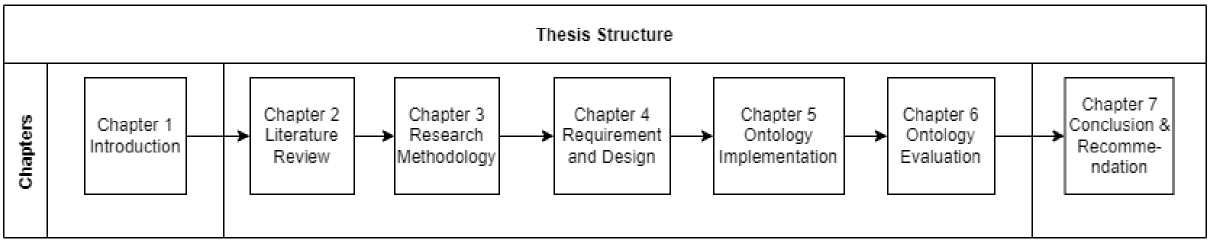

Figure 1.1: Thesis Structure

CHAPTER TWO
LITERATURE REVIEW

This chapter reviews the literature on the study's major areas. The literature review was focused on the key areas of the study based on the aim, which was to develop a domain ontology that can support information retrieval and promote intelligent reasoning on services in the informal business sector in South Africa. The informal sector in general; the informal business sector in South Africa in particular; informal retrieval; domain ontology concept; ontology requirements; ontology development methodologies; Modularisation of ontology; ontology methods and tools; and ontology evaluation methods are among the key areas examined. The chapter also reviews other researchers' work that is relevant to the study.

2.1 Informal Sector

The International Labour Organization (ILO) defines the informal economy as "employment in the informal sector" and "informal employment" that is not legally or practically regulated by formal rules and regulations (International Labour Organisation, 2003). It is characterized by forms of operation that are different from those of formal companies or enterprises, which are the most recognizable units of production in economic theory. Informal sector practices tend to be unregulated by the government and not included in regular national income accounts (International Labour Organisation, 2003; Maluleke, 2019).

In its 1972 Kenya report, the ILO described the informal sector as activities of hardworking individuals who are "not recognized, documented, protected, or controlled by the state authorities." These activities have the following characteristics: ease of entry, dependence on local resources, small scale operations, labour-intensive and adaptable technology, expertise learned outside of conventional education systems, and uncontrolled and competitive markets (International Labour Organization, 2002). Since then, the concept of the informal sector has expanded and evolved. At the 15th International Conference of Labour Statisticians (ICLS) in 1993, it was decided that production units or firms should define the informal sector for statistical purposes, rather than employment connections.

The ICLS definition also suggests that informal sector businesses be classified based on one or more of the following criteria:

(1) Non-registration with the government for tax purposes,

(2) Non-registration of employees in accordance with labour laws,

(3) Small company size in terms of personnel.

Informal sector businesses sell a wide range of goods, such as fruits, foods, drinks, clothes, electronic devices, appliances, and other products, as well as provide various services, such as haircuts, plumbing, and beauty treatments (Mramba et al., 2015). Despite the widespread presence of the informal sector, the scale, structure, industry, and employment of its activities are still not well-documented (Benjamin et al., 2014). The relationship between the informal economy and the formal economy, as well as the overall economy, is crucial for job creation, income growth, poverty reduction, and reducing inequality (Maluleke, 2019).

The informal sector is a major component of economic development and accounts for 60% of the world's workforce (Maluleke, 2019). Many people in developing countries participate in informal economic activities (Mramba et al., 2015). The informal sector has global significance and can help many developed countries grow and evolve (Rogan and Skinner, 2017). While its contribution to the overall economy varies by region, it plays a significant role in the economic development of all World Bank Group client countries, according to some experts (Benjamin et al., 2014). However, the informal sector creates a significant information gap as, by necessity, all or some aspects of informal economic activity are off the official record or not formally recorded.

2.2 South African Informal Business Sector

The National Development Plan (NDP) of South Africa in 2012 emphasizes the importance of the informal sector in national development, recognizing that it plays a crucial role in tackling unemployment, poverty, and inequality. The informal sector contributed 6% to South Africa's GDP in 2017 and 18.3% of employed individuals were employed in the informal sector according to the 2019 Quarterly Labour Force Survey. The growth of the informal sector is influenced by factors such as limited access to formal sector opportunities, increasing demand for low-cost goods, and slow economic development.

However, informal business providers face challenges such as a lack of local municipality support, difficulties in record-keeping, low-quality operations, limited access to electricity, and more (Daramola, 2018). To address these challenges, the National Informal Business Upliftment Plan (NIBUS) was established in 2014 to provide support and resources for entrepreneurs.

In South Africa, the informal business sector is a significant contributor to the country's economy, providing livelihoods and employment opportunities to a significant portion of the population. Despite its importance, the informal business sector remains largely unorganized and often operates outside the formal economic structure, making it difficult to understand and quantify its contribution to the economy.

The classification of the informal business sector into specific categories is crucial in understanding its various components and how they contribute to the economy. In the context of this study, the researcher focused on the informal service sector and identified five categories of services offered by informal businesses in South Africa: health and beauty, event services, construction services, automotive services, and transportation services.

Research conducted by the National Regulator for Compulsory Specifications (NRCS, 2018) and other relevant organizations highlights the growth and high demand in various sectors within the informal business sector in South Africa, including the health and beauty sector, event services sector, construction sector, automotive services sector, and transportation sector. Hence, the classification of the informal business sector into specific categories provides a better understanding of its various components and how they contribute to the economy. In this study, this information was used to construct a minimum viable ontology for the informal South African business sector in order to provide a valuable tool for understanding and quantifying the contribution of this sector to the economy.

The use of ICTs is seen as critical for the development and success of small and informal businesses. These technologies have the potential to improve productivity, provide cost savings, and improve customer response times. Given their affordability and widespread availability, it is important to assess their impact on African businesses, especially informal services.

To this end, several initiatives have been launched to promote the use of ICTs in the South African informal sector. Jiyane et al. (2013) proposed using ICTs to improve information sharing within the informal sector. They also noted that women in the informal sector could play a significant role in utilizing technology and ICTs to enhance their businesses by facilitating vendor interaction and accessing business resources domestically and internationally. The adoption of ICTs could lead to positive outcomes for the economy, employment, and trade. Daramola (2018) introduced a project called Technology Support for the Informal Sector of South Africa (TESISSA), aimed at addressing the challenges faced by the informal sector in South Africa. The TESISSA study aimed to create a web/mobile platform that combines cloud-based technology with semantic information to support the informal sector.

The literature review of the South African informal business sector highlights the significance of various factors in the success of informal businesses. By reviewing the literature, the researcher considered six dimensions as important characteristics of peer-to-peer interactions in the informal sector namely, the characteristics of the services being provided, service quality rating, price of the service, possible modes of payment for the service, customer profile, and service provider profile. Characteristics of the services being provided are crucial in determining the success of informal businesses. This dimension encompasses the type of services offered, their uniqueness, and their relevance to the target market (Jiyane et al., 2013). The quality of services offered is another significant factor, as customers are more likely to return and recommend businesses that offer high-quality services (Rogan & Skinner, 2017).

The price of the service is also a critical determinant of success in the informal sector, as many informal businesses serve low-income customers who are sensitive to the cost of goods and services (Maluleke, 2019). Possible modes of payment for the service include cash, mobile payments, and other electronic payment methods, and the availability of different payment options can impact the success of informal businesses (Ndabeni & Maharajh, 2013).

The customer profile and the service provider profile are also essential in the success of informal businesses. The customer profile includes factors such as age, gender, education level, and income, which can affect their demand for goods and services. On the other hand, the service provider profile includes factors such as education level, experience, and entrepreneurial skills, which can impact the quality of services provided (Legodi & Kanjere, 2015).

Thus, the six dimensions: characteristics of the services being provided, service quality rating, price of the service, possible modes of payment for the service, customer profile, and service provider profile, are important considerations in understanding the dynamics of the SAIBUS ontology.

2.3 Information Retrieval

Information Retrieval (IR) was used in business and intelligence applications in the 1960s before search engines became extensively used in everyday life. IR capabilities expanded with advancements in processing speed and storage capacity as they developed with many other computer technologies. The fast transition from traditional library-based methods to obtaining, indexing, and searching information to more automated methods is reflected in the development of such systems (Sanderson and Croft, 2012).

Information retrieval is the process of extracting useful data from large data sets where an information retrieval (IR) system finds relevant information to a user's query. In most cases, an IR system searches collections of unstructured or semi-structured data (such as web pages, documents, pictures, videos, and so on) (Sanderson and Croft, 2012). When a collection grows to a scale where traditional cataloguing approaches can no longer handle it, an IR system is required. Traditional information retrieval approaches cannot produce high-quality search results when data accumulates, and the demand for high-quality retrieval results rises (Yu, 2019).

With the rise of digitised unstructured information and quick worldwide access to large amounts of that information over high-speed networks, search became the only feasible option for locating relevant things from these large text databases, and IR systems became widespread.

There is a lot of data available on the internet. It has many web pages, making it difficult to identify the right information among them. Data must be formally organised for users to access and utilise them. The Semantic Web is a collection of data connected in a way that makes it simple for machines to process (Singh Narula, Singh and Jain, 2014). Technology for information retrieval has been essential to the Web's success. Google and Yahoo, among other web-based indexing and search engines, have significantly altered how we access information. Semantic web technologies must be integrated with Web search engines and information retrieval technology in general to make an impact (Finin et al., 2005; SIngh Narula, Singh and Jain, 2014). Meta search engines, which use machine learning, are the ancestors of intelligent search engines. Knowledge can be marked on a page in a way that allows automatic systems to extract and comprehend it, enabling intelligent information services, customised websites, and search engines with stronger semantic capabilities. Software agents can comprehend knowledge marked up and further derive conclusions pertaining to the domain of interest, thanks to ontologies (Shah et al., 2002; Singh Narula, Singh and Jain, 2014).

2.4 Domain ontology

Before starting the process of developing domain ontology, one has to understand the theory behind the idea of domain ontology. Ontology was earlier limited to the field of philosophy in relation to the discipline of existence (Gruber, 1993; Keet, 2020). The definition of ontology was acquired from philosophy; after that, ontology research has been progressively dominant in the software engineering culture and has now attained a particular role in the theory of database design and computational intelligence(Saad and Shaharin, 2016). In computer science, Ontology is characterised as the standard representation of information, a collection of principles in the domain, and the relationship between those concepts (Nicola, Daniel and Steffen, 2009). Ontology was originally defined as an explicit conceptualisation specification (Gruber, 1993), which includes ideas, associations, and axioms, in which concepts reflect "objects" in the field of interest, relationships create inter-concept associations, and the axioms state the definitions of the concepts and relationships as well as their behavioural restrictions (Nicola, Daniel and Steffen, 2009; Keet, 2020) Thus, the conceptualisation is based on the body of formally represented knowledge.

Conceptualisation is also characterised as a clear worldview we wish to portray for some reason. Also, any knowledgebase, knowledge-based framework, or information-level agent is specifically or indirectly committed to some conceptualisation (Gruber, 1993; Nicola, Daniel and Steffen, 2009; Saad and Shaharin, 2016). A more expressive definition of ontology by (Fikes and Finin, 1991; Keet, 2020) summarises the guidelines on the elements to be considered in the construction of ontology by identifying the primary terms, the relationship between those terms, the definition of the terms and the information to be inferred from them. Gruber (1993) noted that what "exists" for Artificial Intelligence (AI) systems is what can be interpreted. Hence, the concept of ontology allows the author to describe the problem entity, its attributes and the relationships within the informal business domain of South Africa.

Much research employed ontology to characterise empirical knowledge in an organised way to assist in comprehending the development of domain knowledge. For example, in a system ontology, an AI perspective can be defined by describing a set of conceptual terms associating a universe of discourse such as categories, associations, features, or various objects with human-readable textual content explaining the significance of terms and formal axioms that limit the understanding and well-formed usage of those terms (Gruber, 1993; Saad and Shaharin, 2016).

Ontology has several applications in different fields, such as knowledge computing, databases, and digital development information modelling, and is commonly utilised for numerous uses, including natural language analysis, knowledge management, semantic web, intellectual communication intelligence and others (McDaniel and Storey, 2019). One of the main benefits of ontology is its reusability, which ensures that the same ontology may be used to construct other knowledgebase structures that have the same skeleton (Saad and Shaharin, 2016).

In this study, this perception of ontology helps one structure domain knowledge and formally represent the knowledge in a more precise, explicit, consensus-based, and meaningful way.

2.5 Ontology Requirements

Identifying requirements is one of the key components in the development of ontology. The process of gathering the requirements that the ontology must satisfy is called ontology requirements specification. The ontology requirements specification activity's purpose is to define why it is being constructed, what its targeted uses are, who the end-users are, and what requirements the ontology should meet (Suárez-Figueroa, Gómez-Pérez and Villazón-Terrazas, 2009). To make ontology development easier, many scholars recommend using the Competency questions (CQs) as an important stage in the ontology development life cycle (Grüninger and Fox, 1995).

2.5.1 Competency Question

Competency questions (CQs) are interrogatives that help define the scope of an ontology (Grüninger and Fox, 1995; Williams and Horodnic, 2016a). They are plain language sentences that indicate a pattern for the types of questions that people expect an ontology to answer (Usher, Gudes and Parekh, 2016). In addition, they serve as a basis for establishing the use cases and the requirements that an ontology must satisfy. Thus, the concepts of an ontology and the defined relationships amongst them should enable the ontology to respond accurately to user queries. As a result, the answerability of CQs becomes a functional requirement of the ontology. Therefore, to answer and solve these questions, the ontology must have a necessary and sufficient set of axioms. These questions characterise current ontologies and drive the building of new ontologies needed to answer the CQs (Grüninger and Fox, 1995). Freitas, Santana da Silva and Bezerra (2013) agreed that CQs are crucial in the ontology life cycle because they define the requirements of the ontology being developed, much as software requirements are crucial in software development. They further mentioned that CQs also assess ontology's expressiveness. Developers should be able to generate a series of questions that represent users' requirements, and they will be able to define its scope given a set of scenarios about a domain of discourse. CQs can aid in two ways: first, enabling developers to define the major parts and their relationships to establish the ontology vocabulary, and second, providing developers with a simple way of verifying requirements' satisfiability through knowledge retrieval or entailment on its axioms and answers (Freitas et al., 2013).

Generating competency questions demands a grasp of the area of interest, representing the ontology. Gathering data is an important step in making this happen, especially if one does not have a thorough understanding of the subject matter for which the ontology is being built. For example, Gruninger and Fox (1995) used CQs to describe the ontologies established for the advisors in the TOVE ontology and to give recommendations for the design of additional ontologies. They went on to say that these questions might serve as standards for any ontology activity. The author used the CQs in the requirement stage of the SAIBUS, as later demonstrated in chapter 4 of this thesis.

2.6 Ontology Development Methodologies

Previous authors have proposed different ontology-building methodologies. Gruninger and Fox (1995) described a methodology known as the Toronto Virtual Enterprise (TOVE) to guide ontology design, focusing on ontologies to support reasoning in industrial settings. Their approach is considered first to identify problems through interaction with industrial partners. Second, define the ontology's language or vocabulary, objects, properties, and interactions. Thirdly, to define the significance and limitations of the terminologies. Finally, to check the ontology competencies by providing comprehensiveness theorems on competency questions.

This approach has been employed for many applications in the creation of ontologies. For instance, the Enterprise Design Workbench and the Integrated Supply Chain Management Project agents were built using the TOVE approach (Fernández López and López, 1991).

Uschold et al. (1996) proposed an approach called Skeletal for creating ontologies that include defining the Purpose and Scope first. This stage helps to understand why the ontology is being developed, describe its potential application, and define and detail the selection of its expected uses. Secondly, to build the ontology. Thirdly, the Evaluation. At this point, the ontology established would be tested to check if it fits the frame of reference, which could be the criteria of requirements or questions of competencies or the real world. Fourthly, the documentation of the implemented ontology. Lastly, the Guidelines for each phase. Here, the initial guidelines for creating ontologies which include a collection of procedures, approaches, and concepts for each of the above stages, are outlined.

A different methodology called MENTHODOLOGY has been proposed within the Artificial Intelligence Laboratory at Madrid’s Polytechnic University. This approach involves creating ontologies at the level of awareness, identifying the ontology construction process, a development cycle based on evolving concepts, and comprehensive techniques for implementing each process. This ontology approach was initially proposed for the Intelligent Physical Agent, which promotes interoperability across applications of agents. The MENTHODOLOGY approach has been proven to be the most mature, with suggestions that pre-development procedures are required and that all practices and strategies should be documented (Fernández López and López, 1991; Grüninger and Fox, 1995; Fernández-López et al., 1997).

Swartout and Russ (1997) proposed the SENSUS methodology. In this approach, the domain terms are associated with SENSUS, and any irrelevant vocabularies are adjusted in SENSUS. This approach suggests the following steps when building an ontology in a specific domain: (1) A collection of terms is taken as seed; (2) These seed terms are manually connected to SENSUS; (3) All meanings are included in the route from the seed to the root of SENSUS; (4) Terms that might be relevant within the domain but did not yet exist are included; (5) finally, based on the application, the application is added. Some researchers suggest that various methodologies converge.

Noy and Mcguinness (2001) suggested an approach referred to as ontology development 101, which is an iterative process that suggests the following steps: (1) Decide the ontology domain and scope; (2) Suggest reusing current ontologies (3) Define essential ontology concepts (4) Identify classes and class hierarchy (5) Define class property — slots (6) Define slot facets (7) Identify instances.

The Lean Ontology Development (LOD) was proposed by Cummings and Stacey (2018), which involves the LOD principles and recommendations for current ontologies methodologies that can be used with these principles for ontology development. Their approach to ontology development is based on Agile software development and the Lean Start-up philosophy of organisational development. We have discussed the LOD approach in more detail in section 2.10 of this study.

2.7 Ontology Methods and Tools

When developing a domain ontology, there is a need to understand the methods and tools used in ontology development. Many scholars have recognised that ontology development is an iterative process that can begin with a rough initial approach and then evaluate and improve the emerging ontology. The development of ontology may involve the following steps (Kapoor and Sharma, 2010):

- Identify the domain and scope
- Suggest the reuse of a current ontology

- Define essential terms
- Categorise classes and hierarchies of classes
- Define class attributes and restrictions for these attributes
- Build class instances.

Ontology development requires the use of different tools for its design and implementation. Many open-source tools help build ontologies, such as Protégé, WebODE, SWOOP, Apollo, and IsaViz. Protégé was chosen as the editor for implementing our solution in this study. Protégé is an Ontology editor focused on the experience that includes a graphical user interface (Musen, 2015). It is chosen because it offers a much greater meta-modelling functionality, allows domain ontologies to be constructed, and enables customisation of data entry types to enter data (Corcho, Fernández-López and Gómez-Pérez, 2003; Kapoor and Sharma, 2010). Therefore, this enables the classes and properties to be described, including their hierarchies, variables, limitations of variable value, and relationships in the study's domain.

One must have an ontology specification language to create an ontology. The Web Ontology Language (OWL) is the generally accepted standard among many ontology languages for representing and sharing information in the semantic web context (Hitzler and Rudolph, 2012). OWL is part of the Framework of the Semantic Web that consists of RDF, RDFS, and SPARQL.

2.8 Modularisation of ontology

Conceptually, Modularisation refers to a situation in which a whole item exists but can also be used as a collection of components. Thus, in the context of ontology, a Module is a sub-ontology that can be connected and associated with each other by interacting with some other sub-ontologies (Bakhshandeh et al., 2013). This concept facilitates the development of the large ontology, as each module independently built and managed can help minimise the scope of the development and facilitate ontology integration and reasoning. For example, the Financial Industry Business Ontology (FIBO) is a large ontology built from a series of sub-ontologies in which each model has a basic financial area (Petrova, Tuzovsky and Aksenova, 2017). Therefore, the development of an ontology for the southern African informal market may use the modular ontology theory throughout the development process due to the complexities of the informal sector environment. In general, modular ontology is based on three concepts which include: (1) Upper domain ontology (UDO), (2) Domain-specific ontology (DSO), and (3) Ontology integration. The upper domain ontology helps define general concepts broadly and at a top level between the domains. Domain-specific ontology is a form of ontology specialised in-depth in a collection of definitions presented in the upper ontology

(Doran, 2009). Ontology integration is a method of constructing new ontologies by discovering similar concepts between two or more ontologies.

Two major topics are covered in ontology modularisation, including module extraction and modular development (Osman, ben Yahia and Diallo, 2021). Module extraction focuses on decomposing existing ontologies into separable modules using various logical criteria, whereas modular development is concerned with the development of ontologies in modules, like modular software engineering (Romero et al., 2016). For this study, we are concerned with modular development, in which we used ontology design patterns and OWL import mechanisms to help our modular development process.

2.9 Ontology Design Pattern

Ontology design patterns (ODPs) are successful reusable solutions to a recurring ontology modelling situation (Blomqvist et al., 2015). In theory, they are like software engineering design patterns. An ODP is a small, self-contained ontology or a template used to represent and possibly tackle a modelling problem (Mazimwe, Hammouda and Gidudu, 2019). Patterns are generally expressed explicitly using a logical axiomatisation like the Web Ontology Language OWL (Blomqvist et al., 2015). There are different types of ontology design patterns, such as the ones described in the semantic web portal for ontology design patterns (Gangemi, 2005)

2.10 Lean Ontology Development

LOD is an approach to developing a new product that proposes a technique to be implemented to develop the product. It is based on the Lean product development concept, a life cycle process called Build-Measure-Learn (León and Farris, 2011). It can be executed in 4 stages shown in Figure 2.1:

(1) **Learn:** the problem to be addressed is identified in this stage. To make development and testing easier, objectives and requirements must be well stated and documented, just as with any other software (and ontology) development.

(2) **Build:** a minimum viable ontology (MVO) is developed to collect users' input as early as possible during the development process. This is based on the MVP concept (Rancic Moogk, 2012). The importance of development speed is one of the lean principles. Speed is connected to the notion of gathering user feedback as early as possible throughout the development process. The real purpose of an MVP is to offer your concept or ideas so that you may gather user input as soon as possible. It enables the development team to get the most verified information possible about clients and their product expectations with the least amount of work.

(3) **Measure:** During this phase, users and developers can test the MVO. Users and the development team will be able to do standard testing in case studies, formal analysis, and monitoring. All these activities will convey information and criticism to the development team in a way that is not exclusively impacted by the development team's point of view.

(4) **Learn:** The principles in the current MVO may be changed to meet user expectations better, and the next iteration of the MVO can be created by giving the product greater capabilities.

Several directions open at this stage, including updating the current MVO to help fulfil user expectations and adding additional features to the ontology to create the next MVO iteration. The LOD principles which are used to guide the methodologies used for ontology development are:

i. **The Continuous Development principle**: The LOD understanding of ontologies recognises ontologies as an ongoing and iterative process; therefore, it recommends the importance of incorporating versioning into the process of development ontology adopters can prepare to modify and evolve along with ontology.

ii. **The Minimum Viable Ontology via Prioritisation (MVO) principle**: Given the introduction of versioning, this concept highlights the importance of prioritisation to allow the creation of a Minimum Viable Ontology and lead to the first ontology iteration, which can then be made available to the community of users.

iii. **Community Evaluation principle**: LOD suggests that it is essential to understand and engage with the users' group. Therefore, the ontology-building process must be identified and reported for the testing of the ontology.

iv. **Ontology as Application Programming Interface (API)**: This highlights the adoption of ontology as an API by ontology developers to be used as an interactive approach that consists of ontology-building descriptions, protocols, and tools.

v. **Reuse principle**: One of the main advantages of ontology is considered to be this concept of reusing. Cummings and Stacey (2018) argued that ontology developers could often not be confident about reusing other ontologies because of several considerations, such as the large scale of the current ontology or irrelevant components. This can lead to the formulation of a new ontology.

vi. **Sustainability**: This theory is the culmination of all previous principles for developing a sustainable ontology.

Cummings and Stacey (2018) suggested some recommendations for current ontology approaches that can be applied in the context of LOD, which aims to demonstrate how different methodologies can be applied to LOD but is by no means an exhaustive list. Among many of the methodologies that can be used in the context of the LOD principles includes:

Development of competency questions: This can apply the LOD principle of Minimum Viable Ontology via Prioritisation.

Tree stage Ontology Development: including Upper-Level ontologies, importing ontologies, and developing a new ontology. This can apply the LOD principles of Reuse and Sustainability.

Evaluation and Sustainability: This can use ontology principles such as API, community Evaluation, and Sustainability.

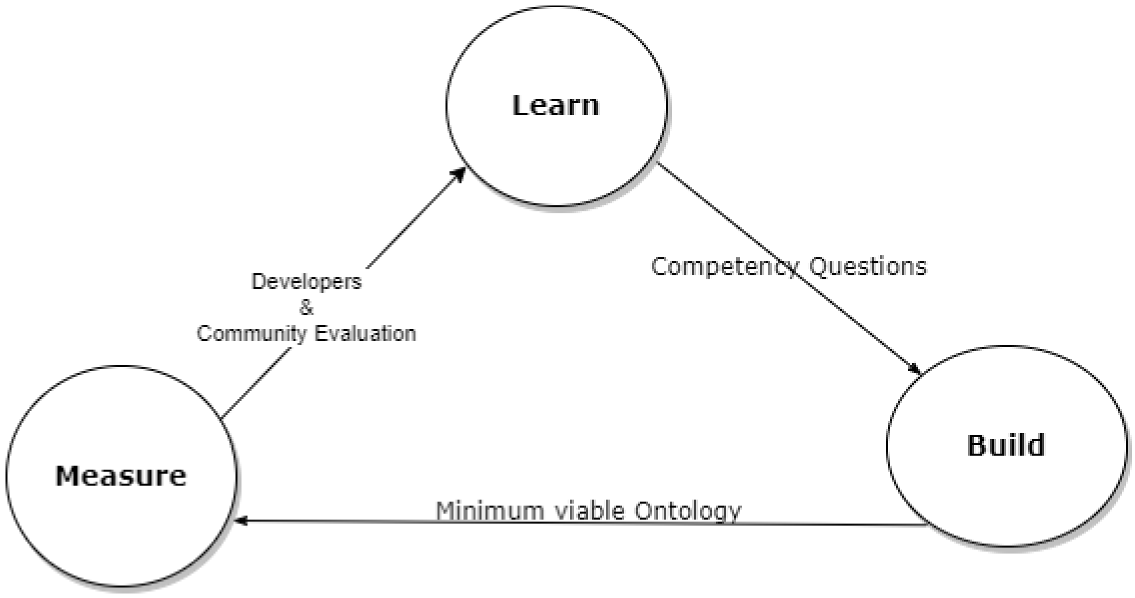

Figure 2.1: Lean Ontology Development (LOD) lifecycle

Although the literature covers a wide variety of ontological methodologies, none of the methodologies compared to the Institute of Electrical and Electronics Engineers (IEEE) standard are fully mature. Some scholars suggest the convergence of more than one methodology. Noy and McGuinness (2001) and a study by (Alsanad, Chikh and Mirza, 2019) supported the argument that a specific correct method is used to build ontologies; nonetheless, we may mix different approaches. For instance, a study by Li and Larsen (2017) combined two methods for developing an ontology. Firstly, METHONDOLOGY has been used to construct ontology at the knowledge level. Secondly, the 101 approaches have been used to create phases for implementing ontology. However, certain complications of the reality that none have fully achieved maturity must be considered by considering the efforts needed to arrive at a situation along the lines of unifying methodologies, approaching knowledge and software engineering together. For this research, the LOD theory and methodology will be considered because the development of domain ontology for the informal South African field is broad, considering the timeframe for the completion of this study, which requires the creation of smaller sub-ontologies that can be combined with other sub-ontologies. Therefore, implementing an Agile approach, such as Lean Ontology Development, will allow the iterative process during ontology development.

2.11 OWL Import Mechanism

The OWL import technique shares many similarities with object-oriented programming import techniques but with some key distinctions. Unlike conventional software systems, where classes and methods are local to a module, OWL axioms and entities are always global, affecting any entity in either importing or imported modules (Rector et al., 2012). As a result, without rules to limit their scope, OWL modules have the potential to interact in unexpected and unwanted ways. Moreover, the OWL imports process is unstructured and low-level.

For example, the impact of OWL imports is global, and the order in which modules are imported is meaningless (O'Connor, 2009; Rector et al., 2012). The axioms are simply collected when one OWL module imports another. As a result, the order in which modules are imported is irrelevant; numerous paths to an imported module are possible. The OWL imports are a suitable mechanism when modules are developed or maintained in separate files. OWL provides ways to distinguish modules such as the IRIs and namespaces. Developing each module in a separate file, considering the IRIs and/or namespaces, as well as the OWL imports mechanism, is a suitable solution for developing a large ontology.

2.12 Ontology Evaluation Methods

Ontologies can be used to give heterogeneously structured databases and various systems equivalent semantics because they explicitly express domains, which are composed of the entities, attributes, and connections that actually exist in the real world. Thus, ontologies enable semantic interoperability and integration in organisations across various fields, including major achievements in the life sciences. However, users are unclear about which ontologies can assist them in resolving their specific data, application, or service issues. Enterprises and communities are unsure if large ontologies created by combining or mapping together smaller ontologies would increase conceptual uncertainty or will enable wider semantic operability for their pooled data and complicated applications (Obrst *et al.*, 2006). It is, therefore, necessary to evaluate the quality of the ontologies.

Ontology evaluation is well-described as a realistic assessment of the characteristics or substance of ontology relating to a frame of reference during every step of their life cycle (Fernández López and López, 1999).

The evaluation process aims to determine what the ontology describes accurately, what it does not, and what it does inappropriately. The ontology evaluation process can be carried out by considering two different contexts, including ontology validation and verification. Ontology verification refers to ensuring that the ontology's definitions are implemented appropriately, and ontology validation refers to whether the definitions' meanings accurately reflect the context in which they were developed (Gomez-Perez, 2001). There are different criteria for evaluating ontology. Some identified criteria are consistency, completeness, conciseness, expandability, and sensitiveness.

Mazumder et al. (2017) summarised some criteria into application-based, metric-based, and domain coverage. Many studies have suggested tools and methods for ontology evaluation, most of which rely on determining quality. The OntoClean suggested by (Guarino and Welty, 2009) and the Oops approach proposed by (Poveda, Suárez-Figueroa and Gómez-Pérez, 2010), for instance, both examine the errors or absence of errors of ontology. Other methods such as OntoQA (Arpinar *et al.*, 2005), OQuaRE (Duque-Ramos *et al.*, 2011), and OntoMetric (Lozano-Tello and Gómez-Pérez, 2004) were suggested with metrics to be used to achieve the overall assessment of the ontology quality. The metrics are designed to assess key elements of ontologies and their ability to represent knowledge. Instead of simply labelling an ontology as effective or ineffective, metrics define a specific feature of it since, in most situations, the ontology's construction is heavily influenced by the domain in which it is constructed (Arpinar *et al.*, 2005).

Mazumder et al. (2017) suggested using the OOPS tool and the reasoner HermiT 1.8.3. Online tool OOPS uses automatic detection to find probable faults or pitfalls (Keet, Suarez...2013). An ontology reasoner called HermiT 1.8.3 can decide whether an ontology is consistent. Additionally, it reveals linkages of class subsumption and offers effective reasoning. In this study, we have used the OOPS tool to evaluate the consistency, completeness, and conciseness criteria. We also used the metric-based evaluation in which we used the OntoQA ontology evaluation approach. This metric-based ontology provides its metrics in two dimensions, one of which is the schema metric, which evaluates the performance of a schema by assessing its structure in modelling a real-world domain. Another is the knowledgebase metric for determining if a populated ontology by analysing its substance is a richly detailed and accurate representation. We have further considered the evaluation by answering competency questions presented in 4.2.2.

2.12 Related Work

The use of domain ontology for knowledge representation and semantic web technology provides answers to different problems. This section provides an overview of related studies on the topic of a domain ontology for the informal sector.

2.12.1 Domain Ontology for the Informal Business Sector

Ontology has already been acknowledged and implemented in the informal sector. Mazumder et al. (2017) implemented a new ontology to represent the informal sector workers' skills, expertise, qualifications, and personal information, as well as the jobs that can be given in the informal sector. They divided their ontology into two sections, including the worker profile defined by the class Worker and the employment profile defined by JobOffer and Employer classes and their relationships with other classes such as Education, Certificate, and Skill. Their ontology was evaluated using standard metrics and based on the domain coverage. Their review illustrates how the proposed model surpasses current ontologies for either formal or informal economic services in work recruiting. While this suggested ontology encompasses some of the worker profile and role profile of the informal sector, there is still a wide field of the informal market sector that is not yet covered, such as the ontology that would cover the types of services provided within this sector, the business profile, service rating, as well as service costs. In this study, the SAIBUS ontology covers this part of the informal sector.

2.12.2 Domain ontology for the formal business sector

Several types of research have been performed on implementing ontology in the formal business field within a specific domain. For instance, Teimourikia and Fugini (2017) presented the implementation of the ontology of the technique for safety management in a smart work setting. Their strategy involved four steps: first, describing the basic steps of the technique for run-time safety management; second, implementing the ontological knowledgebase for workplace safety; third, describing the concept of ontology restrictions depending on the security standards of the organisations; and lastly, communicating sensitive knowledge to each participant in the safety management team.

An ontology-based business intelligence framework was suggested by (Cheng, Lu and Sheu, 2009). Their work tackled the lack of productivity in handling large financial data; the lack of coordination and knowledge exchange between analysts; the lack of a system to manage generated research studies effectively; and the lack of automation in publishing their reports to clients through their websites or emails.

A study by Mccarthy et al. (2013) suggested an approach to tackling risk and regulation reduction in an increasingly competitive and dynamic financial context using semantic-based technology in a shared network setting. Research by Yun et al. (2009) also suggested the creation of an E-learning domain ontology that can be exchanged and reused with the mentioned skeletal methodology (Uschold and King, 1995).

While several studies suggested domain ontology as a response to different issues, relatively few studies concentrate on designing domain ontologies for the informal business field. Some formal business environment approaches have been proposed, showing that developing a domain ontology can answer various problems. Therefore, in this research, we have concentrated on developing an ontology for the South African informal sector.

2.13 Summary

This chapter discussed an all-encompassing and thorough review of the literature relevant to this study's main areas. The study's informal sector, South Africa's informal business sector, information retrieval, domain ontology, requirements for ontologies, ontology development methodologies, ontology methods and tools, ontology modularisation, ontology design patterns, lean ontology development, and related work were all reviewed. This was done to highlight the current gaps in the literature, discuss them in detail, and support the necessity of conducting this study.

CHAPTER THREE

RESEARCH METHODOLOGY

The research methodology used in this study is described in detail in this chapter. Methods and approaches were chosen and implemented in the study based on the research goal of developing a domain ontology that can facilitate information retrieval and promote intelligent reasoning in the South African services business field. The remainder of the chapter is divided into 6 sections: (i) research philosophy, (ii) research approach, (iii) research methods, (iv) research strategy, (v) research design, and (vi) conclusion.

3.1 Research Philosophy

Different study paradigms guide the researcher's action and define the researcher's worldview (Kamal, 2019). These philosophical assumptions include positivism, interpretivism, constructivism, and pragmatism. Each philosophy has a particular viewpoint on the epistemology, methodology, and ontology aspects, as well as the research rhetoric. The philosophical paradigm adopted in this study is pragmatism. This is because pragmatism demands that researchers associate themselves exclusively with realistic problems. The pragmatist emphasises the relevance of activity and the natural world of substance and conceptual artefacts and the experiences with them, including actions and impacts. (Rescher, 2000). Therefore, doing dominates understanding in the human sphere, for all understanding must result from doing. Hence, whatever we understand or know is somehow the result of the investigation of our operation. The pragmatic approach, therefore, directs the author's perspective to concentrate exclusively on practical problems and doing.

3.2 Research Approach

In comparison to the inductive method, where the researcher begins with a collection of abstract speculations about a research topic and seeks to make sense of the phenomenon by analysing a range of a single case, the deductive approach begins with a hypothesis to be able to deduce whether the theory is true or not (Blanche, Durrheim and Painter, 2006). This study followed a deductive approach to assist the researcher in determining whether or not the hypothesis behind the development of an ontology for the informal sector in South Africa to tackle some of the challenges experienced in this sector is valid.

3.3 Research Methodological Choice

The research methodology adopted for this study is a mixed-methods methodology, which incorporates both quantitative and qualitative methods. The quantitative method was used to gather and convert data into numerical forms, and statistical analysis was applied to assess the significance of the results, guided by a deductive approach.

Qualitative methods, on the other hand, were used through a case study approach. The case study aimed to protect the integrity of the narrative data and provide in-depth insights into the services offered by informal businesses in South Africa. The use of a case study allowed the researcher to collect and analyse data from a real-life context, providing a rich and detailed understanding of the informal business sector in South Africa.

The findings from the case study provide a qualitative perspective on the challenges faced by informal businesses, as well as their opportunities for growth and improvement, and were used to validate or contradict the hypothesis, providing evidence-based insights into the informal business sector in South Africa.

3.4 Research Strategy

A research strategy is a comprehensive research plan that directs a researcher in preparing, executing, and evaluating research (Blanche, Durrheim and Painter, 2006). Design science research methodology (DSR) was adopted as a research strategy and was used to define the stages of this research.

3.4.1 Design Science Research

DSR aims to solve a particular problem by creating an artefact in the research context (Hevner et al., 2004). March and Smith (1995) explained the essence of design science by elaborating on the forms of design science practices that involve building, evaluating, theorising, and justifying. They further classified the DSR output as constructs, models, methods, and better theories and argued that these practices are important to ensure that Information Technology research is appropriately conducted and effectively done. This argument motivates the author's selection of the DSR methodology, considering the objectives of this study. Design science research can be structured and summarised in three phases, including identifying problems, designing solutions, and evaluating (Offermann *et al.*, 2009) DSR has been conceptualised in 4 stages (Laumer and Eckhardt, 2012):

i Awareness of the problem: this stage is where the problem is identified and defined.
ii Suggestion: in this phase, a tentative recommendation for a problem solution is drawn from or generated using an acceptable analysis technique from current expertise or hypothesis focused on the problem area.
iii Design and Development: This phase is an innovative stage in which the concept is further designed or developed and creates an actual artefact through several iterations.
iv Evaluation: In this stage, the implemented artefact or prototype is evaluated according to the functional specification.
v Conclusion: In this, a decision on the adopted artefact is made, and a recommendation is made for future studies.

3.5 Research Design

Research design (RD) is a theoretical structure or strategy that directs research practices to achieve sound conclusions (Blanche, Durrheim and Painter, 2006; Creswell, 2009). This stage enables the author to determine the stages to be followed to attain the objectives defined in this research. The DSR phases are summarised and mapped to the lean development process and ontology methodologies following the LOD principles proposed by Cummings and Stacey (2018). Figure 3.1 illustrates the research design process mapping used throughout this study.
Phase 1 of RD: Following the lean development process, the competency question was used to determine the requirements for constructing a domain ontology for the informal South African business sector. The researcher collected data by reviewing some public records, such as official and policy documents, as well as existing classified web platforms like Locanto and Gumtree, to understand the characteristics of informal sectors and to identify the five categories of services in the informal South African business sector: health and beauty, event services, construction services, automotive services, and transportation services. This information was used to determine the requirements for constructing an MVO for the SAIBUS. The competency questions were then generated based on the six dimensions (characteristics of the services being provided, service quality rating, price of the service, possible modes of payment for the service, customer profile, and service provider profile) from the use case. The literature review was used as a data collection instrument to gather information and understand the relevant aspects of the SAIBUS.
At this stage, a conceptual design was introduced to depict the ontology of the South African informal business sector (SAIBUS) at a conceptual level.

Phase 2 of RD: in this phase, we developed a minimum viable ontology (MVO) for the South African informal business sector (SAIBUS).

Phase 3 of RD: in this phase, we evaluated the developed MVO using the OOPS! approach and the standard metrics and by answering the competency questions using the SPARQL query.

Phase 4 of RD: in this phase, we recommended the next version of the MVO.

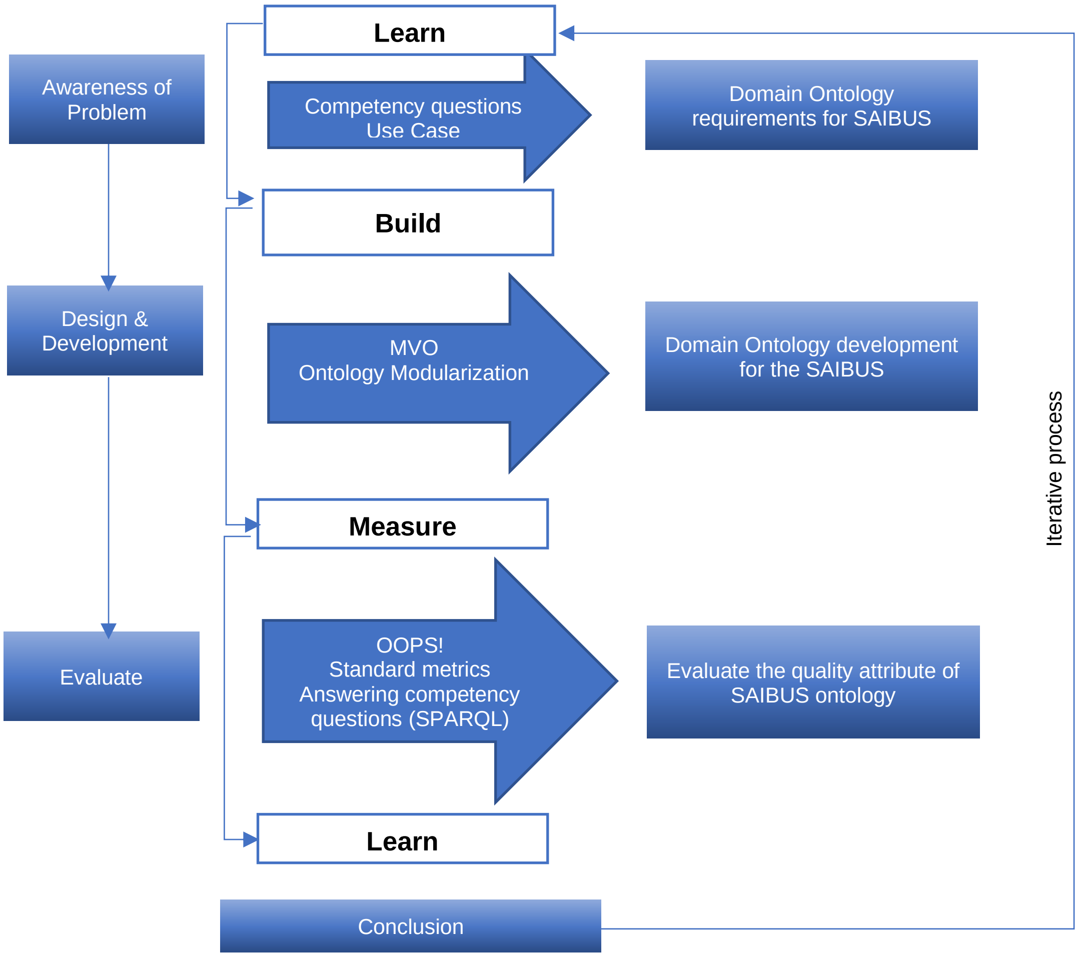

Figure 3.1: Research Design Process

The mapping between the study's objectives and the adopted research design is presented in Table 3.1.

Table 3.0.1: Research Design Process Mapping

Research Design Phase	Objective	Lean Ontology Development process	Methods/Activities
Awareness of the problem	To determine the requirements for developing a domain ontology for the South African informal business sector	Learn	Competency question Use Case
Design & Development	To design and develop a domain ontology for the South African informal business sector	Build	Ontology modularisation Minimum Viable Ontology
Evaluation	To evaluate the quality attributes of a domain ontology for the South African informal business	Measure	OOPS method Standard metrics Answering the competency questions (SPARQL query)
Conclusion	To evaluate if the hypothesis is true or false	Learn	Recommendation of the next version of the MVO

3.6 Summary

We discussed the methodology used in this study in greater detail in this chapter, which is one way this research contributes to achieving its objectives. The selected research philosophies, techniques, methods, and strategies employed for this study are described to build a foundation for the next chapters.

CHAPTER FOUR

ONTOLOGY REQUIREMENTS AND DESIGN

This chapter describes the requirements of the South African informal business sector ontology and the conceptual design of a modular ontology for the South African informal business sector, which will serve as the foundation for the ontology implementation. It outlines the requirements and guidelines for the design and development of the SAIBUS using the competency questions created from the SAIBUS use case scenario described in sections 4.2.1 and 4.2.2 of this chapter.

As part of the LOD approach in the Learn stage, competency questions were used to establish the requirements for constructing a domain ontology for the South African Information Business Sector. This is the first step in the LOD approach selected for this study: to learn about the problem to be solved by developing a minimum viable ontology. These competency questions are based on the SAIBUS use case presented in section 4.2 and serve as a basis for the conceptual design of the SAIBUS ontology.

4.1 Use Cases of the SAIBUS

The South African informal sector has different categories of services (Fig.4.1), and the SAIBUS ontology can support information retrieval on each of these. The purpose of building the SAIBUS ontology is to provide an intelligent knowledge model of the South African informal business sector domain that can be shared and reused to support intelligent search, service recommendation, semantic processing of queries, information retrieval as well as intelligent reasoning by apps (see Fig. 4.2).

The concept of design patterns was applied to depict five categories of the SA informal services domain. Five informal business categories were chosen by reviewing some of the existing classified web platforms, like Locanto and Gumtree, where they categorised some of the informal services and where users may post their services and choose the category they fall under. (Gumtree South Africa, 2022; Locanto, 2022)**.**

First, a model was shown that illustrates how a modular subontology may be created independently and then combined as a component of the SAIBUS ontology.
Each of the modular ontologies in SAIBUS is designed to cover six information dimensions which are the:

i. characteristics of the services being provided;

ii. service quality rating;
iii. price of the service;
iv. possible modes of payment for the service;
v. customer profile; and
vi. service provider profile.

These six dimensions are motivated by the considerations for peer-to-peer interaction between customers and service providers in the sharing economy (Wirtz *et al.*, 2019). These are also the basic information aspects that a multimodal portal for the informal sector must embrace so that it can be a tool for tackling existing challenges of the informal sector, such as lack of organisation, lack of service standards, and low service quality (Daramola, 2018).

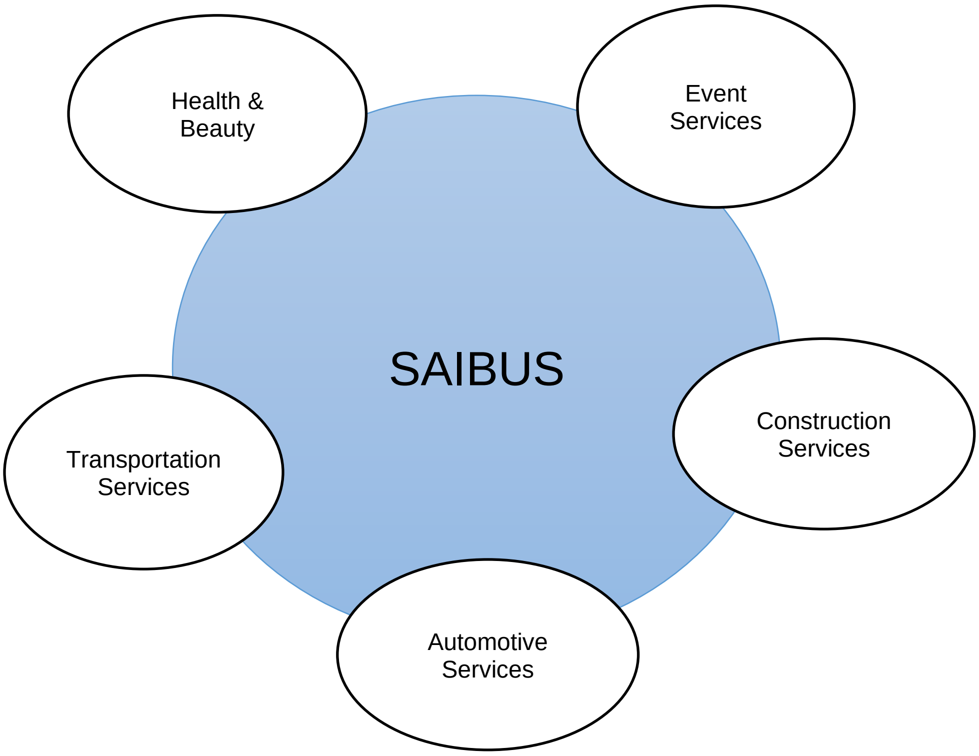

Figure 4.1: Some Service Categories in the Informal Sector

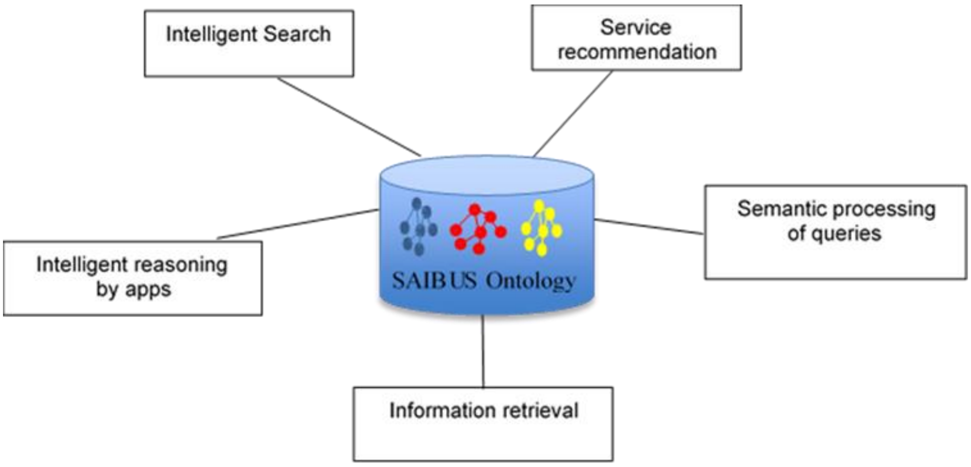

Figure 4.2: Use Cases of the SAIBUS ontology

The selection of these five categories of health and beauty, event services, construction services, automotive services and transportation services in the South African informal business sector was based on a review of some existing classified web platforms such as Gumtree and Locanto, where services posted by informal businesses are categorized. These categories are some of the most in-demand services offered by the informal sector and provide a representative sample of the services provided by informal businesses in South Africa. The choice of these categories was motivated by the need to provide a comprehensive and representative model of the informal sector services domain that can be used to support intelligent search, service recommendation, semantic processing of queries, information retrieval, and intelligent reasoning by applications. The categorization was also based on the consideration of the characteristics of the services being provided and the six information dimensions covered in the sub-ontologies. These dimensions are crucial to understanding the services offered in the informal sector and include characteristics of the services, service quality rating, price, payment options, customer profile, and service provider profile.

4.2 Competency Questions of the SAIBUS

As stated in the literature, competency questions are interrogatives that define an ontology's scope. Driven by use cases and the literature, the identified competency questions were presented in Table 4.1, which were mapped with six dimensions that each sub-ontology should satisfy. These CQs are important in the ontology life cycle because they help the researcher define the requirements of the ontology, and as a result, the CQs become a functional requirement of the ontology.

Table 4.1: SAIBUS Dimensions and Competency questions mapping

S/n	Dimension	Competency Questions
1	Informal service attributes	• What is the type of service being provided? • What is the service description? • To which category does the service belong? • In which location is the service? • What is the nature of the service (in place; service/nomadic service)?
2	Informal service quality rating	• What level of quality is expected from the service?
3	Informal service pricing	• How much does the service cost? • What are the prices for the range of available services?
4	Informal service transactions	• What are the modes of payment (prepaid, post-paid) for the service?
5	Prospective customer profile	• Who is requesting the service? • What can be known about the customer (address, rating by service providers, gender)
6	Service provider profile	• Who is providing the service? • What can be known about the service provider (address, customer rating, gender, years of service)

4.3 A Modular Design of the SAIBUS

After identifying the competency questions, a conceptual design of the South Africa Informal Business Sector ontology was presented at this stage to assist with the representation of the ontology of the SAIBUS at a conceptual level. This design employed ontology modularisation modelling theory and design patterns. Cummings and Stacey (2018) stated that the LOD methodology does not exclude ontology developers from adopting approaches such as ontology design patterns. However, they have advised using them in the learn stage of the LOD if they do not violate the LOD principles. Therefore, this conceptual design is considered part of the learn phase to present a prototype of the gathered requirements by defining components and modules of the SAIBUS and supporting the implementation of the ontology while following the LOD principles.

The South African informal sector is a vast domain; hence the development of the SAIBUS ontology can consist of sub-modules that can be built independently and integrated to support information retrieval. The modelling of the SAIBUS modular ontology design follows the steps suggested by Shimizu (2018). These are:

i Define the use case of the SAIBUS
ii Generate competency questions and scope the problem. In this step, it was decided what could be modelled now and what should be modelled later
iii Identify key notions from the use case
iv Instantiate key notions with patterns
v Add axioms
vi Create the OWL file
vii Evaluate the designed ontology

The following sections explain how steps (i) – (iv) covered have been followed to realise the SAIBUS modular ontologies. Steps (i) – (ii) involve defining the use case and generating the competency questions already covered in section 4.2 of the requirements of the SAIBUS.

4.3.1 Identification of key notions and pattern selection for SAIBUS modules

The key concepts from the use cases and competence questions were identified for the modules at this stage, and the ontology design pattern that should be used for each key concept was selected. By considering the generated competency questions, the following use case scenario of the SAIBUS modular ontology was identified: i) finding the category or type of service being provided; ii) finding the service descriptions; iii) searching for service in a specific location; iv) retrieving recommended service based on the price and quality of the service; v) finding information on the competency level of a service provider; vi) finding a customer's profile; viii) finding a service provider's profile.

The SA informal sector was modelled using patterns corresponding to sub-ontologies of the SAIBUS modular ontology since the ODPs are effective reusable solutions to a recurrent ontology modelling issue. Each pattern is a module that defines which ontological entities belong to which module, pattern, or ontologies. The existing ODPs were used for some of the identified key notions listed in Table 4.2. The Comprehensive Modular Ontology IDE (CoModIDE) plugin within the Protégé ontology editor was to model the SAIBUS modular ontology (Mazimwe, Hammouda and Gidudu, 2019). The CoModIDE is a plugin that allows for graphical drag-and-drop ontology building with an incorporated Ontology Design Pattern library.

Each pattern used in modelling the important concepts is described below:

- i **Stub pattern** Was used for Service Description, Service Category, Service sub-category, mode of payment, and Service Rate. The goal of this meta pattern is to serve as a kind of placeholder for future ontology expansions in circumstances where finer-grained modelling would be detrimental at this time. Still, more information may be needed for future extensions (Mazimwe, Hammouda and Gidudu, 2019).
- ii **Place pattern** was used to represent the Service Location. This pattern defines where something is located, even though it is unclear which location relationship should be described. It has a class called Place and Object. Its object properties include isLocationOf and hasLocation (Wirtz *et al.*, 2019).
- iii **Price pattern** was used to represent the Service Pricing, which denotes an entity's price in a particular currency. The Price pattern has three object properties: hasCurrency, hasPrice, and isPriceOf, as well as a currency class and a Price class (AldoGangemi, 2010).
- iv **ProspectiveCustomer** this pattern was built to represent information about the person attempting to seek the service. Figure 4.3 shows the ODP pattern for the Customer pattern.
- v **ServiceProvider pattern** this pattern was built to aid in representing information about the person providing the service. Figure 4.4 shows the ODP pattern for the ServiceProvider pattern.

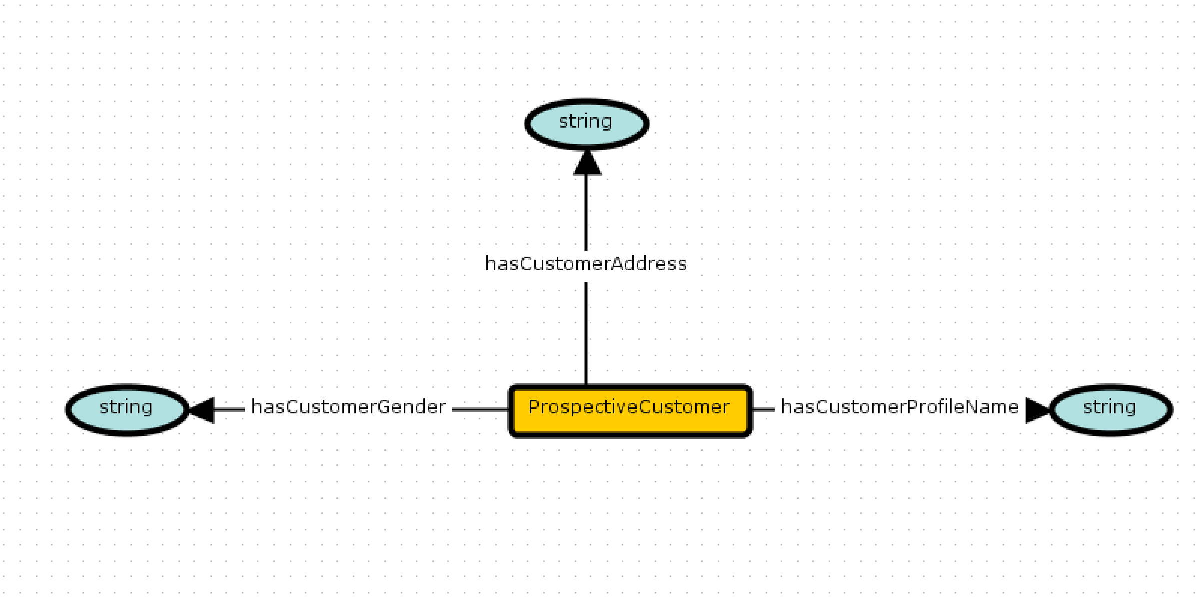

Figure 4.3: Prospective Customer Pattern

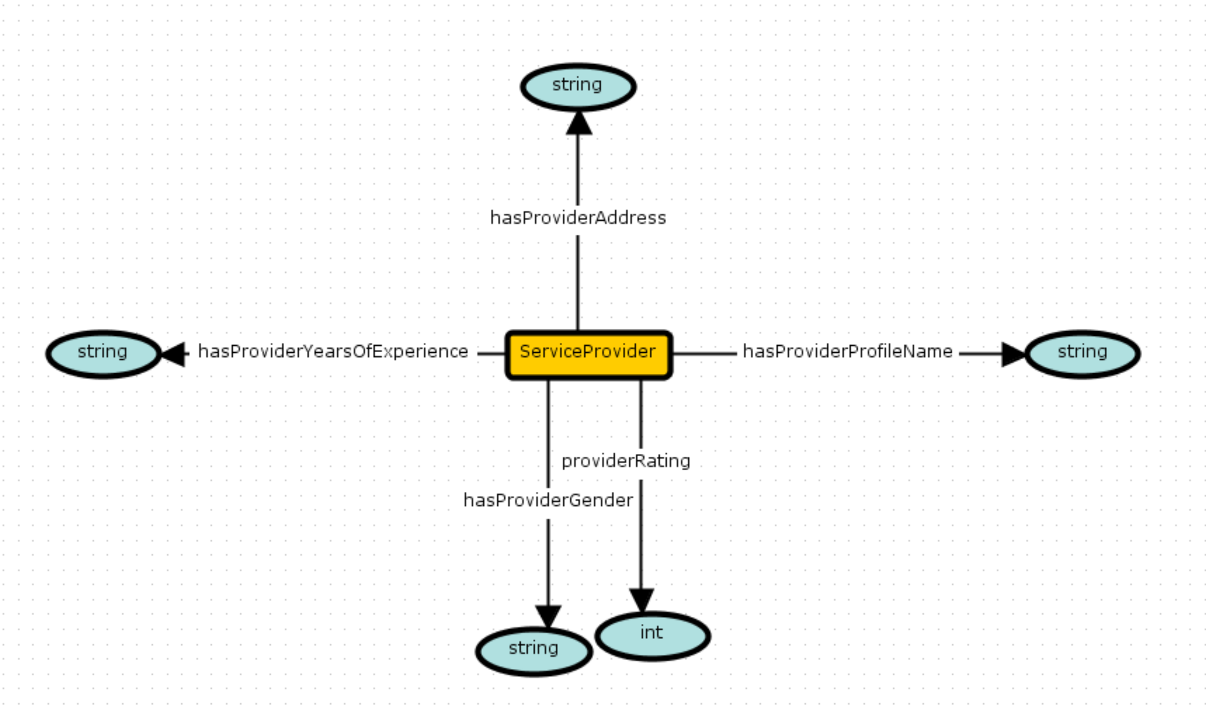

Figure 4.4: Service Provider Pattern

Table 4.2: Key notion and Pattern Mapping

Key Notion	Pattern
ServiceDescription	Stub
ServiceProvider	ServiceProvider
ServiceLocation	Place
ProspectiveCustomer	ProspectiveCustomer
ServicePrice	Price
Service Subcategory	Stub
ModeOfPayment	Stub
ServiceRate	Stub
ServiceNature	Stub

4.3.2 Instantiate key notions with patterns

Each of the core concepts determined in the prior phase was represented by a schema diagram. Each notion was instantiated with the corresponding pattern as specified in Table 4.2.

4.3.3 Module assembly and naming convention revision for SAIBUS classes, properties, and individuals

After instantiating each module, they were connected in this phase. As stated in Fig 4.5 to Fig 4.9, the same steps were used to model all the five sub-ontologies of the SAIBUS using the same patterns by repeating steps 43.2 to 4.3.5.

For instance, the instantiated modules for Health and Beauty were put together as one category within the SAIBUS. The Health and Beauty sub-ontology of the SAIBUS contains defined modules such as Service Sub-category, Prospective Customer, Service Provider, Mode of Payment, Service Rate, Service Description, nature of the service (nomadic or in-place) and Service Price. The Service Subcategory, for instance, has sub-categories of health and beauty such as massage service, manicure service, pedicure service, hairdressing service, makeup service and more. Each of these modules was instantiated from the patterns and put together in this stage to produce the Health and Beauty sub-ontology shown in Figure 4.5.

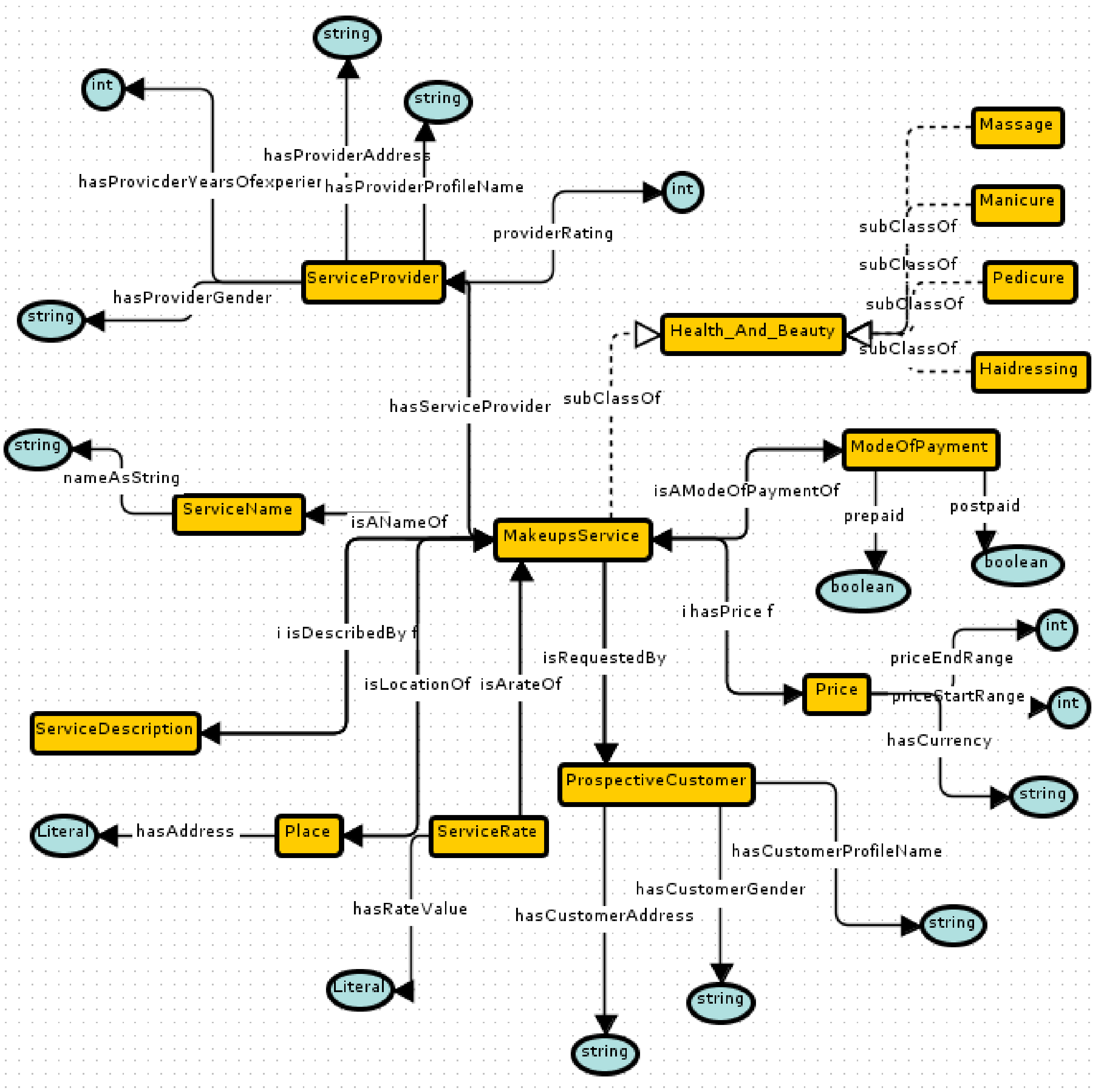

Figure 4.5: Health and Beauty ontology model

Secondly, Event Services was modelled as one of the categories of the SAIBUS. The Event services sub-ontology of the SAIBUS also has all the modules instantiated from the ODPs. On top of the common defined modules, the Event services have sub-categories such as catering, decoring, wedding planning, lighting service and djing service, which have been modelled and presented in Figure 4.6.

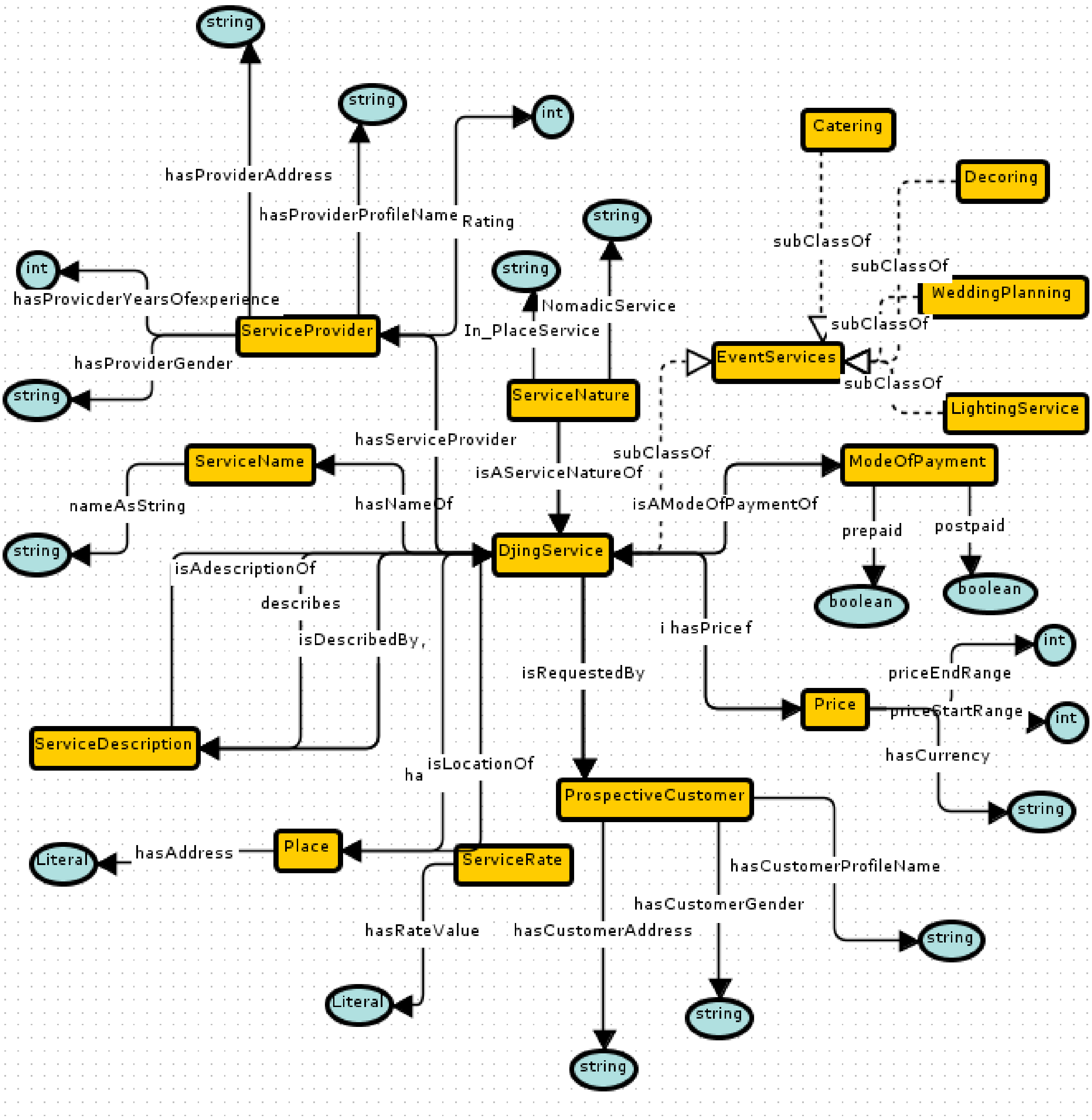

Figure 4.6: Event Service ontology model

Thirdly, a model of the SAIBUS category for automotive services was created. There are several subcategories of automobile services, including towing, tyre, brake, and engine service. Figure 4.7 shows the Automotive Services modelling.

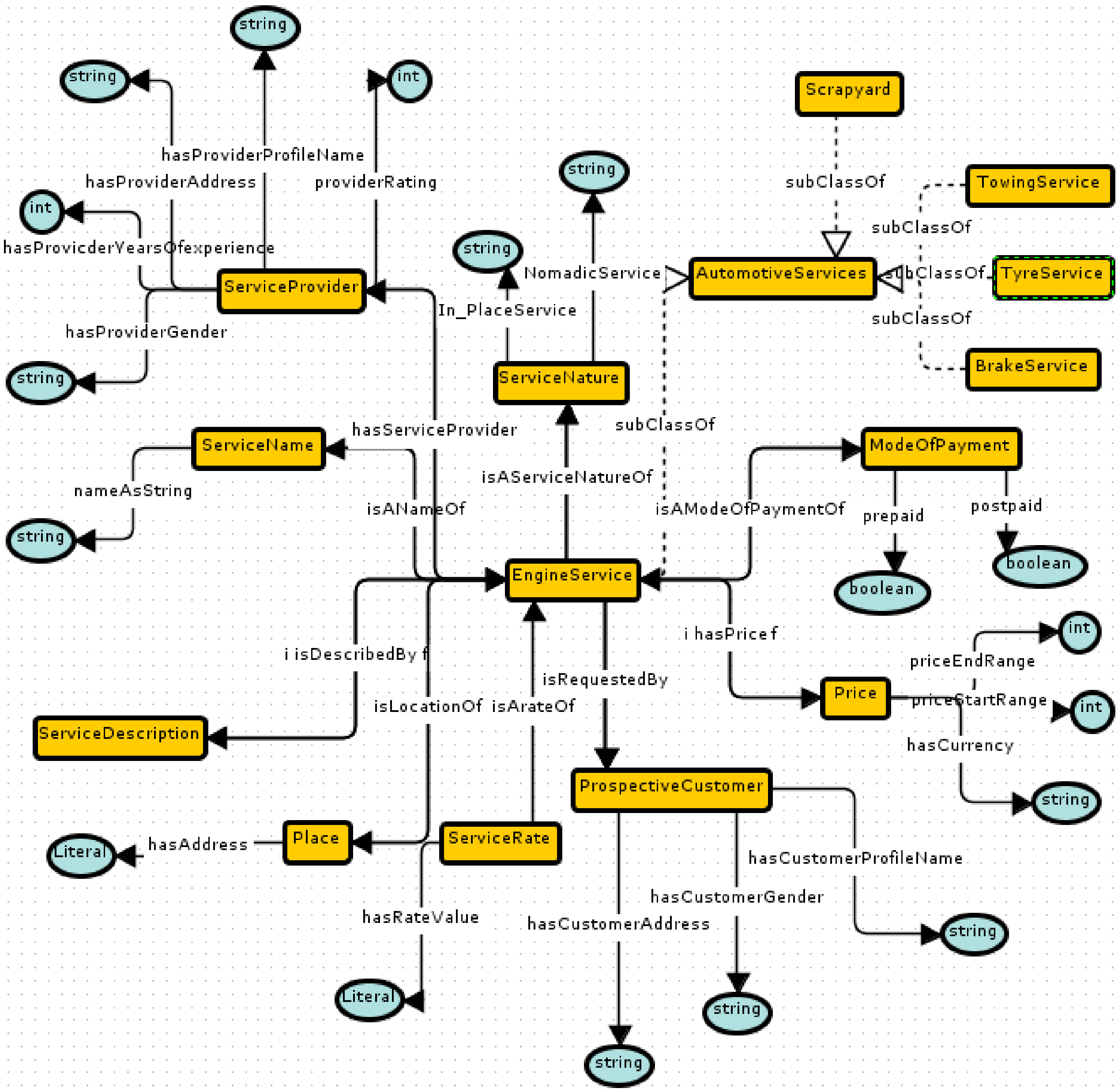

Figure 4.7: Automative Service ontology model

Fourth, a model of the SAIBUS category known as "transportation services" was built. Sub-categories of transportation services include delivery service, taxi service, school transportation service, and cab tour. Figure 4.8 shows the model for transportation services.

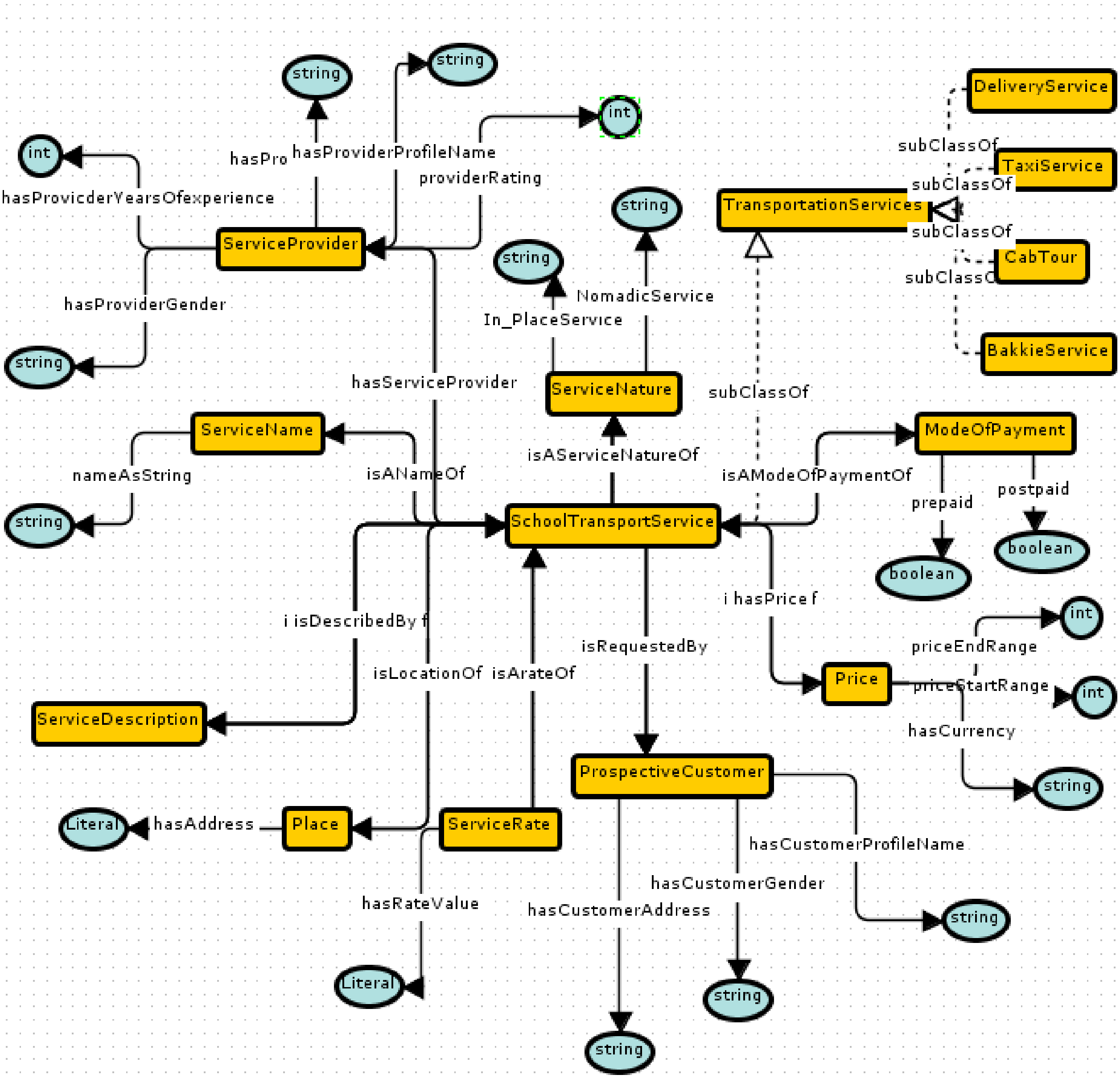

Figure 4.8: Transportation Service ontology model

Lastly, Construction Services was modelled. This is also one of the categories within the SAIBUS and can have sub-categories such as roofing service, pool installer service, painting service, tiling service, and paving service. The construction services also combine all the instantiated key concepts. The construction services model is shown in Figure 4.9.

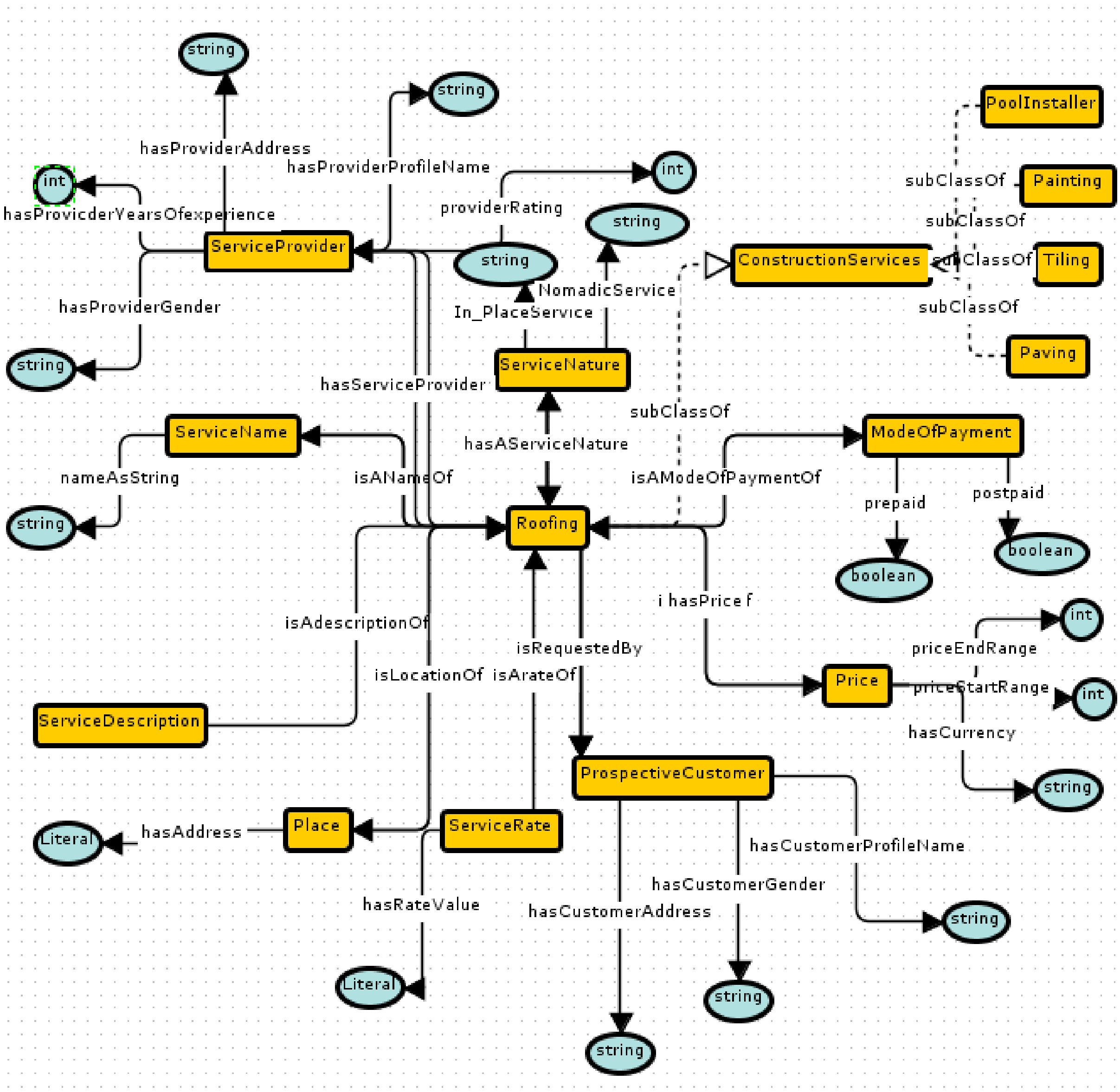

Figure 4.9: Construction Service model with Roofing category

Because each SAIBUS category includes several subcategories, each subcategory can be represented as a service. The assumption is that each service is a subclass of its superclass and is delivered to a specific location. Each service has a price, a prospective customer, a service provider, a service description, and a service quality.

4.3.4 Add axioms for modules

Most of the axioms were derived from the axioms provided by the patterns, and others were added from this stage, as seen in the schema diagram using the OWL Manchester syntax. For a deeper grasp of this, Description Logics (DLs) statements were employed to describe the axiomatisation of the patterns that constitute the foundation of the SAIBUS ontology while avoiding severe commitment at this level of modelling (Grüninger and Fox, 1995).

According to Keet (2020), DLs may be used to represent the interactions between entities in an interest domain and concepts, roles, and individual names are the three different types of entities found in DLs. Roles define sets of binary relations between individuals, concepts denote groups of individuals, and individual names denote single individuals in the domain.

The author further gave an example that an ontology defining the area of people and their familial relationships, for instance, may use terms like *parent* to indicate the (binary) connection between parents and their offspring, individual names like Julia and John to identify the people, Parent to designate the set of all parents, and Female to symbolise the set of all female persons, John (Joan?) and Julia. A DL ontology is a collection of statements, or axioms, each of which must be true in the described context. It does not completely explain any single circumstance or "state of the world." There may be several possible world states that are consistent with the ontology, and these axioms often only capture some information about the situation that the ontology is describing.
In this study, for instance, the stub pattern is logically described as:
$\top \sqsubseteq \forall$ hasValue.xsd: AnyValue/Stub $\sqsubseteq \exists$ hasValue.xsd: AnyValue
Where *hasValue's* range can be any xsd datatype, and *AnyValue* was used to imply that any datatype will suffice in this axiom. The *ServiceDescription*, *ModeOfPayment,* and *ServiceRate* were represented using the stub in the ontology modelling and instantiated with the respective variables since the logic remains the same for all the sub-ontologies of the SAIBUS.
The ServiceProvider pattern can logically be represented as follows:

$\top \sqsubseteq \forall$hasProviderProfileName.ServiceProvider

$\top \sqsubseteq \forall$hasProviderAddress.ServiceProvider

$\top \sqsubseteq \forall$hasProviderGender.ServiceProvider

$\top \sqsubseteq \forall$hasProviderYearsOfExperience.ServiceProvider

Where T can represent the ServiceProvider entity and *hasProviderProfileName*, *hasProviderGender*, and *hasProviderAddress* all have a string range, and *hasProviderYearsOfExperience* has an integer range.

The ProspectiveCustomer pattern can logically be represented as :

$\top \sqsubseteq \forall$hasCustomerProfileName.ServiceProvider

$\top \sqsubseteq \forall$hasCustomerAddress.ServiceProvider

$\top \sqsubseteq \forall$hasCustomerGender.ServiceProvider

The range of *hasCustomerProfileName* and *hasCustomerAddress*, as well as the *hasCustomerGender*, is a string.

The axioms were also created that supported the connection of all the modules. Some axioms using the Health and Beauty sub-domain of the SAIBUS are described in Table 3 as an example keeping in mind that all the sub-ontologies followed the same logical representation.

Table 4.3: Some Axioms for the Makeup Service of H&B

ServiceDescription describes some MakeupsService SubClassOf ServiceDescription
MakeupsService hasLocation some Place SubClassOf MakeupService
ServiceName isANameOf some MakeupsService SubClassOf ServiceName
Price isAPriceOF some MakeupsService SubClassOf Price
ServiceRate isArateOf some MakeupsService SubClassOf ServiceRate
MakeupsService isDescribeBy some ServiceDescription SubClassOf MakeupsService
MakeupsService hasLocation some Place SubClassOf MakeupsService
MakeupsService isPerformedBy some ServiceProvider SubClassOf MakeupsService
MakeupsService isRequestedBy some ProspectiveCustomer SubClassOf MakeupsService
ModeOfPayment isAModeOfPaymentOf some MakeupsService SubClassOf ModeOfPayment

4.3.5 Creating OWL files

After modelling the modules, the sub-ontologies were constructed as a knowledge artefact (OWL file) using the Protégé ontology editor. Each sub-ontologies was saved in a separate OWL file at this stage.

4.4 Summary

In this chapter, the SAIBUS use cases and competency questions were used to provide the requirements for the SAIBUS, and the ontology design pattern was used to build the conceptual designs of the SAIBUS's sub-ontologies, which served as the foundation for the implementation of the ontology that will be covered in more detail in the next chapter.

CHAPTER FIVE

ONTOLOGY IMPLEMENTATION

The SAIBUS's implementation is covered in this chapter. The LOD process is resumed in section 5.1, this time in the Build phase, by presenting the building of an MVO. It demonstrated how the SAIBUS could be built using LOD principles and how its sub-ontologies could relate to its main ontology.

5.1 Ontology Development of the SAIBUS

In this study, a new ontology called SAIBUS Ontology was proposed to overcome some of the challenges experienced in the informal sector, as mentioned in the background and literature of this study. To achieve this, the proposed ontology solution was implemented using Protégé-5.5.0, a free, open-source ontology editor and framework used to create intelligent systems. Keeping in mind the 4 steps of the lean development life cycle, this section of the study covers the second phase, in which the conceptual design of the ontology modularisation was linked with the development of a minimal viable ontology.

During the development stage, the ontology implementation followed the LOD principles discussed in Chapter 2 of this study, such as the continuous development principle, which encourages the importance of versioning into ontologies. The LOD principle on the minimum viable ontology (MVO) via prioritisation was considered, as well as community evaluation, reuse, and sustainability principles. One of the motivations for this methodology was to quickly give a concept, idea, and artefact so that users could provide feedback from the product's first version. This allows the researcher to get feedback as much as possible about users and their expectations by presenting them with the MVO. 5.1.1 to 5.1.3 of this chapter explained how each LOD principle was applied to construct the SAIBUS ontology.

5.1.1 LOD Principle: Continuous Development

According to this principle, Versioning is essential since SAIBUS is a large area, and the building of the ontology will need ongoing and iterative development. Since ontologies will require ongoing and iterative development, versioning must be incorporated into the ontology creation process. So those users can be ready for the ontology to change and develop over time, all ontologies must have roadmaps illustrating the versioning process. It is advised to adhere to the major and minor version numbering approach used by the software industry (Cummings and Stacey, 2018).

5.1.1.1 Ontology versioning of the SAIBUS

Each ontology may contain an ontology IRI. If an ontology has an ontology IRI, it may also have a version IRI, which is used to identify the current version of the ontology. The ontology should be a URL (http://...) that leads to a web page where the ontology may be downloaded (Keet, 2020). In developing the sub-ontologies of the SAIBUS, the ontology version IRI was used in protégé to create each version of the ontology, as presented in Figure 5.1. Each of the sub-ontologies or modular ontologies has been versioned.

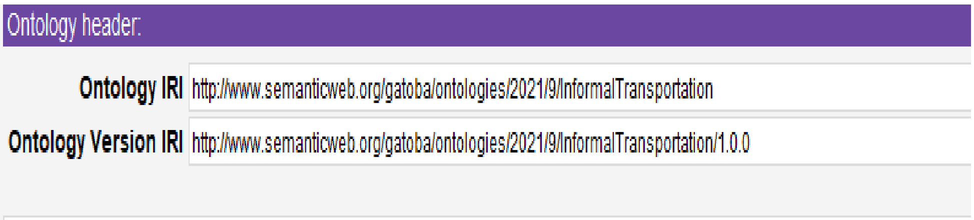

Figure 5.1: Ontology IRI & versioning

A sub-ontology, for example, can be separately maintained or upgraded, and the new version of the sub-ontology can be imported to replace the previous version. Each version of the modular ontology was created in a separate OWL file with an ontology IRI and an IRI for it. Because this is the initial version of each SAIBUS sub-ontology, they all have version 1.0.0, which can later incorporate different versions in the improved artefacts.

5.1.2 LOD Principle: Minimum Viable Ontology via Prioritization of the SAIBUS

Prioritisation brings up the concept of a Minimum Viable Ontology (MVO). To practise this approach, competency questions were used to describe the ontology's required features for what is needed currently for this phase. For this initial MVO, the identified competency questions in 4.2.2 were prioritised based on the use case to provide needed requirements that helped for this first iterative phase. This enabled creating an MVO and releasing the first version of the ontology to the user community. To produce the SAIBUS MVO, each modular ontology of the SAIBUS was modelled separately, and an OWL file was created for each as outlined in section 4.3 and imported the sub-ontologies into the core ontology to create the main SAIBUS MVO. Section 5.1.3.1 of this chapter describes the ontology import process used during the implementation.

Following the modelling of ontologies, CoModIDE is a plugin that enables graphical drag-and-drop ontology building with an integrated Ontology Design Pattern library. During modelling, the CoModIDE plugin helped generate the axioms automatically. Table 5.1 shows some axioms generated from the Heath and Beauty and Event services. As a result, the axioms for each sub-ontology were formed during the modelling step, and Class hierarchies were constructed and reviewed to define super classes and sub-classes of each modular ontology during the modelling stage.

The same approach was used to model and build five modular ontologies that are regarded sub-ontologies of SAIBUS, including the HealthAndBeauty ontology, EventServices ontology, ConstructionServices ontology, AutomotiveServices ontology, and TransportationServices Ontology. Following the modelling and development of these five ontologies, the SAIBUS ontology was created, which serves as the main ontology by importing the SAIBUS sub-ontologies. The next part explains how importing the sub-ontologies into the main SAIBUS ontology was implemented.

Table 5.1: General axioms of the SAIBUS, H&B service and EventService

Ontology Metrics	**SAIBUS**	**HealthAndBeauty**	**EventService**
Classes	CarWash, EventService, Health_And_Beauty, InformalConstruction, InformalManufacturing, InformalMechanical, InformalTrade, InformalTransportation, Module, Pattern, SAIBUS	Health_AndBeauty, MakeupService, Massage, Pedicure, Manicure, HairDressing, ModeOfPayment, Place, Price, ProspectiveCustomer, ServiceDescription, ServiceName, ServiceProvider, ServiceRate	EventService, Djing_Service, SecurityGuard, WeddingPlanner, Decoring, Catering, ModeOfPayment, Place, Price, ProspectiveCustomer, ServiceDescription, ServiceName, ServiceProvider, ServiceRate
Object Property	None	describes,isAModeOfPaymentOf, hasCustomer, hasLocation, hasNameOf, hasPrice, hasServiceProvider, isAModeOfPaymentOf, isANameOf, isArateOf, isDescribedBy, isLocationOf, isPerformedBy, isRequestedBy	customerAddress, customerender, endRange, hasCurrency, hasRateOf, hasValue, postPaid, prepaid, postPaidMode, providerYearsOfExperience, providerProfileName, providerRating, providerYearsOfExperience, serviceDetail, serviceName, startRange, priceStartRange, priceEndRange hasCustomerProfileName, nameAsString
Data Property	None	customerAddress, customerender, endRange, hasCurrency, hasRateOf, hasValue, postPaid, prepaid, postPaidMode, providerYearsOfExperi	customerAddress, customerender, endRange, hasCurrency, hasRateOf, hasValue, postPaid, prepaid, postPaidMode, providerYearsOfExperience, providerProfileName, providerRating, providerYearsOfExperience, serviceDetail, serviceName, startRange, priceStartRange,

		ence, providerProfileName, providerRating, providerYearsOfExperience, serviceDetail, serviceName, startRange, priceStartRange, priceEndRange hasCustomerProfileName, nameAsString	priceEndRange hasCustomerProfileName, nameAsString

In addition, the sub-ontologies of the SAIBUS main ontology use the same naming convention as the SAIBUS sub-ontologies, and when the ontology is imported, the current classes, axioms and other properties that have the same names are overridden with a new imported version.

5.1.2 LOD Principle: Reuse and Sustainability

5.1.2.1 Importing sub-ontologies into the core SAIBUS Ontology

To connect the sub-ontologies to the core ontology, Protégé was used, and the OWL import technique, which falls under the reuse and sustainability principle. The OWL import mechanism was described in section 2.10. The tree import types provided by Protégé's import ontology wizard are shown in Figure 5.2. All the sub-ontologies were imported into a new ontology called "SAIBUS Main," which serves as the main ontology.
Because all the OWL files were kept locally during the development and published later to the ontology repository, the imported ontology contained in a local file option was used for this process. However, one can also import from the web by providing the ontology URL. Figure 5.3 shows the imported ontologies within Protégé, which also provide flexible ways to remove or import more ontology to the main.

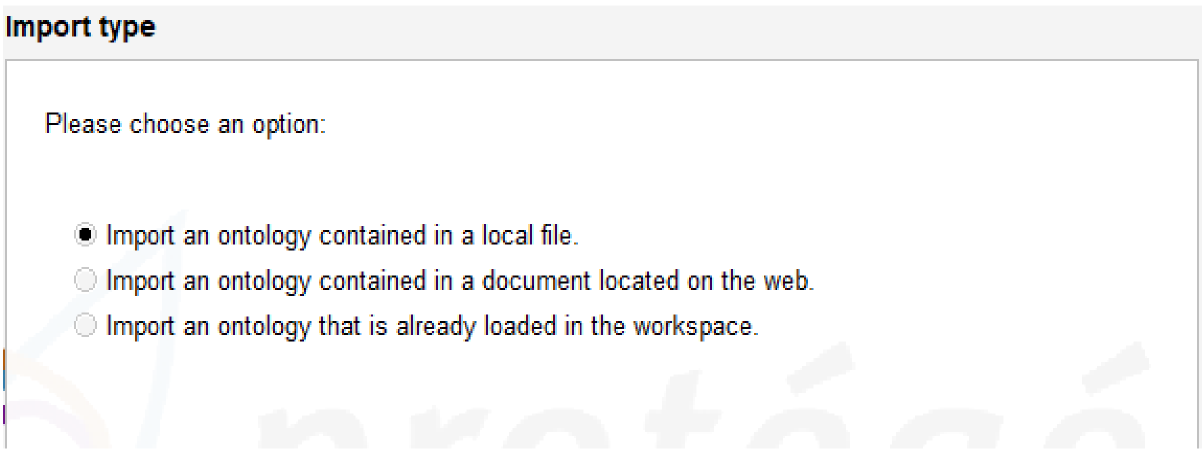

Figure 5.2: OWL Import Type with Protégé

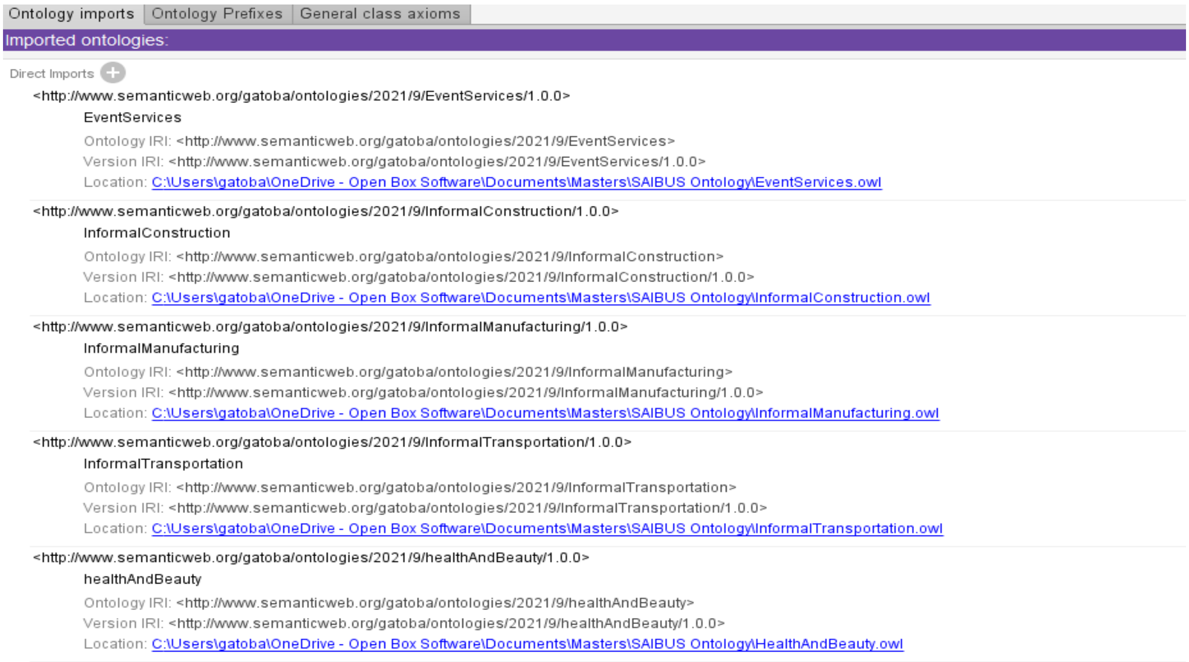

Figure 5.3: Imported Sub-Ontologies within Protégé

Each of the imported ontologies has an ontology IRI and a version IRI. The whole set of axioms, classes, object properties, and data properties has been imported successfully, while the sub-ontologies inside the core ontology share common axioms across ontologies. As an independent module, replacing the ontology with a new version or removing the imported ontology is easy for future import. This same process can be used to import more informal sector sub-ontologies. However, the main ontology needs to be reviewed after the import to maintain consistency.

All ontologies have names and have generated files in the form of OWL/RDF/XML. Figures 5.4 and 5.5 show examples of the OWL imported path within the main SAIBUS ontology. Each line refers to the sub-ontology source loaded into the import destination.

```
<owl:Ontology rdf:about="http://www.semanticweb.org/gatoba/ontologies/2021/9/EventServices">
    <owl:versionIRI rdf:resource="http://www.semanticweb.org/gatoba/ontologies/2021/9/EventServices/1.0.0"/>
</owl:Ontology>
```

Figure 5.4: OWL Import Type within Protégé

```
<Import>http://www.semanticweb.org/gatoba/ontologies/2021/9/EventServices/1.0.0</Import>
<Import>http://www.semanticweb.org/gatoba/ontologies/2021/9/InformalConstruction/1.0.0</Import>
<Import>http://www.semanticweb.org/gatoba/ontologies/2021/9/InformalManufacturing/1.0.0</Import>
<Import>http://www.semanticweb.org/gatoba/ontologies/2021/9/InformalMechanical/1.0.0</Import>
<Import>http://www.semanticweb.org/gatoba/ontologies/2021/9/InformalTransportation/1.0.0</Import>
<Import>http://www.semanticweb.org/gatoba/ontologies/2021/9/healthAndBeauty/1.0.0</Import>
```

Figure 5.5: OWL Import within OWL file

The SAIBUS ontology comprises class axioms imported from the sub-ontologies; each class can be improved and expanded on its own. Classes that reflect the various types and categories of services offered in the informal sector are included below. The relationships between the defined classes and other classes are made up of object properties and different attributes in the form of data properties. Therefore, classes, properties, and individuals inside the SAIBUS ontology can be exchanged between sub-ontologies based on the connections. Following the conceptual design in chapter 4, some of the general class axioms and class hierarchy for the SAIBUS ontology are illustrated in Figure 5.6.

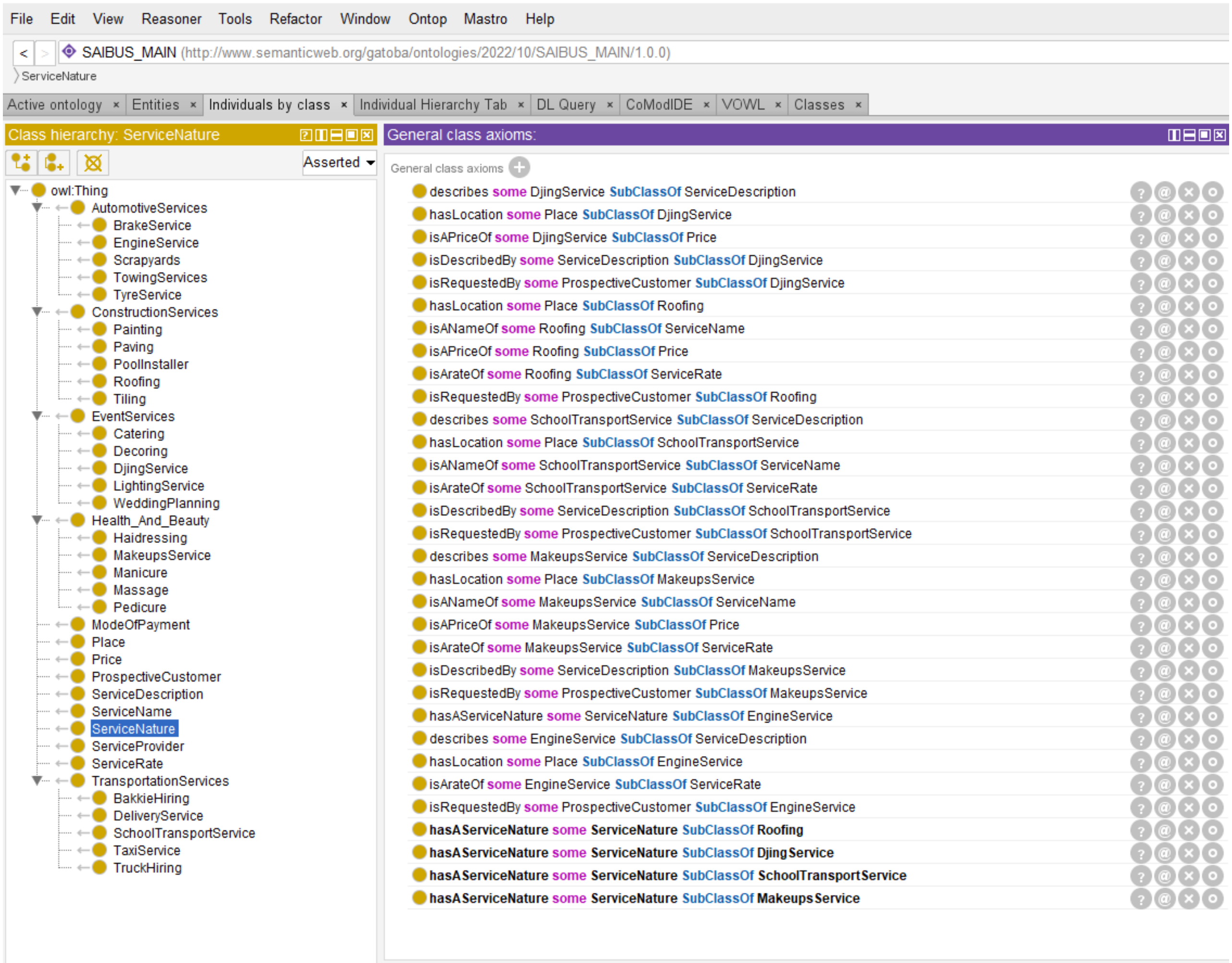

Figure 5.6: SAIBUS Class hierarchy and general class axioms

If a subontology is modified from a location referenced in the SAIBUS main after import, the ontology must be reloaded to include the changes. As a result, it may be modified and used again. A popup window will request to load the subontology file with the new changes if any of the subontology classes are additionally updated from the main class.

The ProtégéVOWL plugin, which is a Protégé plugin for the user-oriented visualisation of ontologies, was used to present each sub-ontology of the SAIBUS. By providing graphical representations for Web Ontology Language (OWL) building blocks used to create a force-directed network structure representing the ontology, implement the Visual Notation for OWL Ontologies (VOWL). This tool was used to visualise and illustrate the sub-ontologies. The ontology visualisation of each sub-ontology is presented in Figure 5.7-5.11. The visualisations display the relationships between classes, the objects and data properties, and the data types used.

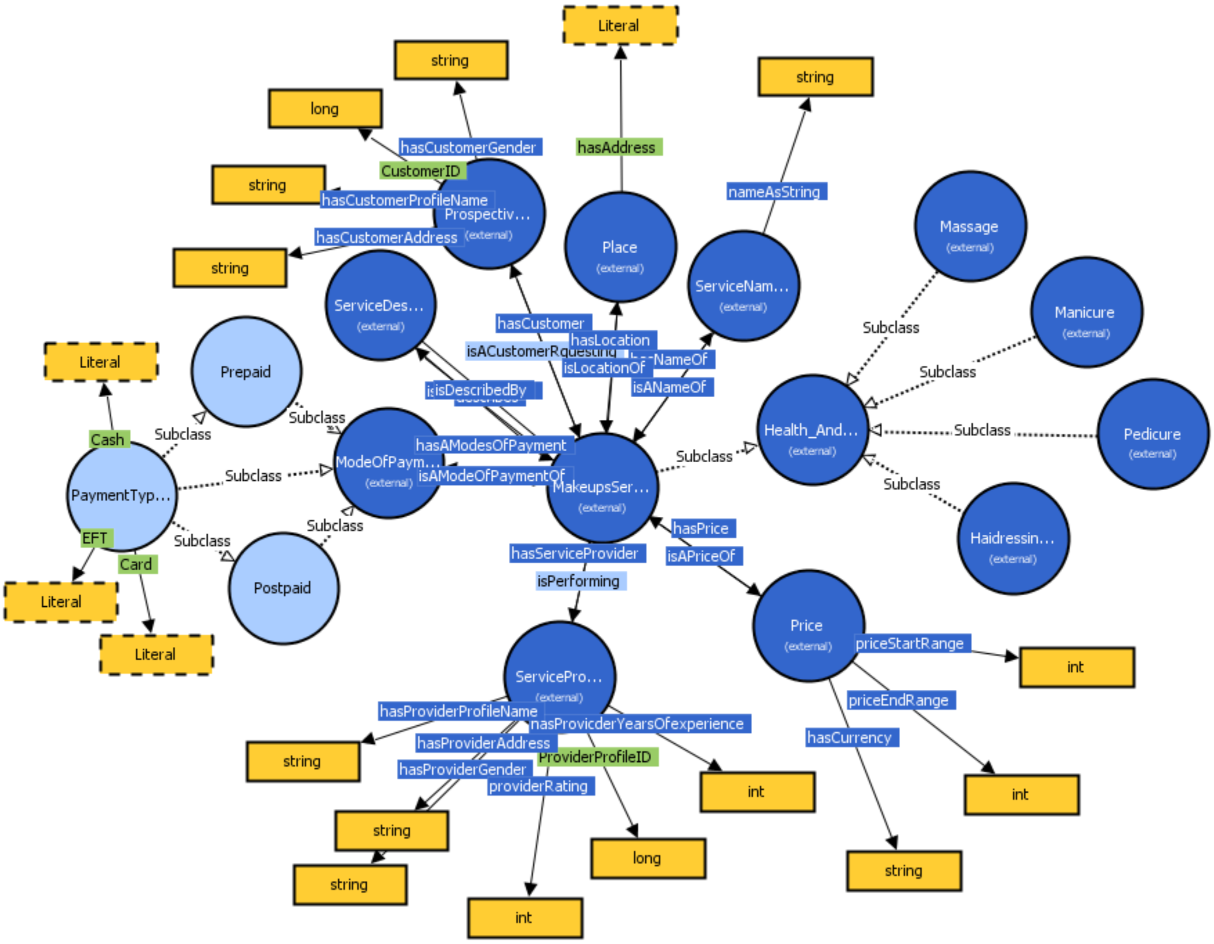

Figure 5.7: Health & Beauty Subontology visualisation

Figure 5.7 presents a visualization of the Health and Beauty subontology using the Ontology VOWL Tool. The visualisation depicts the relationships between classes, object properties, data properties, and data types within the subontology. This visualisation provides a clear and comprehensive representation of the subontology's structure and can be used as a valuable reference for researchers and practitioners working in the health and beauty domain.

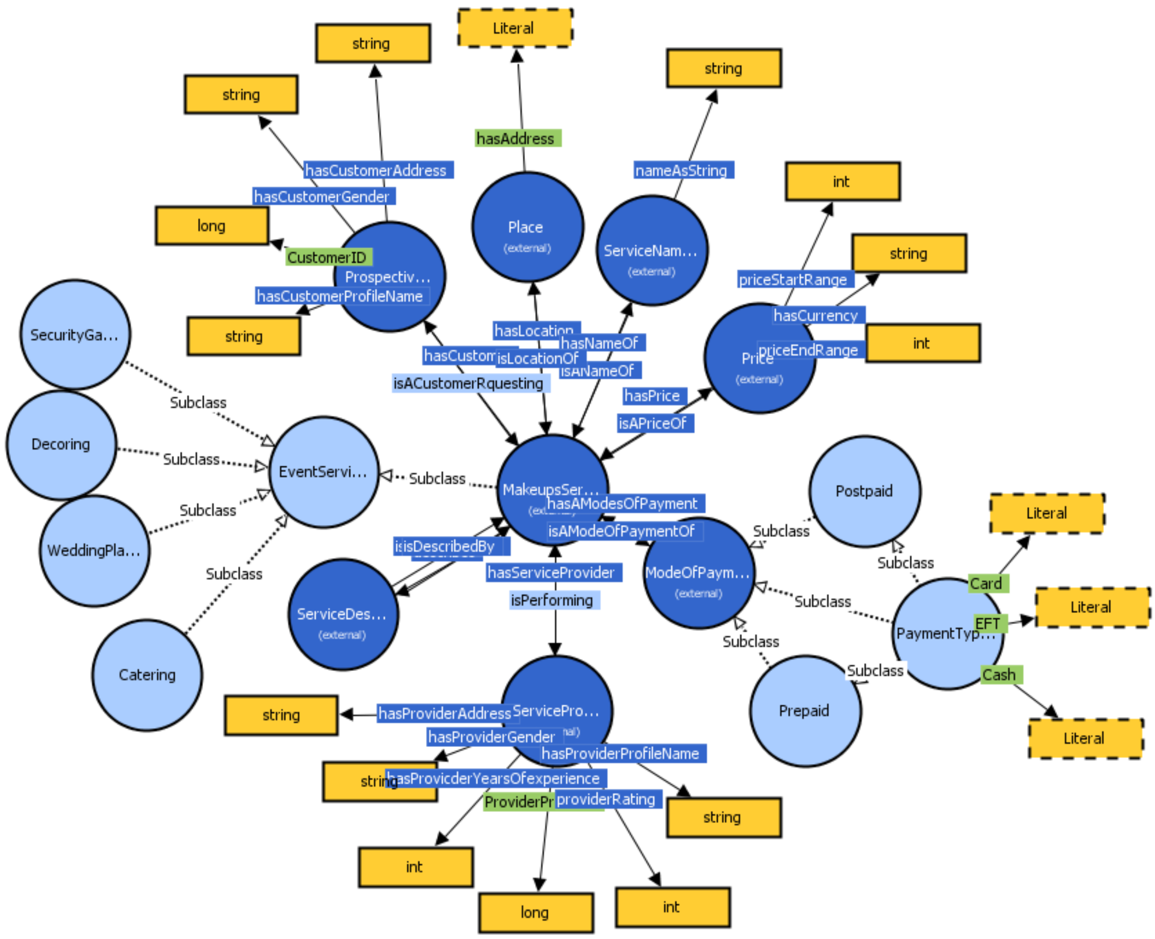

Figure 5.8: Event Service Subontology of SAIBUS

Figure 5.8 presents a visualisation of the Event Service subontology, created using an ontology VOWL tool. This visualisation showcases the intricate relationships between various classes, object properties, data properties, and data types within the Event Service subontology, providing a comprehensive and visually engaging overview of the domain.

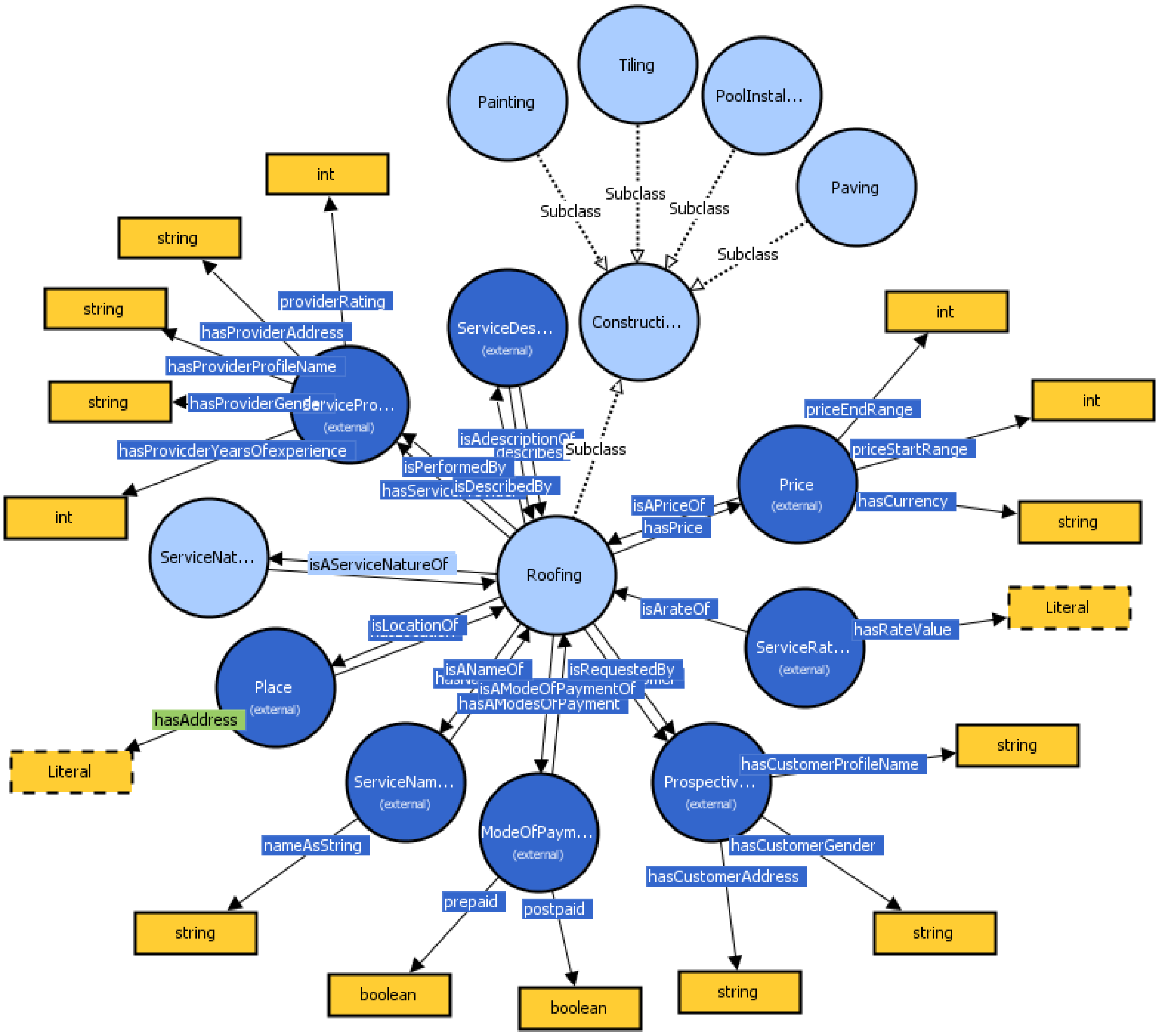

Figure 5.9: Construction Service Subontology of SAIBUS

Figure 5.9 depicts the ontology visualisation of the construction service subontology, generated using the ontology visualisation tool. The visualisation provides a comprehensive overview of the interrelationships between the various classes, object properties, data properties, and data types that are associated with the construction services domain.

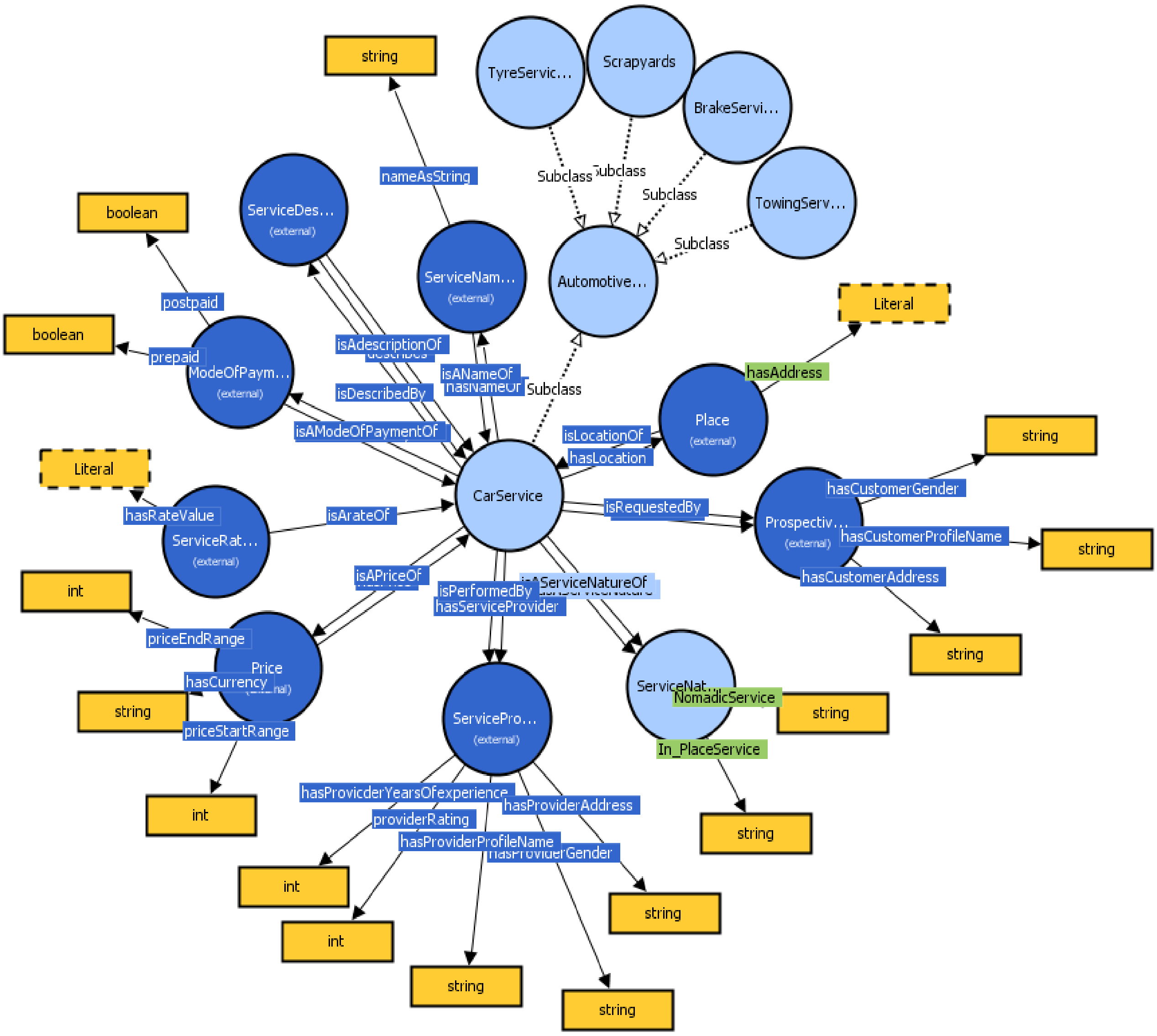

Figure 5.10: Automotive Service Subontology of SAIBUS

Figure 5.10 presents a visualisation of the relationships between classes, object properties, data properties, and data types for the automotive service subontology. The visualisation has been generated using the ontology VOWL tool. The figure provides an overview of the various components of the subontology, including tyre service, scrapyards service, brake service, and towing service. The object properties and data properties that relate to these components are also shown. The visualisation provides a useful tool for understanding the structure of the automotive service subontology and the relationships between its different components.

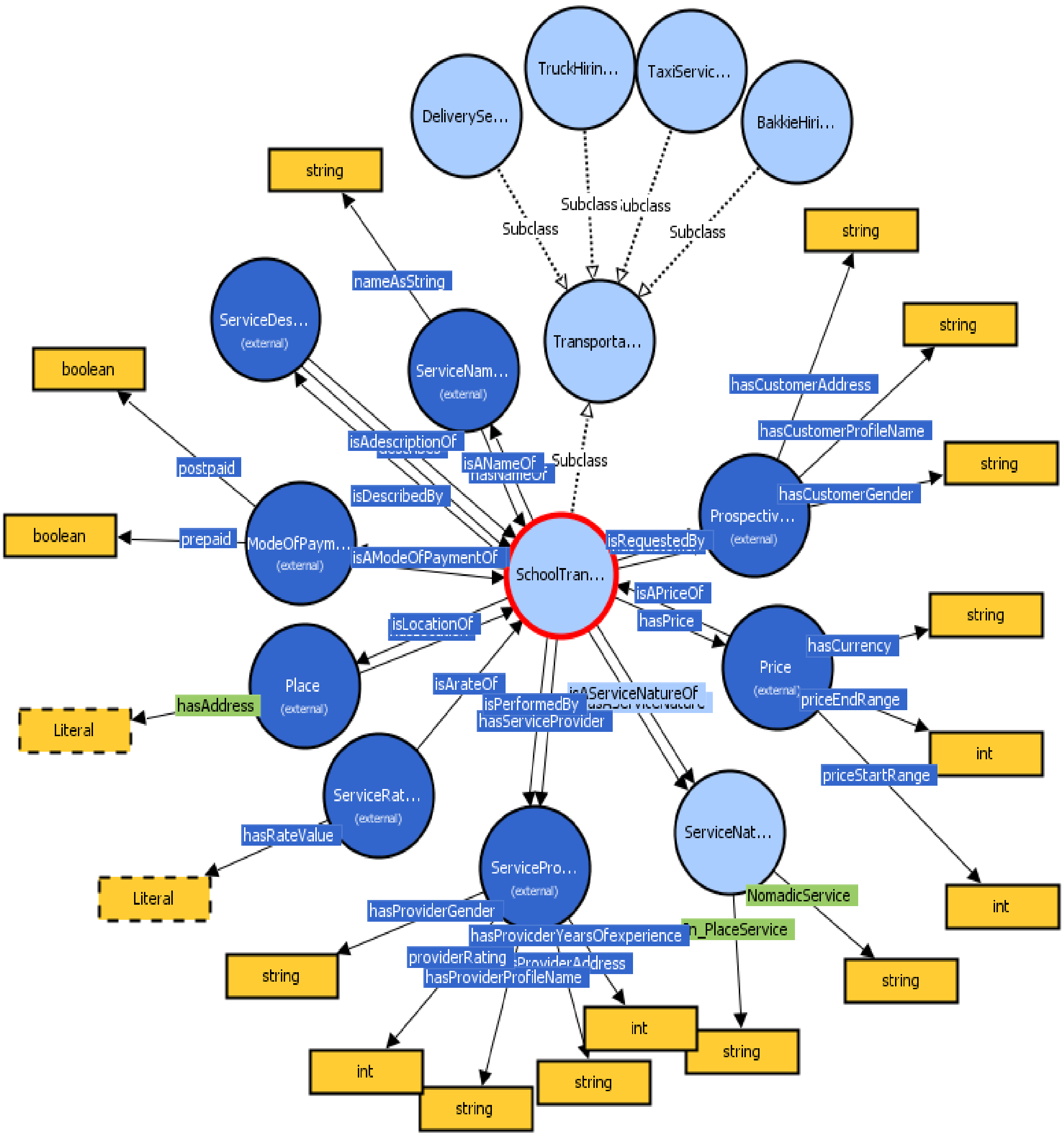

Figure 5.11: Transport Services Subontology of SAIBUS

Figure 5.11 presents the transport services subontology using the ontology VOWL tool. This visualisation provides an overview of the classes, object properties, data properties, and data types that are relevant to the transport services domain. The classes included in this subontology represent various entities involved in the transportation sector, such as courier services, taxi services, and goods transportation. The object properties show how these entities are related to each other, for example, how goods are transported by courier services or how passengers are transported by taxi services. The data properties and data types capture important characteristics of these entities, such as the weight and volume of goods transported or the fare charged for taxi services.

5.1.3 Creating Individuals for the SAIBUS ontology

Individuals, also referred to as instances, are the "core" elements of ontology that are most fundamental. Both real objects like humans, animals, tables, cars, chemicals, and planets and abstract objects like words and numbers can be considered individuals in an ontology.

During the development, some individuals were added to exhibit and evaluate the ontology knowledge for demonstration purposes. Figure 5.12 depicts some of the new instances added to the SAIBUS. *Sandra, for instance, is a Service Provider stored within the SAIBUS ontology, delivering Hair treatment service, which is an instance of Hairdressing service, which is a sub-class of the health and beauty sub-domain. Sandra has all Service Provider attributes such as profileID, profileName, serviceRating, Gender, and YearsOfExperience. The instance of hair treatment then has a service location, which is an in-place service performed at the address 34 Main Road TableView, and hair treatment is a service given by Sandra with a price of R400 and an EFT prepaid payment method. Florence is an instance of a prospective customer who has requested hair treatment service and has all customer attributes such as customerAddress, ProfileName, ProfileID, and Gender*. Figure 5.13-5.17 depicts some of the instances and their interconnections.

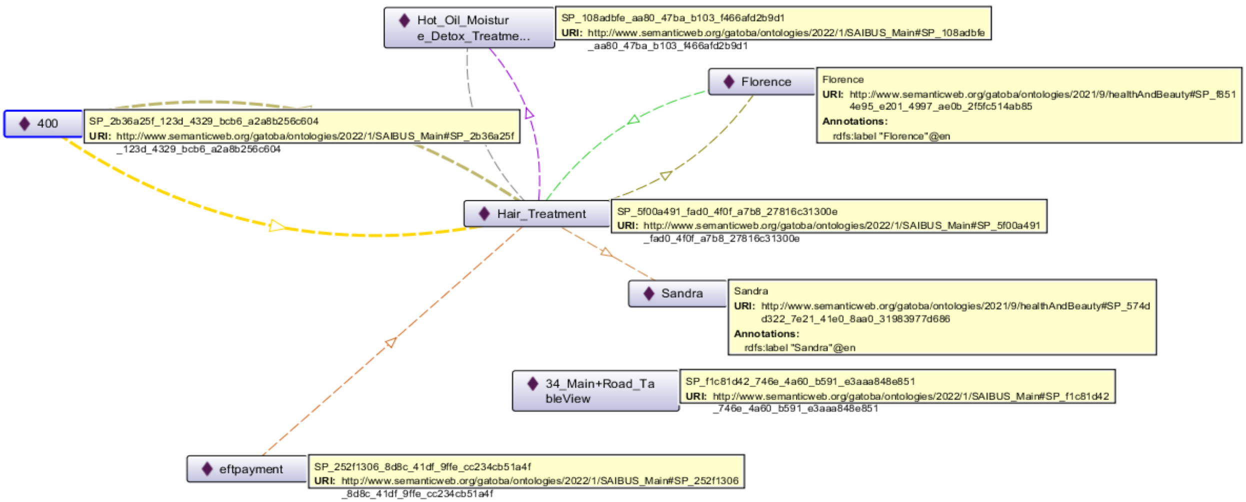

Figure 5.12: SAIBUS Instances and relationships

Figure 5.13 shows the instance relationship between Hair Treatment, an instance of Hairdressing service, and Sandra, an instance of ServiceProvider, by showing that Hair Treatment has a Service Provider named Sandra, using the defined object property hasServiceProvider.

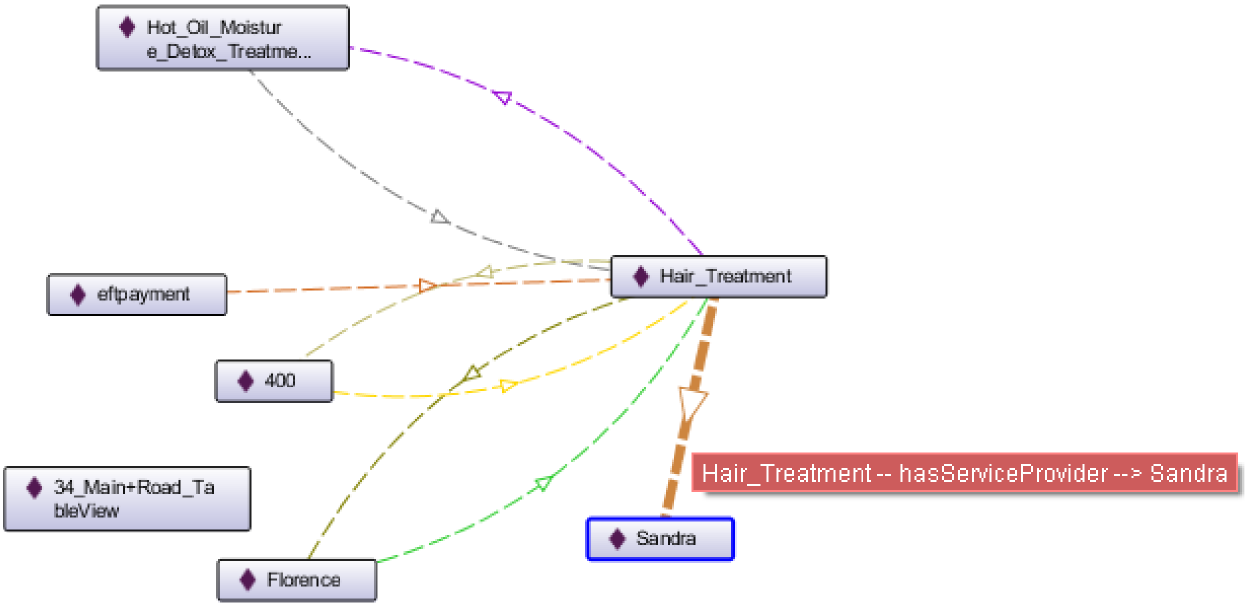

Figure 5.13: hasServiceProvider relationship

Figure 5.14 illustrates another instance relationship between Hair Treatment, an instance of Hairdressing, and Hot oil Moisture Detox Treatment an instance of ServiceDescription, explaining that Hair Treatment is described by Hot oil Moisture Detox Treatment using the defined object property isDescribedBy.

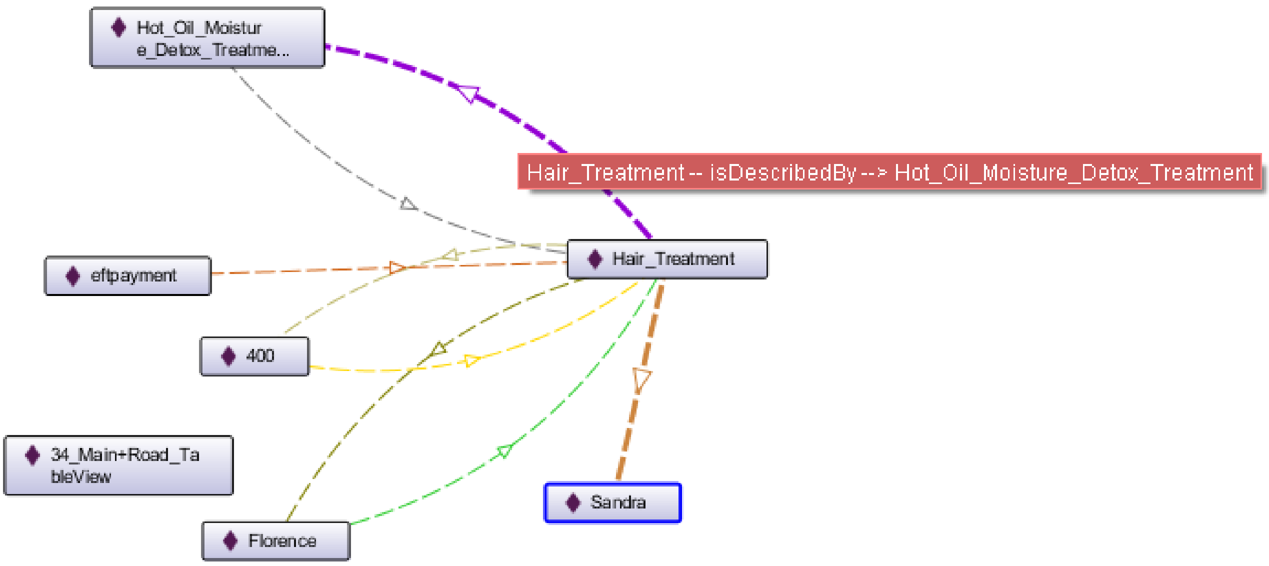

Figure 5.14: isDescribedBy relationship

Figure 5.15 illustrates another relationship between a Hair Treatment, a Hairdressing instance, and an EFT payment, a ModeOfPayment instance. This relationship is demonstrated using the defined object property isAModeOfPaymentOf.

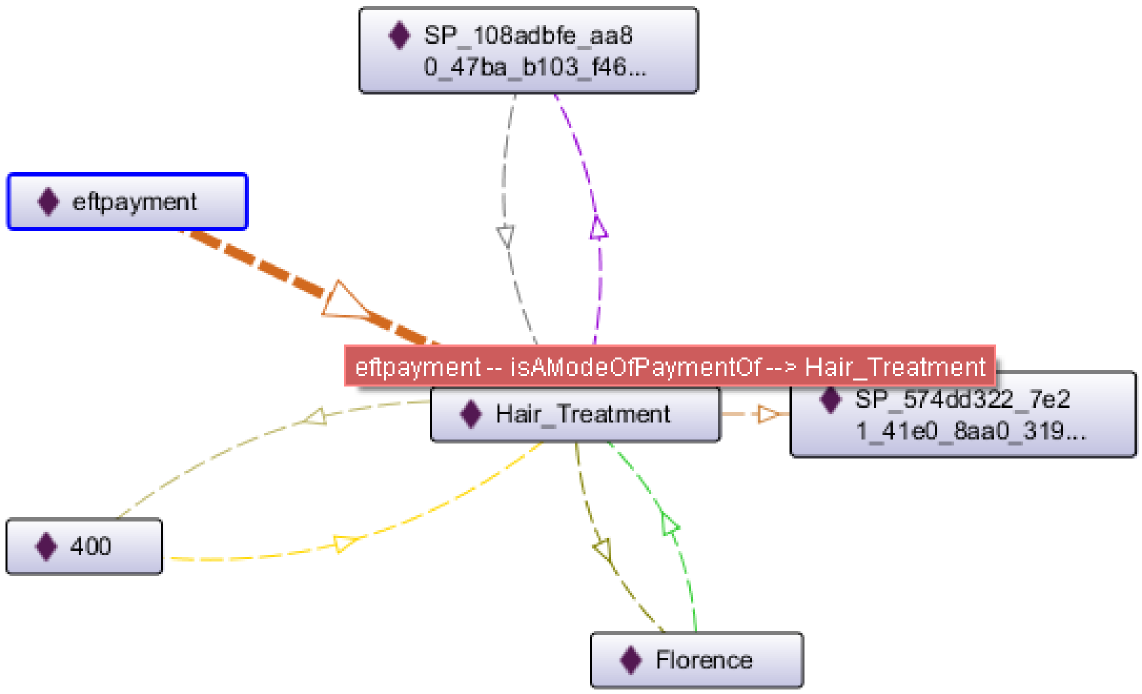

Figure 5.15: isModeOfPaymentOf relationship

Figure 5.16 demonstrates the individual relationship between Hair treatment and 400 as a price of the hair treatment service using the defined isAPriceOF.

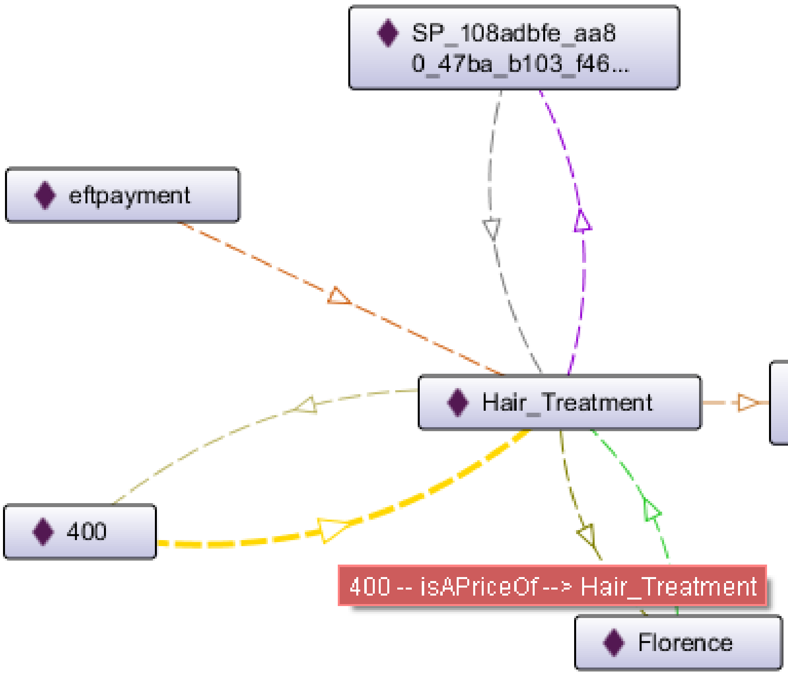

Figure 5.16: isAPriceof relationship

Figure 5.17 shows an instance relationship between Florence, an instance of ProspectiveCustomer, as a customer requesting hair treatment that is an instance of hairdressing service.

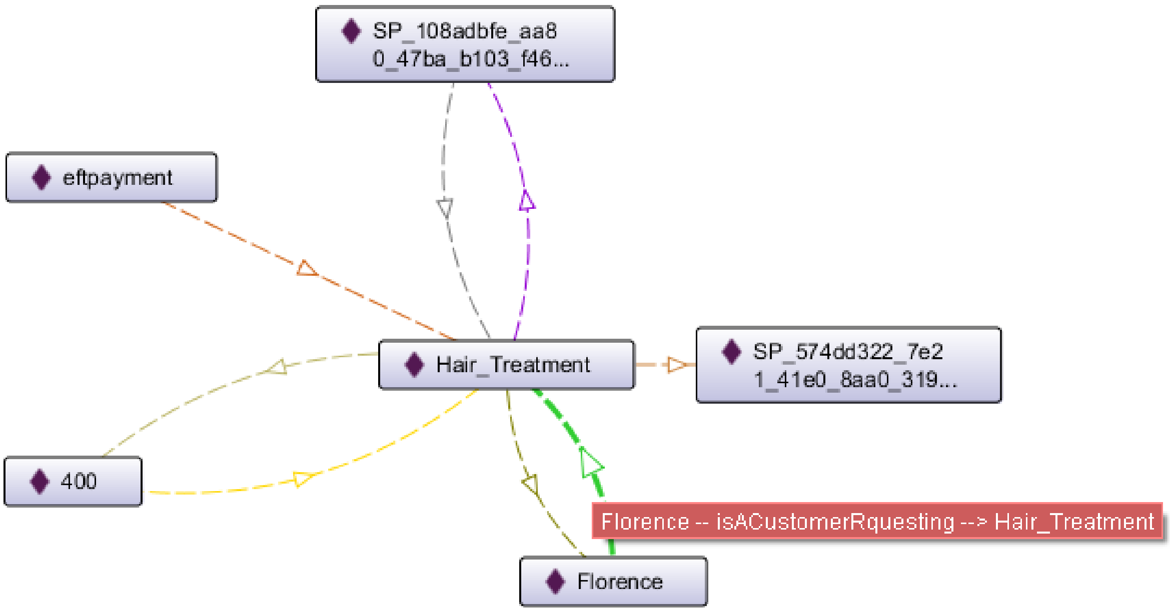

Figure 5.17: isACustomerRequesting relationship

A diagram to illustrate the representation of the SAIBUS ontology was created, and its sub-ontologies were demonstrated in Figure 5.18 and created; an ontology graph was also constructed using the OntoGraf plugin within Protégé shown in Figure 5.19 to represent the SAIBUS minimum viable ontology architecture including classes, objects properties, data properties, as well as some instances. This is due to the limitation of visualisation tools to represent the imported ontologies and their classes, properties, and instances.

The modular ontologies in the SAIBUS architecture that were each individually developed using the modularisation concept and design pattern represent informal sector categories within the SAIBUS. This shows that each modular has classes, objects, and data properties. Each of these modular ontologies is imported into the SAIBUS main ontology using the OWL import mechanism and has shared classes, object properties, data properties, and instances that can be accessed from other classes derived from the modular ontologies. If, for instance, Modular ontology 1 has a class in common with Modular ontology 2, this class automatically becomes a shared class.

As the concepts are built, it also shows the use of SPARQL query to make intelligent searches and find a hairdresser from the SAIBUS ontology. The next section shows how this answered the competency questions and can support intelligent reasoning.

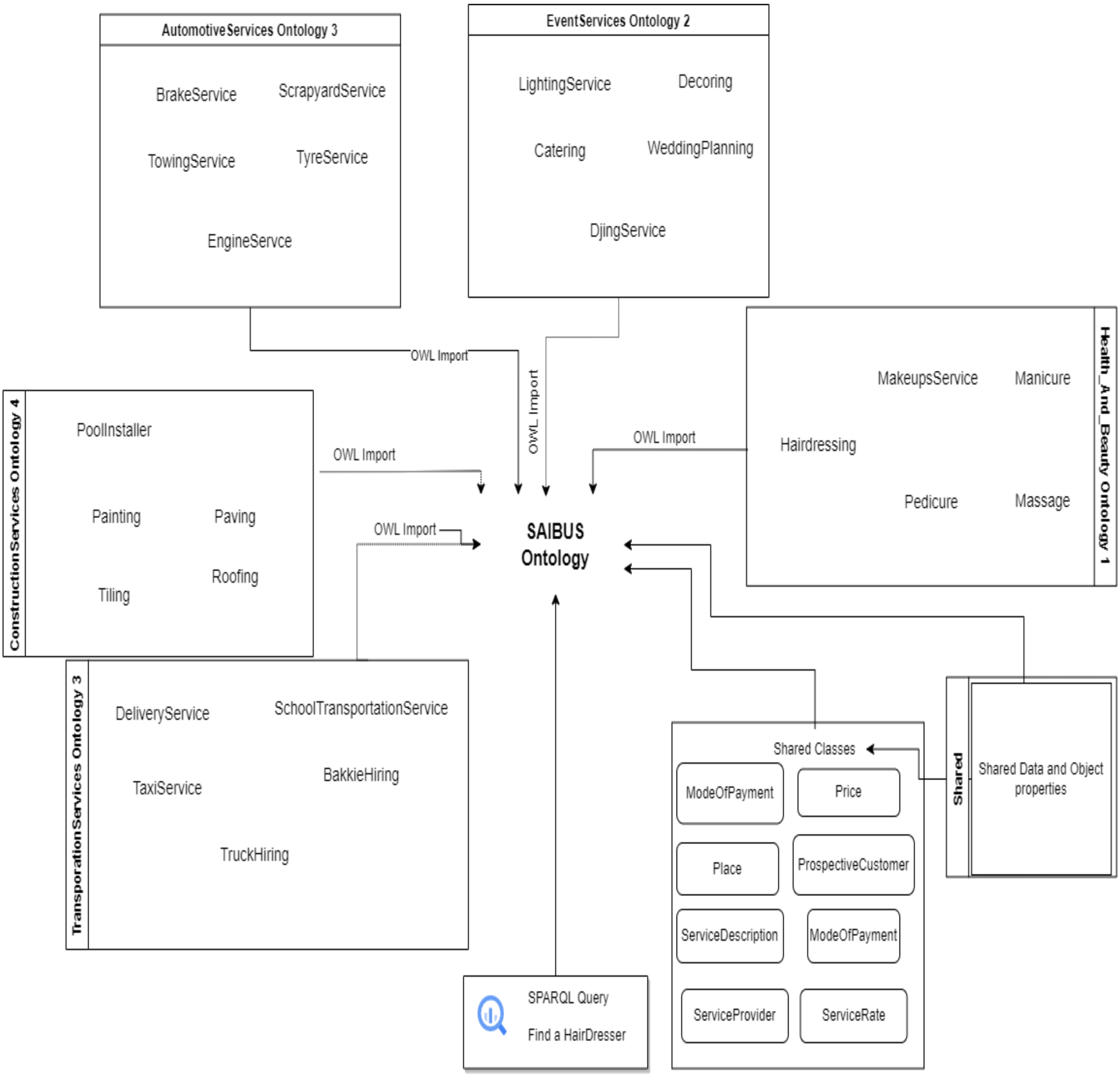

Figure 5.18: SAIBUS MVO Architecture

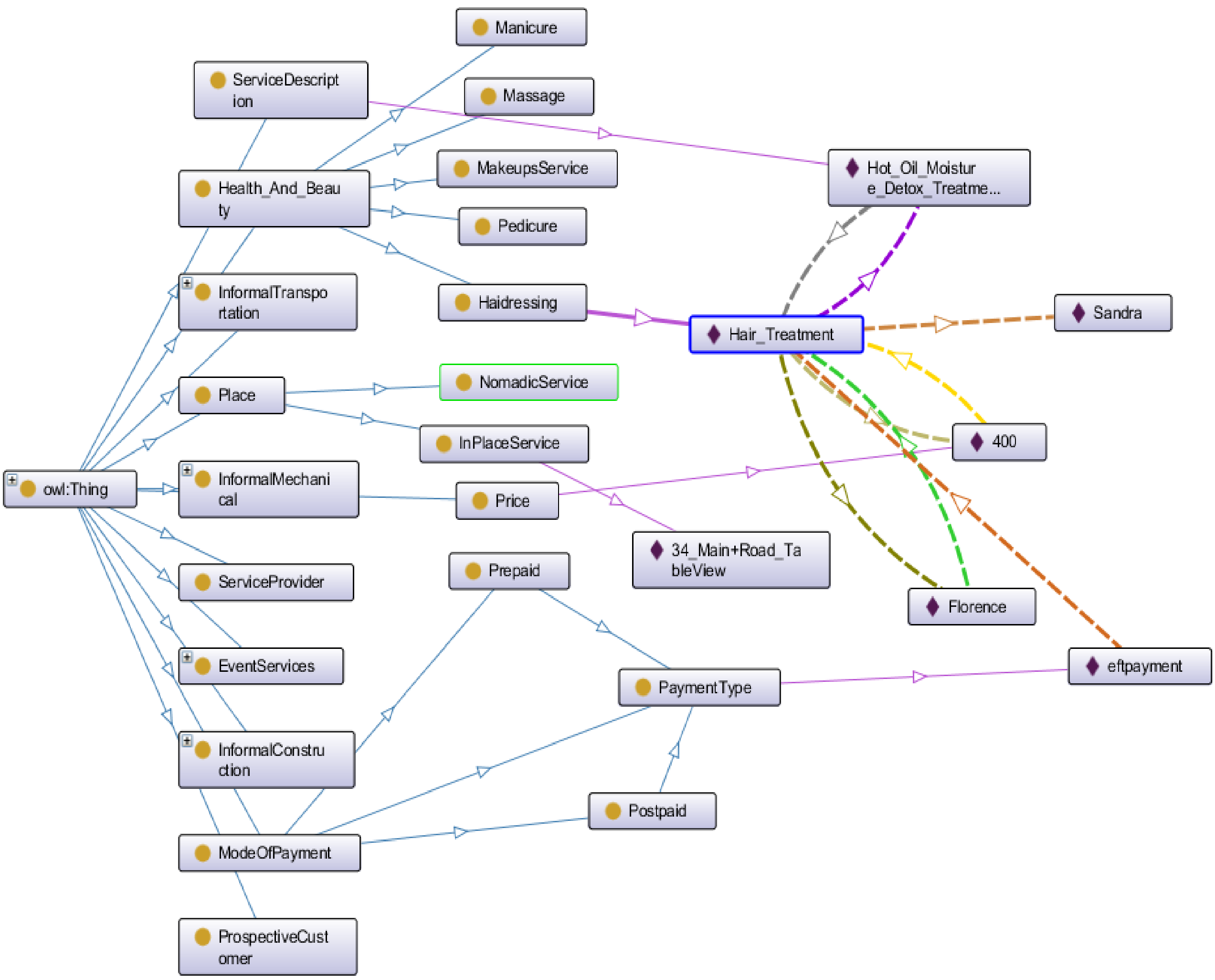

Figure 5.19: SAIBUS ontology graph with instances

5.2 Summary

This chapter presented the implementation of the SAIBUS in which the solution was executed by following the LOD methodologies and by using the import mechanism to import the sub-ontologies of the SAIBUS into the main SAIBUS ontology. After that, individuals (instances) were added to demonstrate and evaluate.

CHAPTER SIX

ONTOLOGY EVALUATION

Before being used or reused in other ontologies or applications, the content of ontologies should be assessed, much like any other resources utilised in software applications. Therefore, before incorporating ontologies into final applications, it is essential to evaluate the ontology's quality and correctness and assess whether it reflects the context in which it was developed (Gomez-Perez, 2001). Evaluating an ontology is crucial for various reasons, including automatically helping the ontology developer discover areas that may require further effort. Ontology validation and verification were performed as part of the evaluation of the SAIBUS ontology. The goal of the validation step was to evaluate the SAIBUS's accuracy and quality. To do this, the OOPS! tool was used, which is covered in 6.1. In 6.2, the validation was conducted to ensure that the meaning appropriately reflected the context for which the ontology was developed. The ontology was verified for compliance with the requirements and competency questions.

6.1 Validation of the SAIBUS Ontology

One way of validating ontologies is to verify whether they adhere to best practices for ontology modelling or, in other words, to determine whether they have any errors or pitfalls. The following ontology evaluation criteria were used in this phase to assess the ontology: consistency, completeness, conciseness, correctness, computational efficiency, and adaptability. An online tool that supports the automatic detection of pitfalls in ontologies is called OOPS! (OntOlogy Pitfall Scanner!) was used. OOPS! helps developers during the validation of the ontology to detect potential pitfalls that might result in modelling issues (Poveda-Villalón, Carmen Suárez-Figueroa and Gómez-Pérez, 2012). Poveda-Villalón, Carmen Suárez-Figueroa, and Gómez-Pérez (2012) presented a list of potential pitfalls, which have been categorised into several ontology quality dimensions.

In the evaluation, the ontology criteria considered as ontology standard metrics were mapped to the pitfalls, along with the description of each pitfall in table 6.1.

Table 6.1: Description of Ontology Validation Metrics of OOPS!

Ontology Category	Metrics	Metric definition	Pitfalls	Pitfalls description
Ontology Correctness	Consistency	Checks that no conflicts are present in or allowed by the ontology.	P05	Define wrong inverse relationships
			P06	Including cycles in a class hierarchy
			P07	Merging different concepts in the same class
	Completeness	Determines if the area of interest is sufficiently covered.	P04	Creating unconnected ontology elements
			P10	Missing disjointness
			P11	Missing domain or range in properties
			P12	Equivalent properties not explicitly declared
			P13	Inverse relationships not explicitly declared
	Conciseness	Focuses on how the ontology excludes or does not permit any inconsistencies.	P02	Creating synonyms as classes
			P03	Creating "is" instead of using rdfs:subClassOf", "rdf:type", or "owl:sameAs"
			P21	Using a miscellaneous class
	Accuracy	This focus on the fact that knowledge must align with the domain knowledge of experts to be evaluated.	P04	Creating unconnected ontology elements
			P10	Missing disjointness
			P9	Missing domain information
Ontology Quality	Computational efficiency	relates to how quickly tools may interact with the ontology.		
	Adaptability	measures the usability of an ontology in many situations		
	Clarity	determines how well the ontology conveys the defined words' intended meaning	P08	Missing annotations
			P22	Using different naming conventions in the ontology

Throughout the development and testing phases, the OOPS! assisted us in identifying and fixing the SAIBUS. Additionally, it contains the prescriptive methodological directions shown in Figure 6.1 for addressing the discovered issues, which helped us improve the SAIBUS until the assessment findings were error-free. For each category, the pertinent criteria were chosen as given in Figure 6.2, and the outcome is displayed in Figure 6.3.

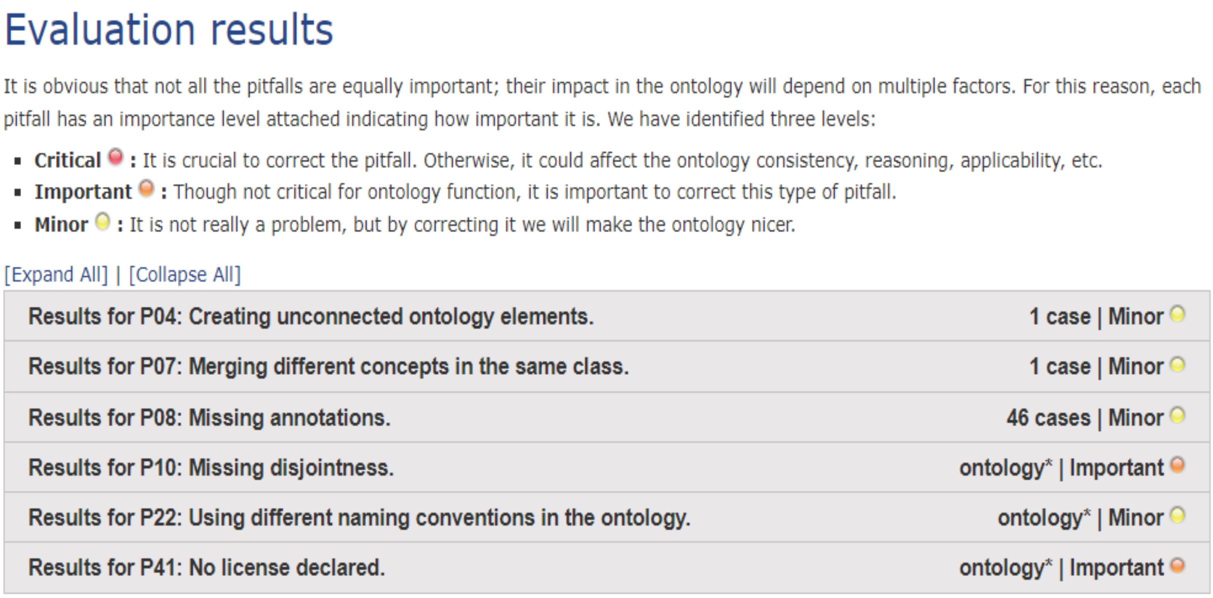

Figure 6.1: SAIBUS Pitfall Prescriptive Directions

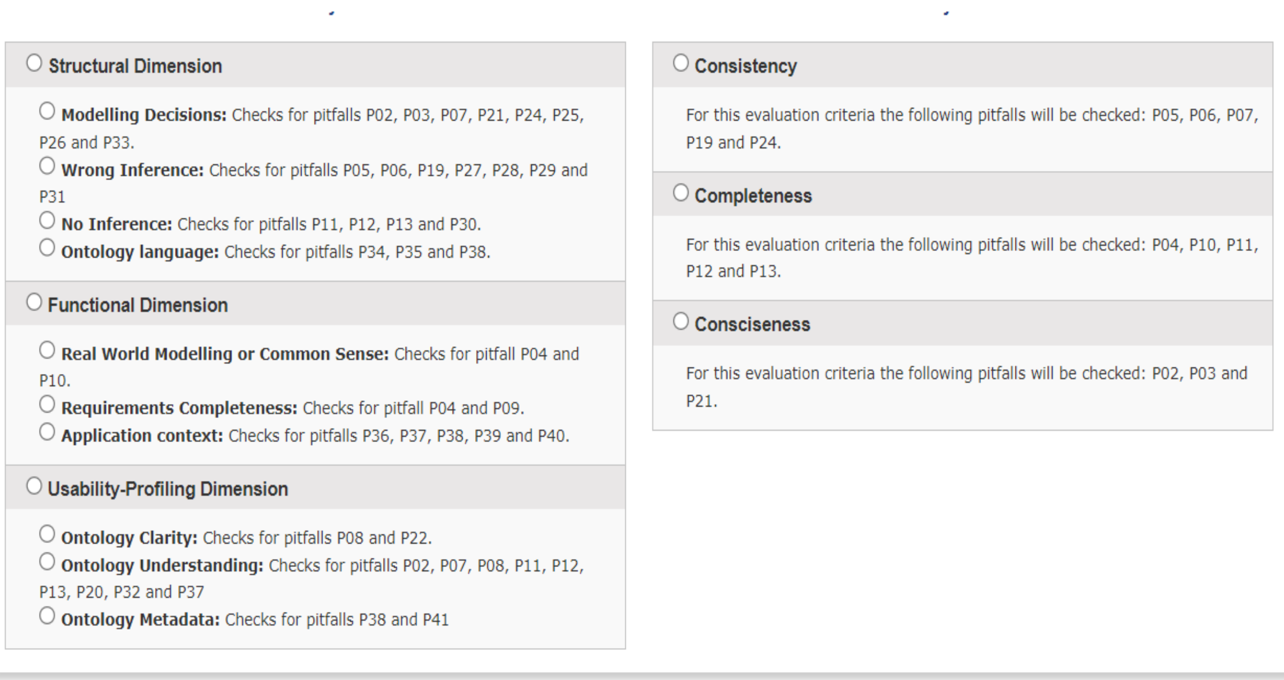

Figure 6.2: SAIBUS Evaluation Criteria

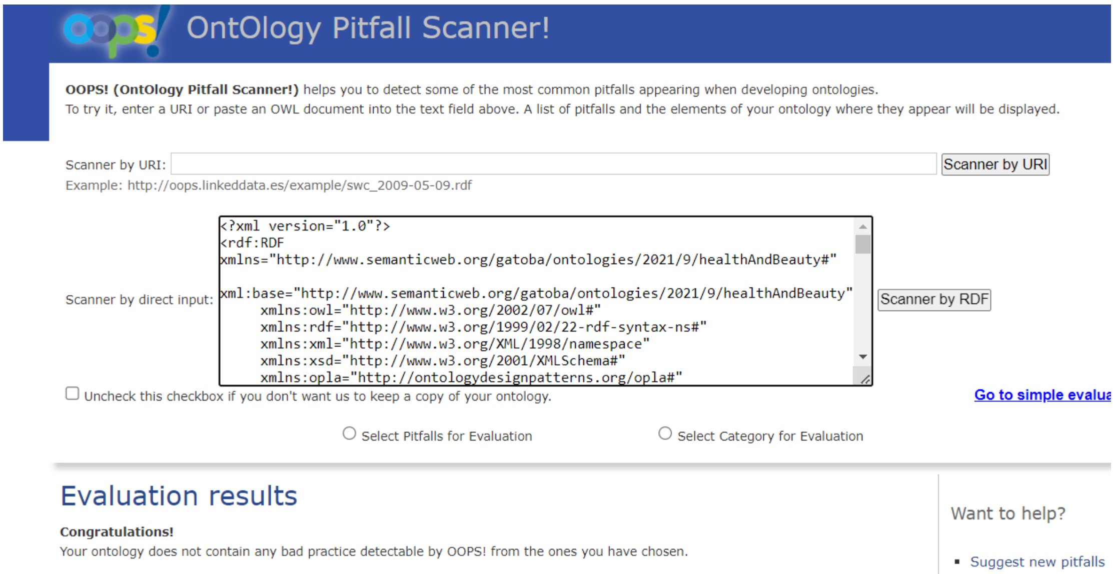

Figure 6.3: SAIBUS Pitfall result

Table 6.2 Summary of evaluation Results

Ontology Category	Metrics	Pitfalls	Score Obtained
Ontology Correctness	Consistency	P05	100%
		P06	100%
		P07	100%
	Completeness	P04	100%
		P10	100%
		P11	100%
		P12	100%
		P13	100%
	Conciseness	P02	100%
		P03	100%
		P21	100%
	Accuracy	P04	100%
		P10	100%
		P9	100%
Ontology Quality	Clarity	P08	100%
		P22	100%

The evaluation results summarised in Table 6.2 validate the ontology against each of its errors and confirm that it complies with the required metrics. After applying a few fixes suggested by the OOPS! prescriptive instructions, the final scores for each pitfall were 100%. In this instance, 100% signifies no errors for the listed pitfalls. This demonstrated that the ontology complied with the standards for consistency, completeness, conciseness, correctness, and clarity.

6.1.1 Validation of Consistency Using HermiT Reasoner

Furthermore, the SAIBUS ontology was validated with the HermiT reasoner in Protégé. HermiT is an ontology reasoner for Web Ontology Language-created ontologies (OWL). HermiT can do many things with an OWL file, such as checking for consistency in the ontology, finding subsumption connections between classes, and more. HermiT uses a variety of optimisations to ensure the effective processing of real-world ontologies. It supports several specialised reasoning services, such as class and property classification, and a range of features outside the OWL 2 standard, such as DL-safe rules, SPARQL queries, and description graphs (Glimm *et al.*, 2010). The HermiT reasoner, which shows a yellow background with protégé, shows that the ontology has no inconsistency. Alternatively, it would provide an explanation and error of the inconsistency within the ontology that would need to be fixed.

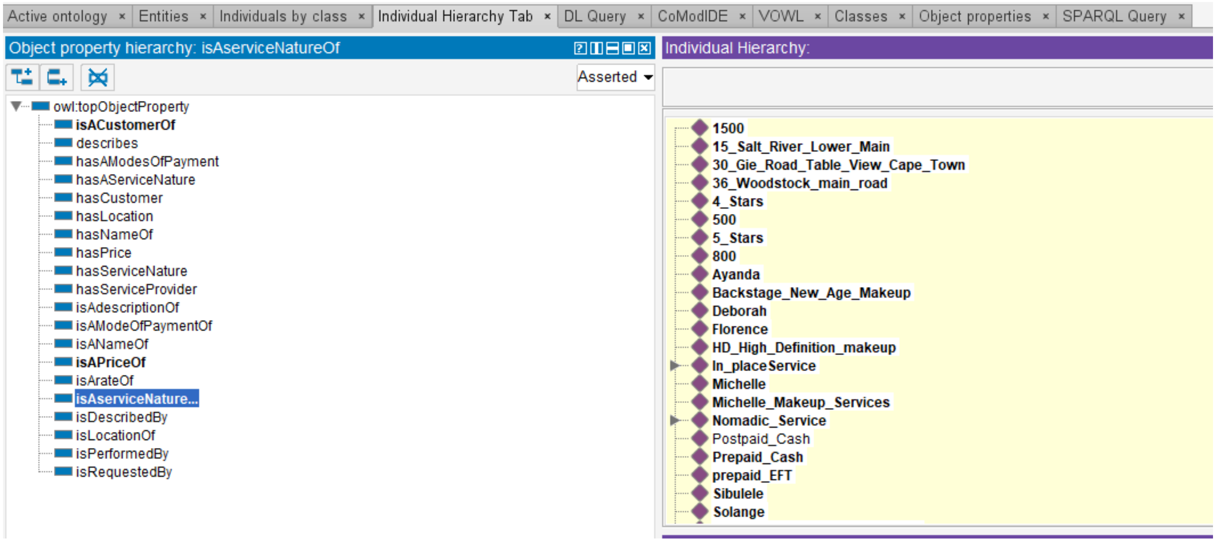

Figure 6.4: HermiT reasoner validation

6.2 Verification of the SAIBUS Ontology

This second evaluation stage was compared to the ontology requirements and the competency questions to assess if the ontology served its intended purpose. To achieve this, the ontology concepts that the SAIBUS covered were mapped to the competency questions and the 6 dimensions defined in Table 4.1, after which the SPARQL query was used to test its ability to retrieve relevant ontological instances. Table 6.2.

Table 6.2: domain ontology concepts mapping to competency questions

Domain Category	Competency questions	Ontology Properties
Informal service attributes	• What is the type of service being provided? • What is the service description? • To which category does the service belong? • In which location is the service? What is the nature of the service (in place service/nomadic service)?	*isAdescriptionOF* *describes* *hasLocation* *hasServiceDescription* *hasServiceCategory* *isLocationOF* *isDescribedBy* *hasNameOf*
	•	
Informal service quality rating	What level of quality is expected from the service?	*isArateOf* *hasARateOf* *hasRateValue*
Informal service pricing	• How much does the service cost? What are the prices for the range of available services?	*isAPriceOf* *hasPrice* *priceStartRange* *priceEndRange* *hasCurrency*
Informal service transactions	What are the modes of payment (prepaid, post-paid) for the service?	*postPaid* *prepaid* *isModeOfPaymentOf* *hasAModeOfPayament*
Prospective customer profile	• Who is requesting the service? What can be known about the customer (address, rating by service providers, gender)	*hasCustomer* *hasCustomerProfileName* *hasCustomerGender* *hasCustomerAddress* *isACustomerRequesting*
Service provider profile	• Who is providing the service? What can be known about the service provider (address, customer rating, gender, years of service)	*hasServiceProvider* *hasProfiderProfileName* *hasProviderAddress* *hasProviderGender* *hasProviderYearsOfExperience* *providerRating*

To verify that the ontology can answer the competency questions, the SPARQL queries were used within Protege to query some instances. The W3C advises using the SPARQL Protocol and RDF Query Language (viz. SPARQL) to describe the RDF (Resource Description Framework) graph, which is a set of triples composed of a subject, a predicate, and an object that serves as the fundamental expression of data stored in an OWL-based knowledgebase

(Dong, Hussain and Chang, 2007). The code below would query data from the SAIBUS ontology, including information on the type of hairdressing service, its description, location, service provider, prospective customers, payment options, and price. Most of the query results returned by SPARQL match the instances inserted into the ontology in Protégé, demonstrating that the ontology is answering the competency questions raised earlier in this thesis**.** The below SPARQL query result from Figure 6.5 - 6.9 shows the result of the six categories of the SAIBUS. Each sub-ontology within the SAIBUS main ontology was queried to verify that it meets the six dimensions mapped to the competency questions.

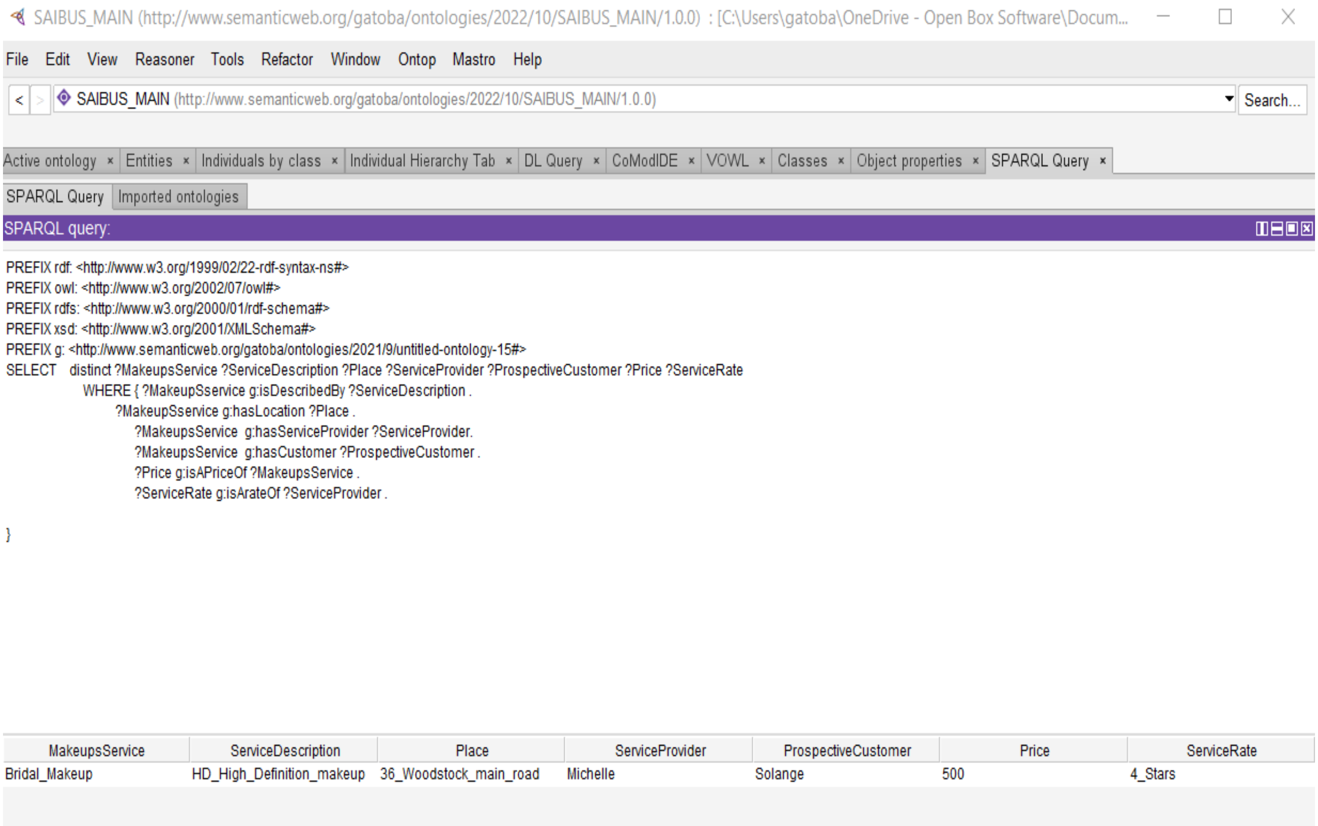

Figure 6.5: Query result of a makeup service

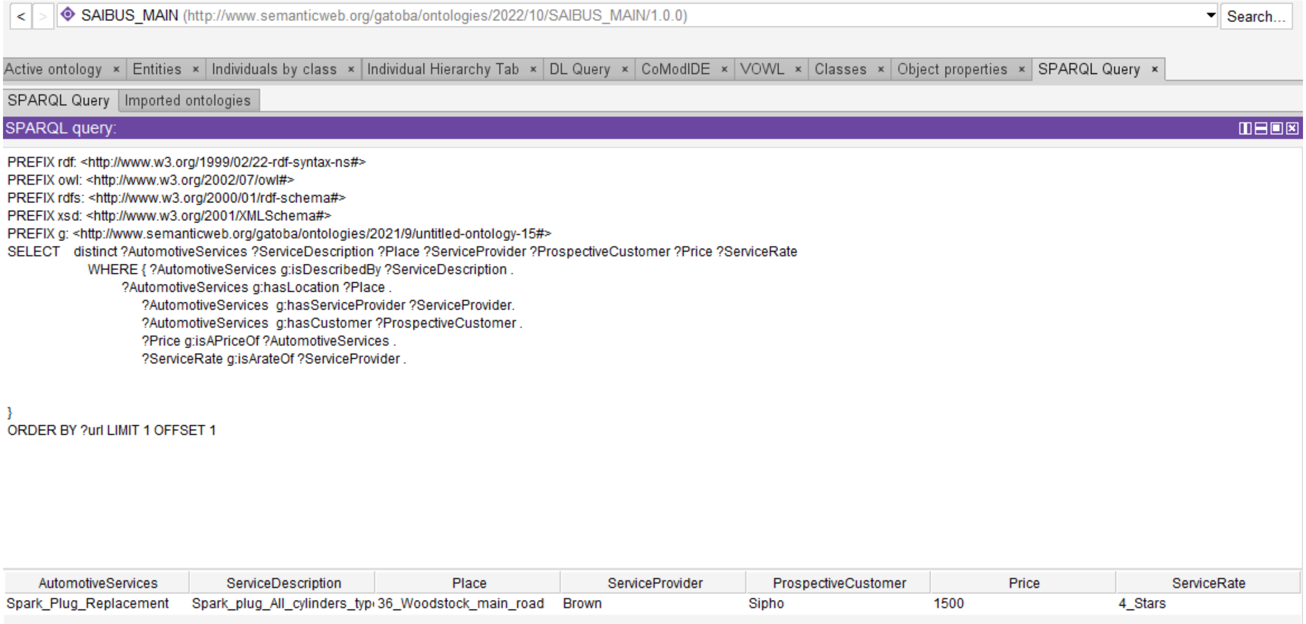

Figure 6.6: Query result of Engine Service

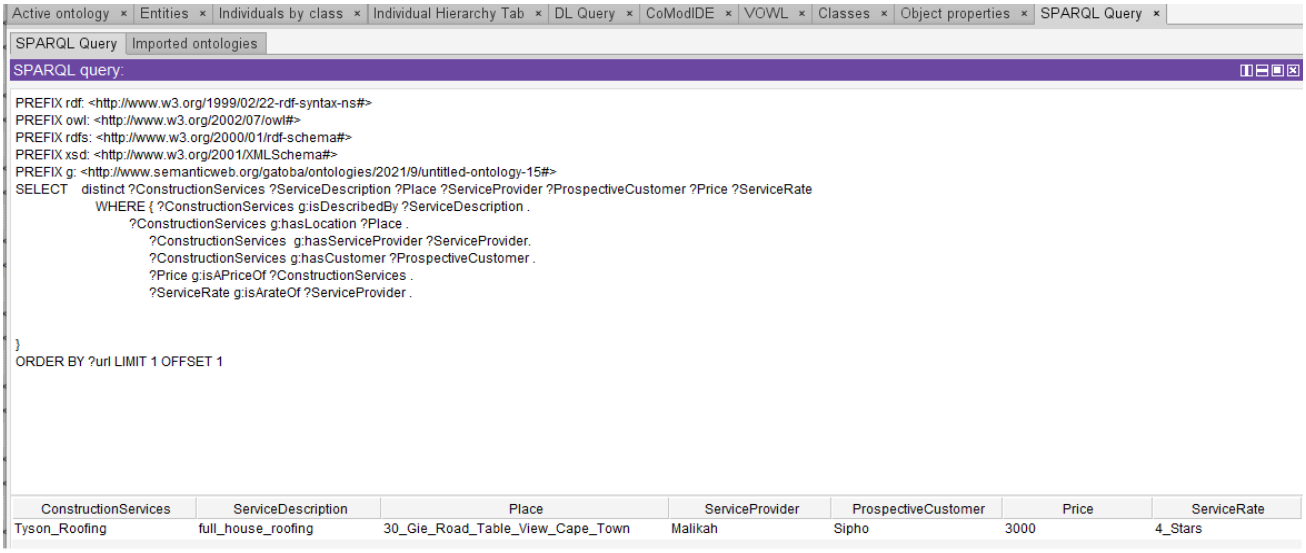

Figure 6.7: Query result of Roofing Service

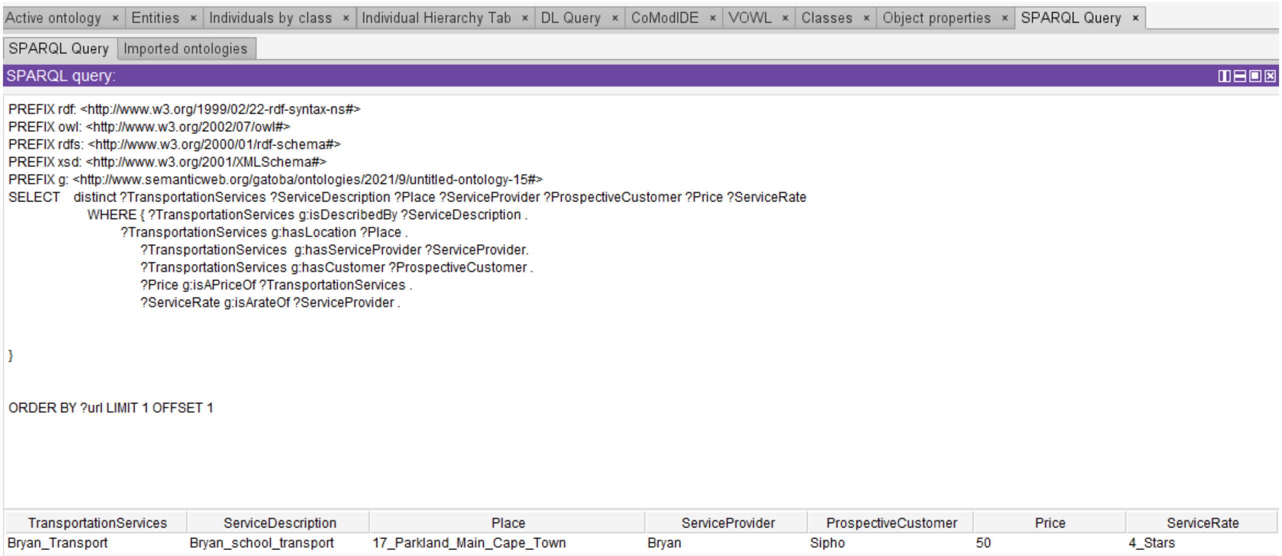

Figure 6.8: Query result of school transportation Service

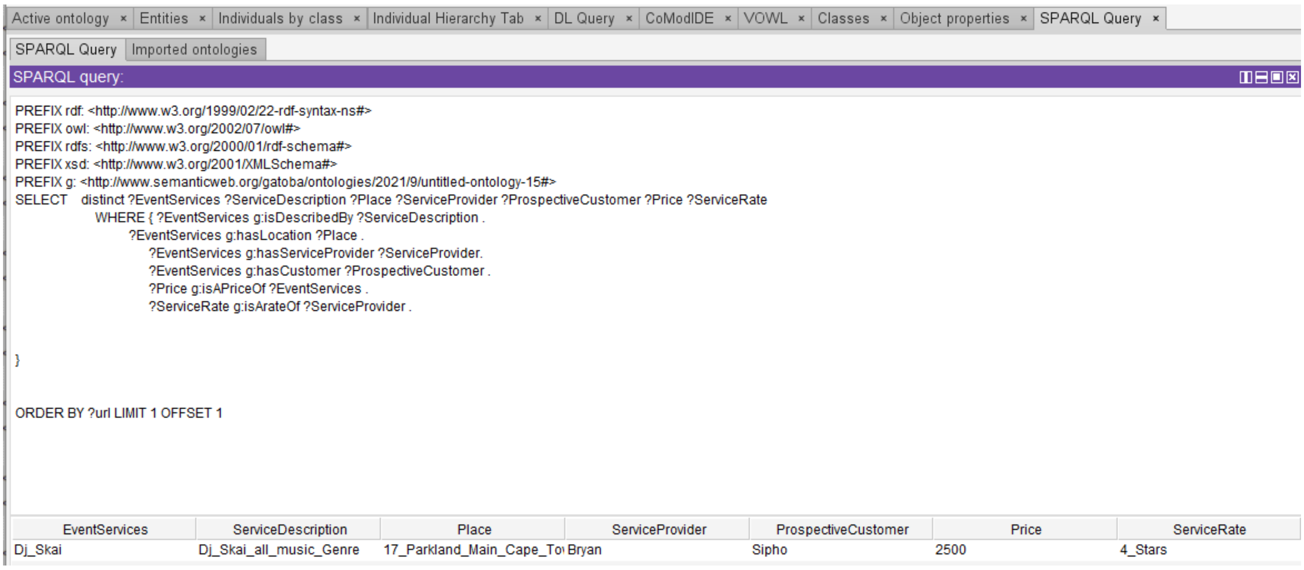

Figure 6.9: Query result of Djing Service

The results show that the SAIBUS ontology constructed a knowledgebase with the 6 dimensions attributes mapped to the competency questions, including Service attributes, service quality rating, service pricing, service transactions, customer profile, and provider profile.

6.3 Summary

In this chapter, the ontology was evaluated using two ontology evaluation approaches, including ontology validation and verification. The ontology was validated using OOPS! Evaluation tools and by satisfying some of the standard metrics. The SPARQL query was used in the verification part to verify if the ontology met its intended purpose.

CHAPTER SEVEN

CONCLUSION AND RECOMMENDATION

The study's recommendations and conclusions are presented in this chapter. There are five sections in the chapter. The study's overview and a brief review of each chapter are presented in section one; the study's benefits and contributions are reviewed in section two; section three contains recommendations section four suggests future research for the study. Lastly, section five drew this chapter to a close.

7.1 Summary of the Study

The study was divided into seven chapters since each part had a different topic. The study was summarised based on the main areas and the objectives section used to achieve the objectives.

First, it discussed the challenges of the South African informal sector and how developing a domain ontology can aid information retrieval in this area. Background information on the study and its limitations and gaps were highlighted by comparing it to existing literature.

The research problem was identified as the lack of knowledgebase to properly equip the web portals with what is needed to build a system that facilitates intelligent reasoning and semantic technology convergence. The study's delineation and significance were then discussed, along with some ethical considerations.

Additionally, the methodology used for this study was explained along with the rationale for the choice.

Objective 1: To determine the requirements for developing a domain ontology for the South African informal business sector.

With the research design on the awareness of the problem, this objective was achieved by reviewing some of the existing literature, and the competency questions were constructed as part of the SAIBUS ontology requirement serving as a basis for the modelling design of the SAIBUS sub-ontologies. This helped to determine what was needed in the iterative stage that the developed solution would provide the answer to and helped the researcher determine steps to take for modelling the SAIBUS ontology by considering the modularisation method.

Objective 2: To design and develop a domain ontology for the South African informal business sector.

With the research design approach, the LOD methodology was followed by developing a minimum viable ontology using the protégé ontology editor to achieve this objective. In this phase, the LOD principles were applied, and the import mechanism helped import the sub-ontologies into the main ontology to form the SAIBUS ontology.

Objective 3: To evaluate the quality attributes of a domain ontology for the South African informal business sector.

To achieve this objective with the research design approach, the artefact was evaluated using the two-ontology evaluation, including ontology validation and verification. In ontology validation, the OOPS! tool was used to evaluate the SAIBUS ontology against existing metrics. In the verification part, the verification was conducted by assessing if the developed ontology answered the competency questions by using the SQARL query to query the ontology knowledge.

7.2 Contribution of the Research

This work has contributions from theoretical, methodological, and practical perspectives.

Theoretical contribution

Theoretically, this study improves the theory and contributes to the literature by specifying the requirements for constructing a domain ontology for the informal business sector in South Africa, which is a first attempt to the best of the researcher's knowledge.

Methodological contribution

This research contributes methodologically by providing an approach combining Lean Ontology Development (LOD) and Modular ontology development methods and evaluating approaches to develop a domain ontology for the South African informal business sector. This approach can also be used to develop more ontologies or sub-ontologies that can be integrated into the domain ontology for the South African Informal Business Sector (SAIBUS).

Practical contribution

This study aimed to create a domain ontology that can support information retrieval and promote intelligent reasoning on services in the informal business sector in South Africa. Therefore, the practical contribution is the domain ontology artefact that was developed which can be shared and reused to empower the web portal of the South African informal sector with intelligent reasoning.

7.3 Recommendation

The LOD methodology, an iterative approach that supports the development of what is needed in the initial iteration, was used in this research to offer a first attempt at developing the SAIBUS. Due to time constraints, conducting a community evaluation which is part of the learn phase of the LOD Methodology, was impossible. Ordinarily, a community evaluation should last for 6 months to enable the effective participation of relevant stakeholders. The community evaluation is recommended for future studies to gather additional requirements for the next iteration of the development of the SAIBUS. Furthermore, to expand the SAIBUS ontology's features, further sub-ontologies should be developed to include more informal sector categories.

Additional software development effort is also needed to develop web and mobile applications that use the SAIBUS as their knowledgebase when performing intelligent search operations. Some existing digital platforms could also be integrated with the SAIBUS ontology to enhance their search capabilities.

7.4 Further Research

The ontology will usually be applied to some sort of application. The application's outputs or performance on the assigned job may be better or worse based in part on the ontology it uses. Therefore, one may contend that a good ontology aids the application in producing positive outcomes for the assigned goal (Jožef et al., 2009; Raad and Cruz, 2015). Therefore, ontologies may be assessed by simply integrating them into an application and analysing the output of the application. So, further research is needed to cover the application-based ontology evaluation and ascertain the in-use qualities of the SAIBUS ontology.

Since many persons working in South Africa's informal economy do so in multilingual settings, the SAIBUS ontology may have challenges supporting information retrieval in this industry due to the language barrier. Future research is therefore required to explore the SAIBUS's ability to handle operations in several languages (Espinoza, Gómez-Pérez and Mena, 2008).

Further study is also needed to expand on the automatic ontology population of the SAIBUS ontology. The Ontology Population is tasked with finding instances of an ontology's non-taxonomic connections and features. It is expensive and time-consuming to hire knowledge engineers. As a result, automated or semi-automatic methods are required (Faria, Girardi and Novais, 2012).

Furthermore, there is also a need to conduct more research to assess the impact of the SAIBUS ontology on the informal Sector of South Africa.

7. References

1. Abdullah Kamal, S.S.L.B. 2019. research paradigm and the philosophical foundations of a qualitative study. people: International Journal of Social Sciences, 4(3): 1386–1394. https://grdspublishing.org/index.php/people/article/view/749.
2. AldoGangemi. 2010. Ontology Design Pattern Submissions:Price.
3. Alsanad, A.A., Chikh, A. & Mirza, A. 2019. A Domain Ontology for Software Requirements Change Management in Global Software Development Environment. IEEE Access, 7: 49352–49361.
4. Arpinar, I.B., Sheth, A., Aleman-Meza, B., Tartir, S., Budak Arpinar, I., Moore, M. & Sheth, A.P. 2005. OntoQA: Metric-Based Ontology Quality Analysis. http://www.cia.gov/cia/publications/factbook.
5. Bakhshandeh, M., Antunes, G., Mayer, R., Borbinha, J. & Caetano, A. 2013. A modular ontology for the enterprise architecture domain. In Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOC. Institute of Electrical and Electronics Engineers Inc.: 5–12.
6. Blanche, M.T., Durrheim, K. & Painter, D. eds. 2006. Research in practice: Applied Methods for the Social Sciences. second edition. Cape Town: University of Cape Town.
7. Blomqvist, E., Hitzler, P., Janowicz, K., Krisnadhi, A., Narock, T. & Solanki, M. 2015. Considerations regarding Ontology Design Patterns Editorial. IOS Press. http://schema.geolink.org/patterns/.
8. Cheng, H., Lu, Y.C. & Sheu, C. 2009. An ontology-based business intelligence application in a financial knowledge management system. Expert Systems with Applications, 36(2 PART 2): 3614–3622.
9. Creswell, J.W. 2009. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Third Edition. New Delhi: SaGE Publications LTD.
10. Cummings, J. & Stacey, D. 2018. Lean ontology development: An ontology development paradigm based on continuous innovation. In IC3K 2018 - Proceedings of the 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management. SciTePress: 367–374.
11. Daramola, O. 2018. Towards a multimodal portal framework in support of informal sector businesses. In Communications in Computer and Information Science. Springer Verlag: 110–119.

12. Dong, H., Hussain, F.K. & Chang, E. 2007. Application of Protégé and SPARQL in the Field of Project Knowledge Management. Personalized Multi-dimensional process framework View project Sensor Cloud Service Composition View project Application of Protégé and SPARQL in the Field of Project Knowledge Management. https://www.researchgate.net/publication/221255025.
13. Doran, P. 2009. Ontology Modularization: Principles and Practice.
14. Duque-Ramos, A., Tomas Fernandez-Breis, J., Stevens, R.D., Aussenac-Gilles, N., Tomás Fernández-Breis, J. & Stevens, R. 2011. OQuaRE: A SQuaRE-based approach for evaluating the quality of ontologies Open PHACTS View project Call For Paper: Special Issue: Application of Ontologies and Semantic Web Technologies in Biomedical Science View project OQuaRE: A SQuaRE-based Approach for Evaluating the Quality of Ontologies. https://www.researchgate.net/publication/285665382.
15. Espinoza, M., Gómez-Pérez, A. & Mena, E. 2008. Enriching an ontology with multilingual information. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 333–347.
16. Etim, E. & Daramola, O. 2020. The informal sector and economic growth of South Africa and Nigeria: A comparative systematic review. Journal of Open Innovation: Technology, Market, and Complexity, 6(4): 1–26.
17. Faria, C., Girardi, R. & Novais, P. 2012. Using domain specific generated rules for automatic ontology population. In International Conference on Intelligent Systems Design and Applications, ISDA. 297–302.
18. Fernández López, M. & López, F. 1991. Overview Of Methodologies For Building Ontologies. http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-18/.
19. Fernández-López, M., Juristo, N., Ferndndez, M. & G6mez-P~rez, A. 1997. METHONTOLOGY: from ontological art towards ontological engineering Empirical Assesment of Test-driven Development View project ML + Linked Data View project METHONTOLOGY: From Ontological Art Towards Ontological Engineering. https://www.researchgate.net/publication/50236211.
20. Fikes, R. & Finin, T. 1991. Enabling Technology for Knowledge Sharing. https://www.researchgate.net/publication/220605114.
21. Finin, T., Mayfield, J., Joshi, A., Cost, R.S. & Fink, C. 2005. Information retrieval and the semantic web. In Proceedings of the Annual Hawaii International Conference on System Sciences. 113.
22. Freitas, F., Santana Da Silva, F., Bezerra, C. & Santana, F. 2013. CQChecker: A Tool to Check the Satisfaction of Description Logic Competency Questions on Ontologies.

CQChecker: A Tool to Check the Satisfaction of Description Logic Competency Questions on Ontologies. https://www.researchgate.net/publication/290105569.
23. Gangemi, A. 2005. Ontology design patterns for semantic web content. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 262–276.
24. Glimm, B., Horrocks, I., Motik, B., Stoilos, G. & Wang, Z. 2010. HermiT: An OWL 2 Reasoner.
25. Gomez-Perez, A. 2001. Ontology Evaluation. http://www.ontoweb.org/.
26. Gruber, T.R. 1993. Substantial revision of paper presented at the International Workshop on Formal Ontology.
27. Grüninger, M. & Fox, M.S. 1995. The Role of Competency Questions in Enterprise Engineering. In 22–31.
28. Guarino, N. & Welty, C. 2009. Towards a methodology for ontology-based model engineering Cite this paper Related papers A formal ont ology of propert ies Towards a methodology for ontology based model engineering. http://www.ladseb.pd.cnr.it/infor/ontology/ontology.html.
29. Gumtree South Africa. 2022. Services. [Online]. Available at: https://www.gumtree.co.za/.
30. Hevner, A.R., March, S.T., Park, J. & Ram, S. 2004. DESIGN SCIENCE IN INFORMATION SYSTEMS RESEARCH 1.
31. Hitzler, P. & Rudolph, S. 2012. OWL 2 Web Ontology Language Primer (Second Edition). http://www.w3.org/TR/owl-primer.
32. Jiyane, G. v, Majanja, M.K., Mostert, B.J. & Ocholla, D. 2013. South Africa as an information and knowledge society : the benefit to informal sector women entrepreneurs. , 79(1): 1–13.
33. Jožef, J.B., Institute, S., Brank, J., Grobelnik, M. & Mladenić, D. 2009. A survey of ontology evaluation techniques X5GON: Cross Modal, Cross Cultural, Cross Lingual, Cross Domain, and Cross Site Global OER Network View project OPTIMUM View project A SURVEY OF ONTOLOGY EVALUATION TECHNIQUES. https://www.researchgate.net/publication/228857266.
34. Kapoor, B. & Sharma, S. 2010. A Comparative Study of Ontology building Tools in Semantic Web Applications. International journal of Web & Semantic Technology, 1(3): 1–13.
35. Keet, C.M. 2020a. An Introduction to Ontology Engineering.
36. Laumer, S. & Eckhardt, A. 2012. Why Do People Reject Technologies: A Review of User Resistance Theories. In 63–86.

37. Legodi, K. & |kanjere. 2015. The Challenges Faced by Informal Traders in Greater Letaba Municipality in Limpopo Province, South Africa.
38. León, H.C.M. & Farris, J.A. 2011. Lean product development research: Current state and future directions. EMJ - Engineering Management Journal, 23(1): 29–51.
39. Li, J. & Larsen, K.R. 2017. unlocking our behavioral knowledge inheritance through ontology learning: a design framework, an instantiation, and a randomized experiment.
40. Locanto. 2022. Services in South Africa. URL: https://www.locanto.co.za/. Access date: 30 October 2022.
41. Lozano-Tello, A. & Gómez-Pérez, A. 2004. ONTOMETRIC: A Method to Choose the Appropriate Ontology. Journal of Database Management, 15(2): 1–18.
42. March, S.T. & Smith, G.F. 1995. Design and natural science research on information technology. Decision Support Systems, 15(4): 251–266.
43. Mazimwe, A., Hammouda, I. & Gidudu, A. 2019. Ontology Design Patterns for Representing Knowledge in the Disaster Risk Domain. In Proceedings - 2019 IEEE 28th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises, WETICE 2019. Institute of Electrical and Electronics Engineers Inc.: 283–288.
44. Mazumder, A., Mukherjee, A., Adhikari, A. & Dutta, A. 2017. Proc. of the 2017 IEEE Region 10 Conference (TENCON), Malaysia, November 5-8, 2017. IEEE.
45. Mazumder, A., Mukherjee, A.J., Adhikari, A. & Dutta, A. Proc. of the 2017 IEEE Region 10 Conference (TENCON), Malaysia, November 5-8, 2017.
46. Mccarthy, J., Vasiliu, L., Grody, A., Muckley, C., Lawrence, D., Zervoudakis, F., Tabet, S., Grondelle, J., Bouras, T., Fernandes, K., Financial, al, McCarthy, J., Grody, A.D., van Grondelle, J., Krintas, T., Xidonas, P., Chourdakis, K. & Koumpis, A. 2013. Industry Ontologies for Risk and Regulation Data (FIORD)-A Position Paper. : 737–744. https://hal.inria.fr/hal-01463269.
47. McDaniel, M. & Storey, V.C. 2019. Evaluating domain ontologies: Clarification, classification, and challenges. ACM Computing Surveys, 52(4).
48. Mramba, N., Apiola, M., Sutinen, E., Haule, M., Klomsri, T. & Msami, P. Empowering Street Vendors through Technology: an Explorative Study in Dar es Salaam, Tanzania.
49. Musen, M.A. 2015. The protégé project. AI Matters, 1(4): 4–12.
50. National Regulator for Compulsory Specifications. 2018. Annual Performance Plan for 2018/19-2020/21.
51. Ndabeni, L.L., Maharajh, R. & Economy, K. 2013. The informal sector and the challenges of development in south africa. , (February): 19–20.

52. Nel, E. & Rogerson, C.M. 2016. The contested trajectory of applied local economic development in South Africa. Local Economy, 31(1–2): 109–123.
53. Nicola, G., Daniel, O. & Steffen, S. 2009. Handbook on Ontologies. Springer Berlin Heidelberg.
54. Noy, N.F. & Mcguinness, D.L. Ontology Development 101: AGuide to Creating Your First Ontology. http://translate.googleusercontent.com/translate_c?anno=2&hl=…opment/ontology101.pdf&usg=ALkJrhjVKWkRnOxyPTCMqp98iRSKKnZ9HQ.
55. O'connor, M. Importing Data into Protégé-OWL.
56. Offermann, P., Levina, O., Schönherr, M. & Bub, U. 2009. Outline of a design science research process. In Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology, DESRIST '09.
57. Osman, I., ben Yahia, S. & Diallo, G. 2021. Ontology Integration: Approaches and Challenging Issues. Information Fusion, 71: 38–63.
58. Petrova, G.G., Tuzovsky, A.F. & Aksenova, N. v. 2017. Application of the Financial Industry Business Ontology (FIBO) for development of a financial organization ontology. In Journal of Physics: Conference Series. Institute of Physics Publishing.
59. Poveda, M., Suárez-Figueroa, M.C. & Gómez-Pérez, A. 2010. Common pitfalls in ontology development. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 91–100.
60. Poveda-Villalón, M., Carmen Suárez-Figueroa, M. & Gómez-Pérez, A. 2012. LNAI 7603 - Validating Ontologies with OOPS! https://moki.fbk.eu/website/index.php.
61. Raad, J. & Cruz, C. 2015. A Survey on Ontology Evaluation Methods. https://hal.archives-ouvertes.fr/hal-01274199.
62. Rancic Moogk, D. 2012. Minimum Viable Product and the Importance of Experimentation in Technology Startups. www.timreview.ca.
63. Rector, A., Brandt, S., Drummond, N., Horridge, M., Pulestin, C. & Stevens, R. 2012. Engineering use cases for modular development of ontologies in OWL. Applied Ontology, 7(2): 113–132.
64. Rescher, N. Realistic Pragmatism: An Introduction to Pragmatic Philosophy.
65. Rogan, M. & Skinner, C. 2017. REDI 3x3 Working paper 28 February 2017 The nature of the South African informal sector as reflected in the quarterly labour-force survey , 2008-2014. , (August): 2008–2014.
66. Romero, A.A., Kaminski, M., Grau, B.C. & Horrocks, I. 2016. Module Extraction in Expressive Ontology Languages via Datalog Reasoning.

67. Saad, A. & Shaharin, S. 2016. The Methodology for Ontology Development in Lesson Plan Domain. International Journal of Advanced Computer Science and Applications, 7(4).
68. Sanderson, M. & Croft, W.B. 2012. The history of information retrieval research. Proceedings of the IEEE, 100(SPL CONTENT): 1444–1451.
69. Shah, U., Finin, T., Joshi, A., Scott Cost, R. & Mayfield, J. 2002. Information Retrieval On The Semantic Web £.
70. Shimizu, C. Modular Ontology Modeling: A Tutorial. http://www.ontologydesignpatterns.org.
71. SIngh Narula, G., Singh, G. & Jain, V. 2014. Information Retrieval (IR) through Semantic Web (SW): An Overview Machine learning View project Ontology Based Information Retrieval View project Information Retrieval (IR) through Semantic Web (SW): An Overview. https://www.researchgate.net/publication/261182515.
72. Suárez-Figueroa, M.C., Gómez-Pérez, A. & Villazón-Terrazas, B. 2009. How to write and use the ontology requirements specification document. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 966–982.
73. Swartout, W. & Russ, T. 1997. Toward Distributed Use of Large-Scale Ontologies t Toward Distributed Use of Large-Scale Ontologies t Bill Swartout Ramesh Patil Kevin Knight. www.aaai.org.
74. Teimourikia, M. & Fugini, M. 2017. Ontology development for run-time safety management methodology in Smart Work Environments using ambient knowledge. Future Generation Computer Systems, 68: 428–441.
75. Uschold, M. & King, M. 1995. Towards a Methodology for Building O n tologies.
76. Usher, W., Gudes, O. & Parekh, S. 2016. Exploring the use of technology pathways to access health information by Australian university students: a multi-dimensional approach. Health Information Management Journal, 45(1): 5–15. http://him.sagepub.com/lookup/doi/10.1177/1833358316639450.
77. Williams, C.C. & Horodnic, I.A. 2016. An institutional theory of the informal economy: Some lessons from the United Kingdom. International Journal of Social Economics, 43(7): 722–738.
78. Wirtz, J., So, K.K.F., Mody, M.A., Liu, S.Q. & Chun, H.E.H. 2019. Platforms in the peer-to-peer sharing economy. Journal of Service Management, 30(4): 452–483.
79. Yu, B. 2019. Research on information retrieval model based on ontology.

80. Yun, H.Y., Xu, J.L., Wei, M.J. & Xiong, J. 2009. Development of domain ontology for e-learning course. In ITME2009 - Proceedings 2009 IEEE International Symposium on IT in Medicine and Education. 501–506.

APPENDICES

APPENDIX A: ETHICAL REQUIREMENTS

The purpose of appendix A was to fulfil the ethical standards outlined in the Cape Peninsula University of Technology research code of ethics. The appendix includes (1) the faculty of Informatics and Design Ethical approval; (2) an introductory letter for the collection of research data from the university and the supervisor.

Cape Peninsula
University of Technology
creating futures

PO Box 1906, Bellville,7535 Symphony Way, Bellville, Cape Town, South Africa
+ 27 (0)21 959 6767 www.facebook.com/cput.ac.za info@cput.ac.za www.cput.ac.za

Office of the Research Ethics Committee
Faculty of Informatics and Design
Room 2.09
80 Roeland Street
Cape Town
Tel: 021-469 1012
Email: ndedem@cput.ac.za
Secretary: Mziyanda Ndede

15 April 2021

Mr Glodi Atoba
c/o Department of Information Technology
CPUT

Reference no: 212028936/2021/8

Project title: Development of a Domain Ontology to Support Information Retrieval on the South African Informal Sector Services

Approval period: 15 April 2021 – 31 December 2022

This is to certify that the Faculty of Informatics and Design Research Ethics Committee of the Cape Peninsula University of Technology approved the methodology and ethics of Mr Glodi Atoba (212028936) for the MTech Information Technology.

Any amendments, extension or other modifications to the protocol must be submitted to the Research Ethics Committee for approval.

The Committee must be informed of any serious adverse event and/or termination of the study.

A/Prof I van Zyl
Chair: Research Ethics Committee
Faculty of Informatics and Design
Cape Peninsula University of Technology

Cape Peninsula
University of Technology

Introductory letter for the collection of research data

Glodi Atoba is registered for the M Tech (IT) degree at CPUT (212028936). The thesis is titled: The Development of a Domain Ontology for The South African Informal Business Sector. The aim of this research is to create a domain ontology that can promote intelligent knowledge services in the Informal business sector in South Africa. The supervisor for this research is:

- Prof Justine Olawande Daramola

In order to meet the requirements of the university's Higher Degrees Committee (HDC) the student must get consent to collect data from organisations which they have identified as potential sources of data. In this case the student will review some of the public documents such as official and policy documents and will use structured interviews to gather data.

If you agree to this, you are requested to complete the attached form and print it on your organisation's letterhead.

For further clarification on this matter please contact either the supervisor(s) identified above, or the Faculty Research Ethics Committee secretary (Ms V Naidoo) at 021 469 1012 or naidoove@cput.ac.za.

Yours sincerely,

Prof Justine Olawande Daramola

4 March 2021

APPENDIX B: PROFESSIONAL EDITOR'S CERTIFICATE

The purpose of appendix b was to show a professional editor's certificate document from a professional and accredited editor of the Cape Peninsula University of Technology that edited this thesis.

NewsPlus Communications (Pty) Ltd
(2016/523137/07)
Unit 15 Silver Willows, 106 Furrow Road, Equestria Pretoria 0184
Cell Phone: +27 62 240 1084
Email: editor@newsplus.ng
odianosent@gmail.com
(Editing, Proofreading, News Writing and Reporting, Speech Writing, Rapporteuring, Web Content Management)

MEMBER,

Professional EDITORS Guild

TO WHOM IT MAY CONCERN

This document certifies that I, Anthony A. Ekata, a member of the Professional Editors' Guild South Africa, Nigerian Guild of Editors, and an accredited editor of the Cape Peninsula University of Technology, Cape Town, South Africa, edited the Master of Technology: Information Technology thesis with the title below for proper Language and Style.

Title: Development of a Domain Ontology to Support Information Retrieval on the South African Informal Sector Services

By

Glodi Mokombati Atoba

Faculty of Informatics and Design

Cape Peninsula University of Technology

Date 7th December 2022

SIGNED

Anthony A. Ekata

+27 62 240 1084