Formal Approach to Uncertainty Modelling in DIS Ontologies

Formal Approach to Information Uncertainty Modelling and Domain Adequacy in DIS Ontologies

By
DEEMAH ALOMAIR,
M.Sc. (Computer Science)

A Thesis Submitted to the Department of Computing and Software
and the School of Graduate Studies of McMaster University
in the Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy
in
Computer Science

McMaster University
Hamilton, Ontario

Doctor of Philosophy (2025)
Department of Computing and Software
McMaster University
Hamilton, Ontario, Canada

TITLE: Formal Approach to Information Uncertainty Modelling and Domain Adequacy in DIS Ontologies

AUTHOR:
Deemah Alomair,
M.Sc. (Computer Science)

SUPERVISOR:
Dr. Ridha Khedri
Professor, Department of Computing and Software
McMaster University, ON, Canada

SUPERVISORY COMMITTEE CHAIR:
Dr. Ryszard Janicki
Professor, Department of Computing and Software,
McMaster University, ON, Canada

SUPERVISORY COMMITTEE MEMBERS:
Dr. Hassan Ashtiani
Associate Professor, Department of Computing and Software,
McMaster University, ON, Canada

Dr. Ridha Khedri
Professor, Department of Computing and Software,
McMaster University

NUMBER OF PAGES: xi, 182

Lay Abstract

Modern intelligent systems rely on ontologies to represent domain knowledge and support automated reasoning. However, these systems often face uncertainty arising from information imperfection (e.g., incompleteness or inconsistency) or uncertainty of relevance. This thesis addresses two key challenges: how to model and reason about uncertainty caused by imperfect information in ontologies, and how to manage uncertainty of relevance by ensuring an ontology's adequacy for its intended domain. To address these problems, it introduces a classification of uncertainty types. Extends the Domain Information System (DIS) framework with quantitative modelling and reasoning capabilities, leveraging possibility theory to manage incomplete knowledge. It also proposes a structured formulation to assess ontological domain adequacy through ontological and data commitment principles. Additionally, it integrates statistical validation techniques to determine the relevance of data-defined concepts. By bridging formal ontology engineering with uncertainty modelling, the thesis lays the foundation for more trustworthy ontology-based systems in data-driven environments.

Abstract

Ontologies play a central role in structuring domain knowledge and enabling automated reasoning within a given domain. However, real-world applications increasingly demand ontologies that can tolerate and represent uncertainty, stemming from either imperfect information or a mismatch between the ontology and its intended domain. This thesis addresses two fundamental types of ontological uncertainty: (1) uncertainty due to information imperfections, such as incompleteness and ambiguity; and (2) uncertainty of relevance, which arises when ontologies fail to capture the semantics of their domain adequately.

To address these challenges, this thesis makes four key contributions. First, it presents a comprehensive survey and classification of uncertainty modelling approaches in domain ontologies, synthesizing a decade of research (2010-2024). It then proposes a formal taxonomy that links types of uncertainty with their appropriate mathematical formalisms for management, and their points of occurrence within the ontologies. Second, it proposes a possibilistic extension to the Domain Information System (DIS) framework that incorporates necessity-weighted formulas to model incomplete information and support flexible, logic-based reasoning. Third, it introduces a novel theory of domain adequacy, based on formal notions of ontological and data commitments, to guide the construction of minimal yet semantically sufficient sub-ontologies. Fourth, it extends this theory to statically defined datascape concepts, developing a practical framework and tooling that enables automated validation of data adequacy through statistical evaluation of real-world datasets.

Altogether, this work advances the theoretical foundation and practical implementation of uncertainty-aware ontology engineering. It demonstrates how to unify data-centric reasoning with formal ontology design, yielding systems that are not only semantically rigorous but also grounded in empirical evidence. The results offer a principled approach to managing uncertainty in ontology-based systems, making them more adaptable, interpretable, and aligned with dynamic, data-driven domains.

Acknowledgements

First and foremost, I thank Allah (God) for granting me the strength, patience, and clarity of mind to complete this journey. Without his guidance and mercy, none of this would have been possible.

I extend my deepest gratitude to my beloved parents, whose unwavering love, endless prayers, and constant encouragement have been the foundation of all my achievements. Their sacrifices, wisdom, and belief in me have shaped every step of my academic and personal life. I also want to thank my dear sisters and brother for their continuous support, understanding, and kindness. Their encouragement and companionship have been a source of comfort and motivation throughout this journey.

To my small family, my dear husband, my shining star Leena, and my sweet son Nawaf, thank you for being my greatest blessings and unwavering sources of love. Your constant support, encouragement, and affection have been the light that guided me through every challenge and the joy that filled my heart. I am endlessly grateful for your patience, understanding, and boundless love, which have given me strength, hope, and inspiration to pursue my dreams.

I am especially grateful to my supervisor, Dr. Ridha Khedri, for his exceptional guidance, insightful feedback, and continuous support throughout this thesis. His mentorship and trust in my abilities have been invaluable to my growth as a researcher.

Contents

List of Figures

Acronyms

DIS Domain Information System. x, 3, 8, 9, 10, 11, 12, 14, 16, 17, 18, 19, 20, 21, 89, 102, 163, 164, 165

DISEL Domain Information System Extended Language. xi, 15, 19, 21, 162, 164, 165

DL Description Logic. 1, 2, 3, 4

DST Dempster-Shafer Theory. 4, 11, 15, 20

KIF Knowledge Interchange Format. 3

ML Machine Learning. 1, 13

NLP Natural Language Processing. 13

OBDA Ontology-Based Data Access. 3

OWL Web Ontology Language. 2, 3, 164

OWL2 Web Ontology Language-2. 2

RDF Resource Description Framework. 3, 164

RQs Research Questions. 9, 10

Declaration of Academic Achievement

I, Deemah Alomair, declare that this thesis titled "Formal Approach to Information Uncertainty Modelling and Domain Adequacy in DIS Ontologies" and the works presented in it are my own. I confirm that I contributed as follows:

- Chapter 1: Introduction.
  This chapter sets the stage for the thesis by presenting the general and specific contexts of the research. It defines the problem statement, outlines the motivation, research questions, objectives, thesis contribution to publications and the structure of the thesis. I authored this chapter in full.

- Chapter 2: A Comprehensive Review of Uncertainty Modelling in Domain Ontologies.
  This chapter provides a systematic and comprehensive survey of uncertainty modelling in domain ontologies. It introduces a structured taxonomy, reviews uncertainty-handling formalisms, and reviews related approaches covering the period of 2010-2024. I co-authored this chapter with Dr.Ridha Khedri and Dr.Wendy MacCaull. I led the systematic review process, conducted the literature analysis and classification, organized regular review meetings, synthesized findings, authored all sections of the paper, and co-authored the manuscript.

- Chapter 3: Possibilistic extension of Domain Information System (DIS) Framework.
  This chapter introduces a novel uncertainty-supported modelling and reasoning framework for the DIS-based ontologies using possibility theory. It presents an extended DIS structure that integrates weighted formulas to represent and reason under incomplete information. I co-authored this chapter with my supervisor Dr.Ridha Khedri. I co-designed the extended DIS model, formalized the weighted components, and collaborated in developing, and implementing the associated reasoning methodology. I co-authored the manuscript.

- Chapter 4: Towards a Cartesian Theory of Ontology Domain Adequacy.
  This chapter introduces a theoretical framework that characterizes domain adequacy in ontologies through the interplay of ontological commitment and data commitment. It establishes formal definitions and properties that govern minimal adequate sub-ontologies. I co-authored this chapter with my supervisor Dr.Ridha Khedri. I contributed to the theoretical formulation, formalized key concepts and claims, and co-authored the manuscript.

- Chapter 5: Data-adequacy of Ontologies with Statistically-defined Concepts: A DISEL plug-in.
  This chapter extends the Domain Information System Extended Language (DISEL) to support data commitment verification for statistically-defined datascape concepts. It presents a system that automates the evaluation of datascape concept adequacy using statistical analysis. I co-authored this chapter with Dr.Yihai Chen, Yijie Wang, and Dr.Ridha Khedri. I developed the *R*-program generation module and contributed significantly to the design, case study, authored all sections of the paper, and co-authored the manuscript.

- Chapter 6: Conclusion and Future Work.
  This chapter concludes the thesis by summarizing the main contributions, reflecting on their theoretical and practical values, and outlining directions for future research. I authored this chapter in full.

- Appendix A: DISEL: A Language for Specifying DIS-based Ontologies.
  This appendix present DISEL syntax and semantics. I co-authored this chapter with Dr.Yihai Chen, Yijie Wang, and Dr.Ridha Khedri. I co-developed the language specification, and co-authored the manuscript.

Chapter 1

Introduction

1.1 General Context

In today's data-driven world, data analytics has become a cornerstone for extracting meaningful insights from vast and complex systems. It encompasses a wide range of methodologies, from classical statistical approaches to more advanced Machine Learning (ML) techniques, each designed to uncover patterns, predict outcomes, and support decision-making across various domains.

Statistical methods form the foundation of data analytics, relying on probability theory, hypothesis testing, regression analysis, and Bayesian inferences [Lista, 2023]. These methods offer well-grounded principles for analyzing data, quantifying uncertainty, and making data-driven decisions. For example, in epidemiology, logistic regression models help estimate disease probabilities based on risk factors, while time-series forecasting techniques enable financial analysts to predict market trends.

Beyond classical statistical techniques, ML has significantly enhanced analytical capabilities by enabling models to learn patterns from data and make predictions [Kenge, 2020]. ML techniques are broadly categorized into supervised learning, unsupervised learning, and semi-supervised learning, among other types [Kenge, 2020]. Despite their power, ML models often operate as *black boxes*, providing little interpretability or explicit reasoning over structured knowledge [Hassija et al., 2024]. This limitation necessitates logic-based approaches that complement data-driven techniques with formalized reasoning.

A key framework in logic-based knowledge representation and generation is Description Logic (DL), which underpins a family of formal ontology languages designed for structured knowledge representation with well-defined semantics such as Web Ontology Language (OWL) [Antoniou and Harmelen, 2009], and Web Ontology Language-2 (OWL2) [Grau, 2009]. It provides the theoretical foundation for defining ontology-based concepts, roles, and individuals in a precise and machine-interpretable manner [Baader et al., 2005]. These languages support several reasoning tasks including, subsumption, consistency verification, and insistence checking. DL-based reasoning allows for the derivation of implicit knowledge, ensuring logical coherence and semantic consistency in structured knowledge bases.

Building on these foundations, researchers and developers sought a tool capable of representing information in a manner that facilitates knowledge extraction, supports logical analysis, and enables explicit and transparent reasoning. Such a tool must keep pace with the current scalability demands of modern knowledge-based systems while satisfying reasoning requirements. In this context, ontology has emerged as one of the leading frameworks for domain knowledge representation and reasoning, particularly in the era of data analytics and advanced computing.

Originally, ontology was considered a branch of philosophical science before it emerged in the field of data analytics. From the philosophical point of view, ontology can be viewed as the study of existing, existence, or nature of being [Berryman, 2019]. On the other hand, in the discipline of ontology engineering, ontology can be defined as a formal, explicit specification of a shared conceptualization [Gruber, 2018]. Informally, ontology structure consists of a set of concepts, their definitions, relationships, instances, properties, and constraints expressed as axioms. Ontology is a popular tool in the field of the semantic web. It is one of the main engines to represent and reason on semantic web knowledge due to its indisputable features. With the use of ontology, information accessibility, re-usability, and interoperability are becoming feasible with lower overhead. Moreover, data heterogeneity issues have lessened. Ontologies have several applications in the semantic web, like in healthcare [El-Sappagh et al., 2018], information retrieval [Jain et al., 2021], decision making [He et al., 2022], and e-commerce [Karthik and Ganapathy, 2021].

Ontology languages are known to be formal languages used to construct ontologies, in such a way that knowledge is encoded, represented, and reasoned upon. There exists a different classification of ontology languages. However, according to [Maniraj and Sivakumar, 2010], ontology languages can be classified into three main categories: logic-based languages (e.g., [Baader et al., 2005]), frame-based languages (e.g., [Xue et al., 2010]), and graph-based languages (e.g., [Corbett, 2008]). The former includes any language that is based on logic, like first-order predicate logic, rule-based logic, and DL. Examples of such languages are OWL, and Knowledge Interchange Format (KIF) language [Genesereth, 1991]. Frame-based languages like F-logic [Kifer et al., 2000]. While the latest is the languages that are represented graphically by one of the graphical means, like a semantic network as in Resource Description Framework (RDF) language [Pan, 2009]. A comprehensive review of ontology languages can be found in [Kalibatiene and Vasilecas, 2011].

From a different perspective, ontology can be used as a component of a system, like in the Ontology-Based Data Access (OBDA) system [Xiao et al., 2018], DOGMA [Jarrar and Meersman, 2008], and DIS [Marinache et al., 2021]. These are examples of frameworks that separate domain knowledge represented by an ontology from data contents that are stored as records in a database format. These frameworks proved to be more suitable for dealing with information integration and expansion processes.

In particular, the DIS framework adopts a bottom-up, data-centric approach to ontology construction. It builds ontologies directly from datasets through a Cartesian construction of concepts, where complex concepts are formed by combining atomic elements. It structurally separates domain knowledge (the ontology) from data views (the dataset), linking the two via a formal mapping operator. This architectural separation grounds the ontology in data, helping reduce mismatches between conceptual and observed data values. Unlike DL-based ontologies, which separate the A-Box and T-Box logically, DIS enforces this separation structurally, leading to improved modularity, transparency, and maintainability in ontology design. DIS is especially useful in domains where ontologies must be generated or adopted from real-world datasets, offering a principled way to align conceptual models with empirical data. Moreover, DIS supports mereological reasoning by employing cylindric algebra to represent datasets and Boolean algebra to construct conceptual mereological structure. This enables robust modelling of parthood relations within structured data [Andrew LeClair, 2025]. In contrast, traditional ontology languages such as OWL are effective for defining conceptual structures [Guarino et al., 2009] but struggle to capture the internal mereological relationships of Cartesian datasets without relying on complex extensions like OBDA systems [Poggi et al., 2008]. Incorporating Cartesian types into OWL for mereological reasoning in non-trivial and can introduce inference ambiguities [Krieger and Willms, 2015]. The DIS framework will be used extensively throughout this thesis. A detailed overview of DIS can be found in [Marinache, 2025]. Moreover, a theoretical background of this framework is discussed extensively in chapter 3, chapter 4, and chapter 5.

1.2 Specific Context

Although classical ontologies have achieved substantial success, their ability to manage uncertainty remains a significant challenge. Classical ontologies are widely used for representing crisp and well-defined knowledge. However, their inherent rigidity makes them ill-suited for handling the uncertain aspect of real-world [Fareh, 2019]. As knowledge representation becomes increasingly complex, uncertainty emerges as a critical challenge in ontology-based systems, affecting both the adequacy of the domain ontologies and the information they encode.

Broadly, two major types of uncertainty arise in ontology-based systems. The first is uncertainty due to information imperfection, which is an inherent aspect of knowledge and manifests in various forms. The literature presents several classifications of information imperfection [Karanikola, 2018, Ma et al., 2013]. However, the most common types of imperfections are incompleteness, imprecision, vagueness, ambiguity, and inconsistencies [Bosc and Prade, 1997, Ma et al., 2013]. Alternatively, some literature [Anand and Kumar, 2022, Ceravolo et al., 2008, Stoilos et al., 2005] considers all the aforementioned types of imperfection as types of uncertainty. Thus, uncertainty may arise from information incompleteness, imprecision, vagueness, ambiguity, or inconsistencies. This type of uncertainty is extensively addressed and discussed in chapter 2, and chapter 3, of this thesis. The second type relates to uncertainty of relevance, which concerns the adequacy of a domain ontology in capturing the necessary concepts, relationships, and instances for a given domain. Addressing this type of uncertainty requires ensuring that the ontology accurately reflects the domain's intended scope and purpose [Waterson and Preece, 1999]. This type of uncertainty is extensively addressed and discussed in chapter 4, and chapter 5, of this thesis. These two major types of uncertainty occurring within ontological systems are illustrated in Figure 1.1.

Currently, different formalisms have been developed to extend ontology languages and their foundational components, such as DL, to effectively manage uncertainty arising from information imperfections. Each formalism is designed to address a specific manifestation of uncertainty, ensuring that different types of imperfections are systematically handled. The most prominent formalisms found in the literature for managing uncertainty in domain ontologies are possibility theory (e.g., [Dubois and Prade, 2015]), Dempster-Shafer Theory (DST) (e.g., [Sentz and Ferson, 2002]), rough set theory

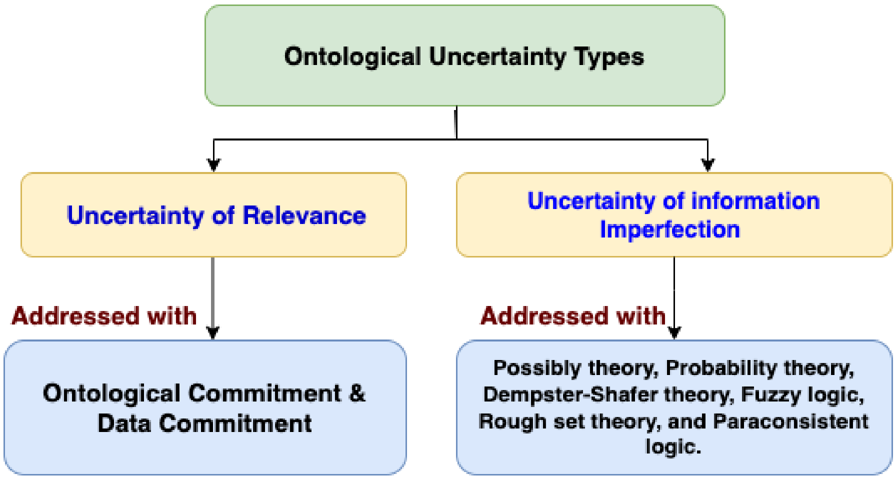


FIGURE 1.1: Major Types of Ontological Uncertainty.

(e.g., [Akama et al., 2020]), paraconsistent logic (e.g., [Priest, 2002]), probability theory (e.g., [Laha and Rohatgi, 2020]) and fuzzy logic (e.g., [Zimmermann, 2011]). Each of these offers distinct methodologies for reasoning under uncertainty, contributing to more flexible and expressive ontology-based systems. A detailed discussion of these formalisms and the specific types of uncertainty they address is provided in chapter 2. Moreover, a developed taxonomy that presents the major types of ontological uncertainty due to imperfect information along with their corresponding formalisms to manage is presented in Figure 1.2.

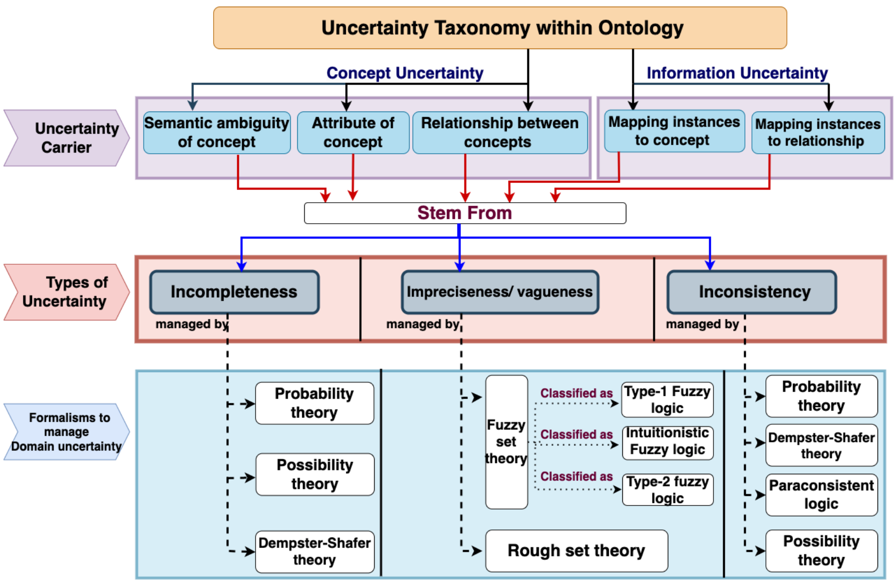


FIGURE 1.2: Ontological Uncertainty Taxonomy

From another perspective, ontological commitment plays a crucial role in addressing the uncertainty of relevance by ensuring that a domain ontology accurately captures the essential concepts and relationships required for a given domain. It establishes agreed-upon assumptions about the nature of these concepts and their interrelations, explicitly defining the ontology's scope and intended use. Ontological commitment helps in identifying a sufficient ontology that aligns with the real-world requirements while also reducing the size and complexity of the ontology. This structured approach facilitates the extraction of a minimal yet comprehensive ontology that remains robust enough to support reasoning within its intended domain, thereby enhancing decision-making processes. Furthermore, aligning ontological commitment with data commitment extends relevance beyond the conceptual level, ensuring that all concepts, relationships, and instances are semantically supported by the underlying dataset.

1.3 Motivation

As knowledge representation and extraction processes continue to evolve, the need for semantically rich and uncertainty-aware systems becomes increasingly important. Ontologies have emerged as a powerful framework for structuring and reasoning about knowledge due to their ability to explicitly define concepts and relationships semantically within a domain. However, classical ontologies are inherently crisp, making them unsuitable for handling uncertainty, a fundamental aspect of real-world knowledge. This limitation has driven extensive research into extending ontological models to incorporate uncertainty-handling mechanisms.

Existing approaches primarily focus on addressing uncertainty due to information imperfections, such as vagueness, incompleteness, and inconsistency, using well-established formalisms like fuzzy logic, probability theory, and possibility theory. These approaches enhance ontologies by making them more flexible and expressive. However, understanding, classifying, and effectively applying uncertainty-handling approaches is a complex and challenging task due to the diverse and growing number of formalisms and approaches proposed in the literature. The field of uncertainty modelling in ontologies is vast and fragmented, making it difficult for researchers and practitioners to identify the most suitable formalism or approach for a given application. There is a pressing need for a comprehensive classification of existing approaches that systematically categorize methods based on the types of uncertainty they address and the reasoning capabilities they provide. Additionally, there is a critical need to develop clear guidelines and a structured framework that explicitly describes where uncertainties might arise in ontology-based systems and how to effectively handle them. Such a framework would provide a systematic roadmap for integrating uncertainty-aware mechanisms into ontologies, ensuring that the uncertainty of information imperfections is addressed coherently and practically.

From a different perspective, most of the existing approaches do not fully address the uncertainty of relevance, the challenge of ensuring that an ontology adequately captures the necessary domain's concepts and relationships. This type of uncertainty affects the validity, completeness, and applicability of an ontology for real-world decision-making and reasoning. To bridge this gap, ontological commitment plays a crucial role in ensuring domain adequacy by extracting the appropriate ontology governing domain scope. Furthermore, aligning ontological commitment with data commitment ensures that the ontology remains not only conceptually relevant but also grounded in real-world datasets. By achieving this, we can move towards more adaptive, systematically meaningful, and practically useful ontology-based reasoning systems.

1.4 Problem Statement

With the scalability of knowledge-based systems and the emergence of big data and its continuous developments, the focus on uncertainty within these systems is increasingly gaining attention. While traditional ontologies are effective in representing well-structured knowledge, their crisp and rigid nature makes them unsuitable for handling uncertainty. The latter is a fundamental aspect of real-world knowledge. To address this limitation, various uncertainty-handling formalisms have been introduced to extend ontology languages and improve their flexibility. However, despite significant progress, several key challenges remain unresolved.

First, the field of uncertainty modelling in ontologies is vast, complex, and fragmented, making it challenging to identify, classify, and apply appropriate uncertainty-handling approaches. The absence of a clear taxonomy and structured framework for uncertainty modelling hinders the selection of the most suitable formalism for specific applications. Researchers and practitioners currently lack systematic guidelines that help in determining which formalism to adopt based on the nature of uncertainty present in a given ontology. Thus, there is a critical need for a comprehensive study that:

1. Clearly defines the different types of uncertainty in ontology-based systems and their sources.

2. Classifies and organizes existing uncertainty-handling approaches, mapping them to the specific types of uncertainty addressed and the formalism adopted.

3. Establishes structured guidelines to assist in selecting the most appropriate formalism for a given uncertainty.

Second, most ontology-based reasoning frameworks assume complete or deterministic information. In practice, however, incomplete information is common, particularly in dynamic and data-driven environments. Existing frameworks do not offer systematic support for modelling and reasoning under incomplete information. While possibilistic logic offers a principled way to handle this form of uncertainty, there remains a gap in its integration into structurally grounded ontology frameworks such as DIS. A unified framework is needed to embed uncertainty tolerance modelling within ontologies, while maintaining ontological coherence.

Third, traditional uncertainty research tends to focus on uncertainty due to information imperfections, such as vagueness, incompleteness, and inconsistency, while largely overlooking the equally critical issue of uncertainty of relevance. Ensuring that an ontology adequately captures the required domain concepts and relationships is essential for its validity and applicability. The lack of formal methods to address the uncertainty of relevance makes it difficult to determine whether an ontology sufficiently represents its intended domain. Ontological commitment provides a structured approach to mitigate this issue, but its integration with ontological-based systems remains under-explored.

Fourth, while ontological commitment ensures that an ontology aligns with domain semantics, a parallel challenge lies in ensuring data-driven adequacy. Data commitment, which ensures that the ontology's concepts are sufficiently supported by available data, remains unexplored. Without mechanisms to assess and enforce data commitment, an ontology may contain concepts that are either irrelevant to the dataset or inadequately instantiated, leading to poor domain coverage or misleading reasoning results. Achieving optimal domain adequacy thus requires a formal understanding of how ontological and data commitments interact, and a method for validating both in a coherent and integrated manner.

Last, with the increasing role of data-driven decision making, ontologies must incorporate statistically-defined concepts derived from observed data. However, current ontology frameworks lack support for verifying whether such data-driven concepts are logically consistent and semantically committed to the ontology. There is a need for a structured mechanism for ensuring the data commitment of statistically defined concepts within the ontological frameworks, which bridges statistical modelling with ontological reasoning, thus enabling seamless integration of statistical evidence into formally data-driven ontologies.

By addressing these challenges, ontology-based systems can evolve into more robust, scalable, and systematically aware frameworks capable of reasoning under uncertainty and supporting reliable decision-making in complex, real-world applications.

1.5 Research Questions

The key Research Questions (RQs) guiding this thesis focus on developing a comprehensive understanding of uncertainty in domain knowledge and its formal modelling within ontology frameworks. Understanding the nature of uncertainty is the first step toward developing effective methodologies for managing it, requiring a clear distinction between different types of uncertainty, including information imperfections and uncertainty of relevance.

Regarding uncertainty arising from information imperfection, a critical aspect is to understand the types of uncertainty introduced into domain ontologies. Equally important is the identification and classification of existing mathematical formalisms that support uncertainty handling, enabling the structured and precise modelling of uncertain aspects within ontology-based systems. This study further aims to establish guidelines and selection criteria for choosing the most suitable formalism based on the type of uncertainty. Additionally, this work explores how uncertainty due to information imperfections can be formally represented within ontology-based frameworks, such as the DIS framework, to support more expressive and adaptable knowledge models. Specifically, this study explores the following set of RQs in this context:

1. RQ-1: How can uncertainty resulting from imperfect information be formally modeled and managed within domain ontologies?
   - RQ-1.1: What are the different types and sources of uncertainties that arise in the context of domain ontologies?
   - RQ-1.2: Where are uncertainty reside within a ontological framework (e.g., concepts, instances)?
   - RQ-1.3: What mathematical formalisms are available for representing and reasoning about such uncertainties?

- RQ-1.4: What selection criteria or guidelines can support choosing the most appropriate formalism based on the nature of the uncertainty involved?

2. RQ-2: How can data-driven ontology frameworks formally represent and manage uncertainty due to imperfect information?

   - RQ-2.1: How can uncertainty resulting from information imperfections be formally represented within the DIS framework?
   - RQ-2.2: What specific type of uncertainty arise in DIS, and which formalism is most suitable for managing it?
   - RQ-2.3: How can reasoning be conducted effectively in an uncertainty-aware DIS framework?

   Another critical aspect is ensuring the adequacy of domain ontologies, particularly in dynamic and evolving domains where concepts and relationships may change over time. This helps in reducing the growing complexity of large and monolithic ontologies. This research investigates how ontological commitment can be structured to address the uncertainty of relevance, ensuring that an ontology remains adequate and representative of its domain. This study also examines the role of data commitment in maintaining ontology relevance. Specifically, it explores how aligning ontological commitment with data commitment ensures that all concepts, relationships, and instances within the ontology are grounded in real-world datasets. Another critical aspect is ensuring the adequacy of domain ontologies for statistically defined concepts, ensuring that statistical learning models and ontological structure remain aligned, thus improving the applicability of such ontologies. Specifically, this study explores the following set of RQs in this context:

3. RQ-3: What constitutes domain adequacy in ontologies, and how can it be ensured?

   - RQ-3.1: What approaches ensure the adequacy of domain ontologies? and what role does ontological and data commitments play in this matter?
   - RQ-3.2: How can domain adequacy be ensured and validated within the DIS framework?
   - RQ-3.3: How can the domain adequacy of an ontology be assessed for statistically defined concepts within the DIS framework?

By addressing these research questions, this thesis aims to develop a comprehensive structure for uncertainty management in ontologies, providing clear guidelines for selecting appropriate formalisms, ensuring ontology relevance and adequacy, and enhancing reasoning efficiency for real-world applications.

1.6 Objectives

The key research objectives driving this study are designed to systematically develop a structured approach for understanding, modelling, and managing uncertainty in ontology-based systems (e.g., DIS). The following objectives define the scope of this research:

Objective 1: Identify and Classify Uncertainties in Ontologies.

This objective focuses on analyzing the different types of uncertainties that arise in ontology-based systems, distinguishing between uncertainty of relevance and uncertainty due to information imperfections. A structured classification will provide a foundation for selecting appropriate modelling techniques. This objective directly addresses Research Question 1.

Objective 2: Investigate Formalism for Uncertainty Handling.

This objective aims to explore and analyze the various mathematical formalisms used to model and manage uncertainty of imperfections in ontologies, including fuzzy logic, probability theory, possibility theory, DST, rough set theory, and paraconsistent logic. This objective directly addresses Research Question 2.

Objective 3: Establish Guidelines for Selecting the Best Uncertainty Handling Formalism.

Given the diversity of approaches for managing uncertainty, this objective aims to develop a clear taxonomy or selection criteria to guide researchers and practitioners in choosing the most appropriate formalism based on factors such as the type of uncertainty and where uncertainty might arise within the ontology frameworks. This objective directly addresses Research Question 3.

Objective 4: Develop a Structured Approach for Representing Uncertainty of Imperfections in DIS Framework.

Building on the classification and formalism selection, this objective focuses on designing a structured approach for modelling uncertainty due to information imperfection in the DIS framework. This approach defines where uncertainty arises, how it can be formally modelled within DIS structure, and how reasoning mechanisms can be adopted to handle uncertainty effectively. This objective directly addresses Research Question 4.

Objective 5: Ensure Ontology Domain Adequacy Through Ontological and Data Commitments.

Since the uncertainty of relevance arises when ontology does not sufficiently capture its intended domain, this objective focuses on ontological and data commitments as methods for ensuring domain adequacy. This involves establishing a structured theory for defining an ontology's scope, aligning it with real-world domain requirements, and preventing conceptual and data-grounded gaps. This objective directly addresses Research Question 5.

Objective 6: Determine Ontology Domain Adequacy for Statistically Defined Concepts.

Statistically defined concepts, those derived from statistical inferences, pose challenges for ontology modelling. This objective aims to explore methodologies for ensuring ontology domain adequacy when integrating such concepts. Specifically, it investigates how to ensure the data commitment of those concepts within the DIS framework. This objective directly addresses Research Question 6.

By achieving these objectives, this research provides a comprehensive framework for uncertainty management on ontologies, offering structured guidelines for formalism selection, ensuring domain adequacy (including statistically defined concepts), integrating data commitment, and enhancing reasoning efficiency.

1.7 Thesis Contributions

This thesis makes several key contributions to the field of ontology-based reasoning under uncertainty, addressing both classical imperfection (e.g., incomplete information) and relevance uncertainty (i.e., adequacy of an ontology to a domain). The contributions are categorized as follows:

1. **A comprehensive survey on uncertainty modelling in ontologies (outcome of Objectives 1-3)**

   To establish a solid foundation for this research, a systematic survey was conducted on uncertainty modelling in ontology-based systems. This review provides a structured analysis of existing approaches, offering insights into the different formalisms employed and their applicability in handling uncertainty. Specifically, it:

   - Introduces a guidance taxonomy that systematically maps different types of uncertainty to appropriate modelling formalism. This taxonomy, presented in Figure 1.2, also identifies where uncertainty may arise within an ontological framework, serving as a roadmap for researchers and practitioners.
   - Provides a theoretical foundation by outlining formalisms commonly used to model and reason about uncertainty in ontologies, ensuring clarity on their theoretical underpinnings.
   - Classifies reviewed papers based on the formalism adopted, facilitating an organized understanding of how uncertainty has been addressed across different formalisms.
   - Analyzes the orientation of each study, distinguishing between foundational research and application-driven approaches, those based on prior modelling approaches.
   - Identifies key motivations and domain area behind uncertainty modelling, specifying the problem under investigation and challenges addressed in each study.
   - Details the modelling approaches applied to integrate uncertainty into ontological frameworks, shedding light on the methodologies used.
   - Examines the reasoning procedures proposed in the reviewed studies.
   - Investigates the integration of ML and Natural Language Processing (NLP) techniques in the reviewed approaches, assessing their role in uncertainty reasoning approaches.
   - Lists supporting tools or languages used in uncertainty-aware ontology modelling, offering practical insights into available resources.

- Illustrates any given evaluation of the reviewed approaches, discussing how they were validated and the effectiveness of their proposed approaches.

2. **A framework for handling uncertainty due to incomplete information in ontologies (outcome of Objective 4)**

   This contribution focuses on addressing uncertainty arising from incomplete information in ontology-based reasoning systems. It proposes:

   - An enhanced DIS framework incorporating quantitative possibility theory to model and reason under incomplete information.
   - A weighted formula extension that spans across all components of the DIS framework to systematically handle uncertainty, based on a solid mathematical foundation.
   - An extended reasoning procedure based on the enhanced framework.

3. **A Cartesian theory of domain adequacy for ontologies (outcome of Objective 5)**

   To address relevance uncertainty (i.e., ensuring an ontology is adequate for its domain), this contribution proposes:

   - A formal distinction between objective reality (representing the objective existence) and datascape reality (representing how the underlying data captures reality).
   - A formal mathematical foundation for assessing optimal domain adequacy, ensuring that the ontology-based system adheres to both ontological and data commitments.
   - The ontological commitment theory and its related properties to identify the minimal sub-ontologies (modules) that maintain domain adequacy.
   - The data commitment theory, its related properties, and its interplay with the ontological commitment.

4. **A framework for data commitment of statistically defined concepts (outcome of Objective 6)**

   The contribution addresses the challenge of integrating statistical data-driven concepts into ontology-based reasoning by:

- Developing a structured approach for data commitment, ensuring that statistically defined concepts align with the semantic and logical foundations of an ontology.
- Bridging the gap between statistical reasoning and ontological inference, making ontology-based systems more flexible and data-driven.
- Extending DISEL with the proposed uncertainty-tolerance framework to enable applying it practically in real-world applications.

1.8 Description of Contributions to Publications

This thesis has been prepared in the "sandwich thesis" format. This section describes the contribution of each of the co-authors to the work.

Chapter 2: A Comprehensive Review of Uncertainty Modelling in Domain Ontologies.

Authors: Deemah Alomair, Ridha Khedri, and Wendy MacCaull.

The idea of conducting a systematic review paper emerged during the first year of the PhD journey. It becomes evident that the domain of uncertainty modelling in ontologies is vast and fragmented, necessitating a structured classification of existing approaches. As researchers in this field, we recognized the importance of comprehensively understanding and categorizing existing methodologies based on key factors such as the type of uncertainty handled and the formalism adopted.

Dr. Ridha Khedri initiated the idea of conducting the review, and as a team (i.e., Deemah Alomair, Ridha Khedri, and Wendy MacCaull), we collaboratively formulated research questions, specified the search period (2010-2024), and defined search keywords. The research questions were iteratively refined until we finalized a set of eleven key research questions. Additionally, the team identified the necessity of snowballing the initial set of retrieved papers to ensure comprehensive coverage of relevant literature.

Two levels of snowballing were conducted to expand the review scope. Deemah Alomair was responsible for the entire snowballing process, leading to the identification of 562 relevant papers for analysis. These papers were subsequently classified based on the formalism adopted to handle uncertainty, resulting in eight categories: fuzzy-based, probability-based, possibility-based, rough set-based, paraconsistent-based, DST-based, combined or hybrid-based, and review-based papers.

To ensure a continuous review process, team members engaged in weekly paper discussions, systematically reading and analyzing papers, classified based on formalisms. Discussions were held to assess the contributions of each paper to the research questions. All analysis and findings were stored in a tabulated format in a cloud-based environment, accessible to all team members. Deemah Alomair was responsible for curating and distributing weekly reading materials to the rest of the team. All the team members read and discussed the papers together.

Moreover, Deemah Alomair proposed an initial guidance taxonomy that works as a roadmap to understand where uncertainty might arise within ontology frameworks and what the best formalisms are to manage them.

The review phase spanned approximately three years (2021-2024). Upon completing the initial analysis, Deemah Alomair was responsible for filtering and synthesizing the results, structuring the findings into the results section of the paper. Following this, Deemah Alomair wrote the entire review paper, drafting all sections and sub-sections before submitting it for feedback to Dr. Ridha Khedri and Dr. Wendy MacCaull. The paper underwent multiple rounds of refinement until the final version was approved by all team members.

This extensive work is presented in chapter 2 of this thesis. The material in this chapter has been accepted for publication in the highly ranked *ACM Computing Surveys* journal, with 2024 Impact Factor: 28.0 (ranked 1/147 in Computer Science Theory & Methods). The manuscript, spanning 61 pages, adheres to the rigorous standards of the journal. The copyright of the material is held by the authors, with publication rights licensed to *ACM*. This review constitutes a major contribution to this field, offering a comprehensive classification and evaluation of uncertainty due to information imperfections modelling approaches in domain ontologies.

Chapter 3: Possibilistic extension of Domain Information System (DIS) Framework.

Authors: Deemah Alomair, and Ridha Khedri.

The idea of developing a comprehensive framework to address uncertainty due to imperfect information emerged at the beginning of the PhD journey. The initial challenges are identifying the specific type of uncertainty we have, determining the most suitable formalism, and effectively integrating that formalism into the DIS framework. To establish a solid foundation, we prioritized conducting a systematic survey to gain a broader understanding of the domain, classify existing approaches, and determine where our framework fits within the landscape. As the survey paper neared completion (2024), we understood our framework focus, clearly identified the type of uncertainty being addressed, and selected the most appropriate formalism to support it.

Through weekly discussions, we systematically formulated our complete framework. We determined that our primary concern was to address uncertainty due to incomplete information and concluded that possibility theory was the most appropriate formalism (a detailed rationale for this selection is provided in the paper). Consequently, we extended the DIS structure by associating a necessity degree with each instance-to-concept mapping, relation, attribute-to-concept mapping, and concept definition. This enhancement embeds uncertainty tolerance directly into all components of the DIS framework, enabling it to represent and reason with incomplete information.

The writing phase spanned approximately one year (2024-2025). Following the initial draft, the paper underwent multiple rounds of refinement until a final version was approved by all team members.

This extensive work is presented in chapter 3 of this thesis. The material was accepted for publication in the *SCITEPRESS* Digital Library as part of the proceedings of the *17th International Conference on Knowledge Engineering and Ontology Development*. The manuscript, comprising 12 pages, meets the rigorous standards expected by the conference and its associated publication venue. The copyright is held by the authors, with publication rights licensed to *SCITEPRESS, Science and Technology Publications*.

This paper constitutes a significant contribution to the field, offering a novel approach to uncertainty modelling in domain ontologies by enhancing the DIS framework with a structured uncertainty-tolerant model. It provides theoretical insights into the integration of weighted formulas and demonstrates how this extension enables more robust and efficient reasoning procedures. This work establishes a foundational framework for reasoning over uncertain domain knowledge within ontology-based systems.

Chapter 4: Towards a Cartesian Theory of Domain Adequacy of Ontologies.

Authors: Deemah Alomair, and Ridha Khedri.

The idea of writing about domain ontology adequacy emerged during the literature review phase. It becomes evident that the uncertainty of relevance is a critical aspect of uncertainty modelling in ontologies that requires exploration. This type of uncertainty is rooted in philosophy and is traditionally addressed through ontological commitment.

To develop a comprehensive understanding, we began by analyzing the domain and identifying its fundamental components. Through weekly discussions, we progressively formulated our theory. During this process, we recognized that the reality of the DIS framework is divided into two primary aspects: the objective and the datascape. Each aspect requires a distinct form of commitment, objective commitment and data commitment, which collectively ensure the domain adequacy of an ontology.

To maintain steady progress, the team engaged in structured weekly discussions, systematically analyzing findings and integrating insights into the paper. The writing phase spanned approximately two years (2023-2025). Following the initial draft, the paper underwent multiple rounds of refinement until a final version was approved by all team members.

This extensive work is presented in chapter 4 of this thesis. The material in this chapter has been submitted, on August 1^{th}, 2025, for publication and now it is under review phase. The manuscript, spanning 27 pages, adheres to the rigorous standards of the journal.

This paper constitutes a significant contribution to this field, offering a novel perspective on domain adequacy in ontologies by examining the interplay between data commitment and ontological commitment. It advances theoretical insights into how minimal sub-ontologies, or modules, can be effectively identified to ensure domain adequacy. By formally linking definitions of ontological and data commitments to domain adequacy and optimal domain adequacy, this work establishes a foundational framework for assessing the sufficiency of ontologies in representing and reasoning about real-world domains.

Chapter 5: Domain-adequacy of Ontologies with Statistically-defined Concepts.

Authors: Deemah Alomair, Yihai Chen, Yijie Wang, and Ridha Khedri.

The idea of investigating domain ontology adequacy for statistically defined concepts emerged during the formulation of our adequacy theory. While we are in the phase of developing the ontological domain adequacy theory, we extended our perspective to interoperate statistical language into the DIS framework. A key question arose: What if datascape concepts are defined using statistical language? Addressing this question highlighted the need for a comprehensive framework that integrates both regular and statistical concepts. Consequently, we designed a system framework and developed a language to ensure data adequacy for statistically defined concepts.

Through weekly discussions, we progressively refined our theory and systematically proved its properties. During this process, we proposed an approach to verify the relevance of ontology concepts that are defined using terms involving data elements and statistical language. The proposed approach enhances the domain adequacy of a given ontology concerning a specific dataset, leading to a more refined and relevant sub-ontology. Furthermore, we automated the process of assessing domain ontologies' adequacy by generating R programs that run within a DISEL plug-in. The latter is an ontology specification language based on DIS framework [Wang et al., 2022]. The automation system is implemented as a DISEL editor plug-in. Through a case study on a weather ontology, we illustrate the approach's application and demonstrate how the automated verification process results in a smaller, more domain-adequate ontology.

The writing phase spanned approximately two years (2023-2025). Deemah Alomair was responsible for generating the R-programs that process the dataset to compute the statistical measures or samples necessary for verifying the relevance of concepts. Yijie Wang extended this work by integrating the $R-$based library into the DISEL plug-in, ensuring compatibility and seamless integration. Additionally, Wang verified the relevance of statistical datascape concepts to the domain based on the underlying dataset. Deemah Alomair drafted the initial manuscript and submitted it for feedback from Dr. Ridha Khedri. Following the initial draft, the paper underwent multiple rounds of refinement until a final version was approved by all team members.

This extensive work is presented in chapter 5 of this thesis. The material in this chapter has been submitted, on August 14^{th}, 2025, for publication and now it is under review phase. The manuscript, spanning 30 pages, adheres to the rigorous standards of the journal.

This paper constitutes a significant contribution to the field, offering a novel perspective on the domain adequacy of the statistically defined datascape concept within the DIS framework. It examines the data commitment of these special types of concepts and derives the corresponding ontological commitment, collectively advancing theoretical insights into how minimal sub-ontologies, or modules, can be effectively identified to ensure domain adequacy.

1.9 Thesis Outline

This thesis is structured into multiple chapters to establish a comprehensive framework for addressing uncertainty in ontology-based systems. The organization of the thesis is as follows:

Chapter 1: Introduction. This chapter presents the fundamental purpose and scope of the research. It introduces the general context of data analytics, followed by a discussion of the research motivations, problem statement, and key challenges. The research questions and objectives that guide this study are outlined, and a summary of the main contributions is provided, linking them to relevant publications. The chapter concludes with an overview of the thesis structure.

Chapter 2: A Comprehensive Review of Uncertainty Modelling in Domain Ontologies. This chapter provides an in-depth survey of existing approaches to uncertainty modelling in ontology-based systems. It categorizes uncertainty-handling approaches based on their underlying formalisms, including fuzzy logic, probability theory, possibility theory, rough sets, paraconsistent logic, DST, and hybrid approaches. The chapter also introduces a guidance taxonomy, which serves as a roadmap for selecting suitable uncertainty-handling formalism in ontology frameworks. This work is considered a major contribution to this research as it covers the vast literature in this field.

Chapter 3: Possibilistic extension of Domain Information System (DIS) Framework. This chapter introduces and extends the DIS framework to incorporate quantitative possibility theory to handle uncertainty due to incomplete information. The chapter explains the motivation for choosing possibility theory, details the integration of weighted components into the DIS structure, and presents the mathematical formalization of the enhanced framework. It also explores how this extended framework improves the reasoning and expressiveness when dealing with incomplete or imperfect information in ontology-based systems.

Chapter 4: Towards a Cartesian Theory of Domain Adequacy of Ontologies. This chapter develops a novel theoretical framework for ensuring domain adequacy in ontology-based reasoning. It introduces the concepts of ontological commitments and data commitment, formalizing their roles in maintaining relevance and completeness within an ontology. The chapter distinguishes between objective reality and datascape, demonstrating how these aspects influence ontology domain adequacy. The proposed framework provides a theoretical foundation for addressing relevance uncertainty, ensuring that ontologies remain semantically appropriate and robust.

Chapter 5: Domain-adequacy of Ontologies with Statistically-defined Concepts. This chapter presents a framework that bridges the gap between statistical data representations and ontology-based reasoning. It focuses on data commitment mechanisms that ensure statistically derived concepts are accurately integrated into the ontological structure. The chapter provides a modular approach for handling such concepts, ensuring that ontologies remain adequate when incorporating concepts derived from statistical analysis.

Chapter 6: Conclusion and Future Work. This chapter summarizes the key contributions and findings of the thesis. It reflects on the key values of the proposed uncertainty modelling frameworks, discussing their theoretical and practical implications. Additionally, the chapter outlines the suggested future research directions and potential improvements.

Appendix: DISEL: A Language for Specifying DIS-based Ontologies. This appendix presents the DISEL language, designed for specifying ontologies within the DIS framework. It introduces the core syntactical constructs of DISEL and demonstrates their use through a simplified weather ontology example. Since DISEL was employed in the specification work discussed in chapter 5, it is included here to ensure the thesis remains self-contained and accessible. The inclusion of this material supports the understanding of the tools and methodologies underlying the contributions and findings of the thesis.

1.10 Conclusion

This chapter outlined the fundamental purpose and scope of this research, providing a structured foundation for the study. It began by establishing both the general and specific contexts, introducing key foundational concepts such as data analytics, ontology, and uncertainty types to contextualize the research within the broader domain. The chapter then presented the main motivations behind this work, emphasizing the challenges and limitations of the current field. This was followed by a clear articulation of the problem statement, defining the research gap that this study aims to address.

The discussion then transitioned into the core research questions that underpin the investigation. These questions were formulated to systematically explore how different uncertainty types impact reasoning within ontology-based systems and how formal methods can be leveraged to enhance expressiveness and accuracy in reasoning under uncertainty. To address these questions, the key research objectives were outlined, detailing the specific goals that guide this study.

A significant portion of this chapter was dedicated to positioning the research contributions within the field. A comprehensive overview of the contributions was provided, explaining how each aligns with the research objectives and demonstrating their dissemination through publications. Finally, this chapter concluded with a detailed outline of the thesis structure, providing a roadmap for the chapters that follow.

Bibliography

Luca Lista. *Statistical Methods for Data Analysis*. Springer Cham, 3 edition, 2023. DOI: https://doi.org/10.1007/978-3-031-19934-9.

Rohit Kenge. Machine Learning, its limitations, and solutions over IT. *International Journal of Information Technology Modeling and Computing*, 11:73–83, 10 2020. DOI: 10.5958/0975-8089.2020.00009.3.

Vikas Hassija, Vinay Chamola, Atmesh Mahapatra, Abhinandan Singal, Divyansh Goel, Kaizhu Huang, Simone Scardapane, Indro Spinelli, Mufti Mahmud, and Amir Hussain. Interpreting black-box models: A review on explainable artificial intelligence. *Cognitive Computation*, 16(1):45–74, 2024. DOI: 10.1007/s12559-023-10179-8.

Grigoris Antoniou and Frank van Harmelen. *Web Ontology Language: OWL*, pages 91–110. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. DOI: 10.1007/978-3-540-92673-3_4.

Bernardo Cuenca Grau. OWL 2 web ontology language: Direct semantics. 2009. URL `https://api.semanticscholar.org/CorpusID:62253854`.

Franz Baader, Ian Horrocks, and Ulrike Sattler. *Description Logics as Ontology Languages for the Semantic Web*, pages 228–248. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005. DOI: 10.1007/978-3-540-32254-2_14.

Donna R. Berryman. Ontology, epistemology, methodology, and methods: Information for librarian researchers. *Medical Reference Services Quarterly*, 38(3):271–279, 2019. DOI: https://doi.org/10.1080/02763869.2019.1623614.

Tom Gruber. *Ontology*, pages 2574–2576. Springer New York, New York, NY, 2018. ISBN 978-1-4614-8265-9. DOI: 10.1007/978-1-4614-8265-9_1318.

Shaker El-Sappagh, Francesco Franda, Farman Ali, and Kyung-Sup Kwak. SNOMED CT standard ontology based on the ontology for general medical science. *BMC Med Inform Decis Mak*, 18(1):76, 2018. DOI: https://doi.org/10.1186/s12911-018-0651-5.

Shivani Jain, KR Seeja, and Rajni Jindal. A fuzzy ontology framework in information retrieval using semantic query expansion. *International Journal of Information Management Data Insights*, 1(1):100009, 2021. DOI: https://doi.org/10.1016/j.jjimei.2021.100009.

Lisheng He, Wenjia Joyce Zhao, and Sudeep Bhatia. An ontology of decision models. *Psychological Review*, 129(1):49, 2022. DOI: 10.1037/rev0000231.

RV Karthik and Sannasi Ganapathy. A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce. *Applied Soft Computing*, 108:107396, 2021. DOI: https://doi.org/10.1016/j.asoc.2021.107396.

V Maniraj and R Sivakumar. Ontology languages – a review. *International Journal of Computer Theory and Engineering*, 2(6):887, 2010. DOI: 10.7763/IJCTE.2010.V2.257.

Yunjiao Xue, Hamada H. Ghenniwa, and Weiming Shen. A frame-based ontological view specification language. In *The 2010 14th International Conference on Computer Supported Cooperative Work in Design*, pages 228–233, 2010. DOI: 10.1109/CSCWD.2010.5471972.

Dan R. Corbett. *Graph-Based Representation and Reasoning for Ontologies*, pages 351–379. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008. ISBN 978-3-540-78293-3. DOI: 10.1007/978-3-540-78293-3_8.

Michael R. Genesereth. Knowledge interchange format. In *International Conference on Principles of Knowledge Representation and Reasoning*, 1991. URL `https://api.semanticscholar.org/CorpusID:32333782`.

Michael Kifer, Georg Lausen, and James Wu. Logical foundations of object-oriented and frame-based languages. *Journal of the ACM*, 42(4):741–843, 01 2000. DOI: 10.1145/210332.210335.

Jeff Z. Pan. *Resource Description Framework*, pages 71–90. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. DOI: 10.1007/978-3-540-92673-3_3.

Diana Kalibatiene and Olegas Vasilecas. Survey on ontology languages. In *International Conference on Business Informatics Research*, pages 124–141, Berlin, Heidelberg, 2011. Springer, Springer Berlin Heidelberg. DOI: https://doi.org/10.1007/978-3-642-24511-4_10.

Guohui Xiao, Diego Calvanese, Roman Kontchakov, Domenico Lembo, Antonella Poggi, Riccardo Rosati, and Michael Zakharyaschev. Ontology-based data access: a survey. In *in Proceedings of the 27th International Joint Conference on Artificial Intelligence*, IJCAI'18, page 5511–5519. AAAI Press, 2018. ISBN 9780999241127. DOI: https://doi.org/10.24963/ijcai.2018/777.

Mustafa Jarrar and Robert Meersman. Ontology engineering—the DOGMA approach. In *Advances in Web Semantics I*, pages 7–34. Springer, 2008. DOI: https://doi.org/10.1007/978-3-540-89784-2_2.

Alicia Marinache, Ridha Khedri, Andrew LeClair, and Wendy MacCaull. A data-centered framework for domain knowledge representation. In *2021 Reconciling Data Analytics, Automation, Privacy, and Security: A Big Data Challenge (RDAAPS)*, pages 1–8. IEEE, 2021. DOI: 10.1109/RDAAPS48126.2021.9452007.

Ridha Khedri Andrew LeClair. An algebraic approach to ontology modularization and knowledge refinement. *IEEE Transactions on Knowledge and Data Engineering*, pages 1–26, 2025. Under review.

Nicola Guarino, Daniel Oberle, and Steffen Staab. *What Is an Ontology?*, pages 1–17. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. DOI: 10.1007/978-3-540-92673-3_0.

Antonella Poggi, Domenico Lembo, Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini, and Riccardo Rosati. Linking data to ontologies. In *Journal on Data Semantics X*, pages 133–173, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg. DOI: https://doi.org/10.1007/978-3-540-77688-8_5.

Hans-Ulrich Krieger and Christian Willms. Extending OWL ontologies by Cartesian types to represent n-ary relations in natural language. In *Proceedings of the 1st Workshop on Language and Ontologies*, London, UK, April 2015. Association for Computational Linguistics. URL `https://aclanthology.org/W15-0401/`.

Alicia Marinache. *Syntax and Semantics of Domain Information System and its usage in conjecture verification*. PhD thesis, School of Graduate Studies, McMaster University, Hamilton, Ontario, Canada, August 2025.

M. Fareh. Modeling incomplete knowledge of semantic web using bayesian networks. *Applied Artificial Intelligence,*, 33(11):1022–1034, 2019. DOI: https://doi.org/10.1080/08839514.2019.1661578.

Loukia Karanikola. *Managing Uncertainty and Vagueness in Semantic Web*. PhD thesis, National and Kapodistrian University of Athens, Department of Informatics and Telecommunications., Athens, Greece, 2018.

Z. M. Ma, Fu Zhang, Hailong Wang, and Li Yan. An overview of fuzzy description logics for the semantic web. *The Knowledge Engineering Review*, 28(1):1–34, 2013. DOI: 10.1017/S0269888912000306.

Patric Bosc and Henri Prade. *An Introduction to the Fuzzy Set and Possibility Theory-Based Treatment of Flexible Queries and Uncertain or Imprecise Databases*, pages 285–324. Springer US, Boston, MA, 1997. ISBN 978-1-4615-6245-0. DOI: 10.1007/978-1-4615-6245-0_10.

Sanjay Kumar Anand and Suresh Kumar. Uncertainty analysis in ontology-based knowledge representation. *New Generation Computing*, 40(1):339–376, 2022. DOI: https://doi.org/10.1007/s00354-022-00162-6.

Paolo Ceravolo, Ernesto Damiani, and Marcello Leida. Which role for an ontology of uncertainty? In *International Workshop on Uncertainty Reasoning for the Semantic Web*, volume 423, 2008. URL `https://api.semanticscholar.org/CorpusID:13393169`.

Giorgos Stoilos, Giorgos B Stamou, Vassilis Tzouvaras, Jeff Z Pan, and Ian Horrocks. Fuzzy OWL: Uncertainty and the semantic web. In *OWLED*, volume 188, 2005. URL `https://api.semanticscholar.org/CorpusID:1250151`.

A Waterson and A Preece. Verifying ontological commitment in knowledge-based systems. *Knowledge-Based Systems*, 12(1):45–54, April 1999. ISSN 0950-7051. DOI: 10.1016/S0950-7051(99)00007-6.

Didier Dubois and Henry Prade. *Possibility Theory and Its Applications: Where Do We Stand?*, pages 31–60. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. ISBN 978-3-662-43505-2. DOI: 10.1007/978-3-662-43505-2_3.

Kari Sentz and Scott Ferson. *Combination of evidence in Dempster-Shafer theory*, volume 4015. Sandia National Laboratories Albuquerque, 2002. DOI: 10.2172/800792.

Seiki Akama, Yasuo Kudo, and Tetsuya Murai. *Overview of Rough Set Theory*, pages 7–60. Springer International Publishing, 2020. DOI: 10.1007/978-3-030-29566-0_2.

Graham Priest. *Paraconsistent Logic*, pages 287–393. Springer Netherlands, Dordrecht, 2002. ISBN 978-94-017-0460-1. DOI: 10.1007/978-94-017-0460-1_4.

Radha G Laha and Vijay K Rohatgi. *Probability theory*. Courier Dover Publications, 2020.

Hans-Jürgen Zimmermann. *Fuzzy set theory—and its applications*. Springer Science & Business Media, 2011. DOI: https://doi.org/10.1007/978-94-010-0646-0.

Yijie Wang, Yihai Chen, Deemah Alomair, and Ridha Khedri. DISEL: A language for specifying dis-based ontologies. In *The 15th International Conference on Knowledge Science, Engineering and Management (KSEM 2022)*, Lecture Notes in Artificial Intelligence, pages 1–16, Singapore, 2022. Springer. DOI: 10.1007/978-3-031-10986-7_13.

Chapter 2

A Comprehensive Review of Uncertainty Modelling in Domain Ontologies

This chapter presents a comprehensive literature review of uncertainty modelling approaches in domain ontologies, covering the period from 2010 to 2024. It addresses the first three objectives of this thesis: to identify and classify the different types of uncertainties that arise in ontologies, and their handling formalisms, and to provide clear guidelines for the best formalism selection. The review is guided by a set of guiding questions that inform the analysis, including the motivations for modelling uncertainty, the formalisms employed, and the types of uncertainty addressed in the selected studies.

This work provides the theoretical and contextual foundation for the subsequent contributions in the thesis. It introduces a structured taxonomy that identifies the primary carriers of uncertainty in ontologies and relates them to both the types of uncertainty (e.g., incompleteness, imprecision, or inconsistencies) and the corresponding modelling formalisms, such as possibility theory, fuzzy logic, or probability theory.

A Comprehensive Review of Information Uncertainty Modelling in Domain Ontologies

DEEMAH ALOMAIR[*], McMaster University, Canada
RIDHA KHEDRI, McMaster University, Canada
WENDY MACCAULL, St. Francis Xavier University, Canada

Domain ontologies are essential for representing and reasoning about knowledge, yet addressing information uncertainty within them remains challenging. This review surveys approaches to modelling information uncertainty in domain ontologies from 2010 to 2024. It categorizes modelling formalisms, identifies information uncertainty types, and analyzes how information uncertainty is integrated into ontology components. It reviews reasoning techniques and emerging methods, including Machine Learning and Natural Language Processing. The review examines languages, tools, and evaluation strategies. The purpose is to map the landscape of information uncertainty modelling in domain ontologies, highlight research gaps and trends, and provide structured guidance for selecting suitable approaches.

CCS Concepts: • **Computing methodologies → Ontology engineering**; **Description logics**; **Probabilistic reasoning**; **Vagueness and fuzzy logic**; **Reasoning about belief and knowledge**; • **Information systems → Ontologies**.

Additional Key Words and Phrases: Information Uncertainty modelling, Ontological uncertainty taxonomy, Possibility theory, Dempster-Shafer theory, Rough set theory, Paraconsistent logic, Probability theory, Fuzzy set theory.

ACM Reference Format:
Deemah Alomair, Ridha Khedri, and Wendy MacCaull. 2024. A Comprehensive Review of Information Uncertainty Modelling in Domain Ontologies. *ACM Comput. Surv.* XX, X, Article XXX (December 2024), 62 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

In the era of data-driven analytics, domain ontologies have emerged as pivotal tools for modelling structured knowledge across various domains, enabling applications in fields such as healthcare [1], information retrieval [2], decision making [3], and e-commerce [4]. An ontology, as formally defined in the literature [5], is an explicit, formal specification of a shared conceptualization. Informally, a domain ontology, in most cases, consists of a set of concepts, relationships among these concepts, individuals, and constraints, representing knowledge within a specific domain.

While classical ontologies are widely used in representing crisp, well-defined knowledge, their inherent rigidity makes them ill-suited for uncertain information [6]. Information uncertainty remains an inherent aspect of knowledge and manifests in various forms, such as incompleteness, vagueness, ambiguity, imprecision, and inconsistency [7]. Consequently, researchers have extended classical ontologies and their underlying structures, such as Description Logic

[*] Also with Imam Abdulrahman Bin Faisal University.

Authors' Contact Information: Deemah Alomair, McMaster University, Hamilton, Ontario, Canada, alomaird@mcmaster.ca; Ridha Khedri, McMaster University, Hamilton, Ontario, Canada, khedri@mcmaster.ca; Wendy MacCaull, St. Francis Xavier University, Antigonish, Nova Scotia, Canada, wmaccaul@stfx.ca.

Manuscript submitted to ACM

(DL), to handle uncertainty using several formalisms. Specifically, possibility theory (e.g., [8]), Dempster-Shafer Theory (DST) (e.g., [9]), rough set theory (e.g., [10]), paraconsistent logic (e.g., [11]), probability theory (e.g., [12]) and fuzzy logic (e.g., [13]). Although classical ontologies have achieved substantial success, their ability to manage information uncertainty remains a significant challenge.

1.1 Motivation

The challenge of incorporating information uncertainty into ontological frameworks is well-documented [6, 14], and a range of approaches have been proposed for representing and reasoning about information uncertainty in domain ontologies. However, a significant gap remains when considering review papers: no existing review systematically examines information uncertainty modelling across all major uncertainty formalisms (e.g., possibility theory, DST, rough set theory, paraconsistent logic, probability theory, and fuzzy set theory). Existing reviews typically focus on individual information uncertainty formalisms, such as fuzzy-based ontologies [15, 16, 17] or Bayesian Network (BN)-based probabilistic ontologies [14, 18]. For the most part, these reviews are descriptive and lack systematic methodologies. An exception is the systematic review presented in [19], which concentrates exclusively on BN-based probabilistic ontologies. Our search process did not reveal any review, beyond BN, that addresses probabilistic ontologies more broadly. An exception is the review [20], which offers broader coverage addressing fuzzy, probabilistic, and possibilistic-based ontologies.

1.2 Contributions

This systematic review addresses a critical gap in the literature by providing a comprehensive review of information uncertainty modelling in domain ontologies. First, it introduces a novel taxonomy that offers a structured framework for understanding and addressing different types of information uncertainty within domain ontology frameworks and the formalisms to handle them. Second, the review summarizes the theoretical foundations of six information uncertainty formalisms: possibility theory, DST, rough set theory, paraconsistent logic, probability theory, and fuzzy set theory, providing essential insights for researchers and practitioners. Third, it presents a systematic methodology to address specific guiding questions about various aspects of information uncertainty modelling within domain ontologies. It outlines the guiding questions and the inclusion and exclusion criteria. It elaborates on the results of the guiding questions in a structured manner. The review synthesizes and compares contributions from related surveys and reviews, critically analyzing their scope and methodologies.

Overall, this study provides a systematic review that encompasses the full spectrum of information uncertainty modelling formalisms within domain ontologies to serve as a practical roadmap for ontology engineers and developers, particularly those focused on designing, modelling, or reasoning about ontological systems under uncertain information. This review clarifies the current state of the field and reveals gaps, limitations, and emerging trends. It aims to facilitate informed decision-making in the selection of uncertainty-modelling techniques, ultimately contributing to the development of more expressive and resilient ontology-based systems. This review provides a robust and consolidated reference point for the literature concerned with modelling uncertain knowledge within domain ontologies.

1.3 Outline of Paper

The remainder of this paper is organized as follows: The theoretical foundations and formalisms for addressing information uncertainty are provided in section 2. The review methodology, including the guiding questions and inclusion/exclusion criteria, is detailed in section 3. Findings and answers to the guiding questions are presented in section 4. Related reviews are summarized in section 5. Discussion of key insights is provided in section 6, and a conclusion is drawn in section 7. In the supplementary materials, we provide detailed categorizations of the approaches based on their formalism and scope of focus, along with the bibliography of the excluded studies.

2 Preliminaries and Definitions

This section overviews the theoretical background on uncertainty types and their management formalisms. The nature of uncertainty is detailed in subsection 2.1, while relevant formalisms are discussed in subsection 2.2.

2.1 Information Imperfection and Domain Ontology

Uncertainty is a multifaceted concept without a single universal definition, with multiple classifications proposed across disciplines and contexts [21]. In modelling and decision-making, a common distinction is based on the nature of uncertainty, namely into *aleatory and epistemic. Aleatory uncertainty* arises from intrinsic randomness or variability inherent in the observed world, reflecting irreducible nondeterministic behaviour. It is therefore often referred to as irreducible uncertainty, inherent uncertainty, variability, or stochastic uncertainty. In contrast, *epistemic uncertainty* results from lack of knowledge or insufficient information about the system or environment and is, in principle, reducible through additional data collection or improved modelling. In the literature, it is also referred to as reducible uncertainty, subjective uncertainty, or cognitive uncertainty [22]. Another prominent classification focuses on the occurrence of uncertainty within system design, distinguishing among *data uncertainty, model uncertainty, and structural uncertainty* [23]. Within this scheme, the underlying nature of uncertainty (aleatory or epistemic) may vary depending on the specific case. Data uncertainty can be further divided into, for example, algorithmic uncertainty [24] and experimental uncertainty [25]. From an information-theoretic perspective, uncertainty is understood to originate from deficiencies in the available information itself, often referred to as uncertainty-based information or imperfect information [26]. Such information deficiencies are incompleteness, imprecision, vagueness, ambiguity, or inconsistency. It is worth noting that classification schemes differ on how to situate these deficiencies: some, such as [22], regard all forms of imperfect information as epistemic uncertainty. while others [27, 28] treat imperfect information more broadly, encompassing both aleatory and epistemic aspects, which may manifest in different forms (e.g., inconsistency, incompleteness, vagueness, and imprecision). In this review, we focus primarily on imperfect information, which, for simplicity, we will hereafter refer to as *uncertainty*. We provide a detailed description of its different forms below.

Incompleteness arises from partial knowledge about the domain, leading to uncertainty across possible interpretations. To address this, an estimation degree is calculated for possible worlds. For example, while we cannot be certain it will snow tomorrow, we can estimate its likelihood. For further details on incompleteness, refer to [29].

Imprecision is the lack of exactness or specificity in representing information, leading to approximate or qualitative descriptions. For example, phrases like "the temperature is around 25 degrees" or "John's weight is either 65 or 67 kg" illustrate imprecision. For further details on imprecision, refer to [30].

Vagueness, also known as fuzziness, occurs when a term or proposition lacks a precise evaluation or clearly defined boundaries. In other words, there is no universally accepted meaning for the concept. Terms such as "old," "young," and "tall" may vary depending on the context and do not always imply a certain assessment. Graded propositions under fuzzy set theory are used to provide a partial truth of such propositions. For further details on vagueness, refer to [31].

Ambiguity occurs when a term has multiple meanings. For instance, the phrase "the food is hot" could mean either "warm" or "spicy". Therefore, it is important to determine the appropriate meaning in a specific context. For further details on ambiguity, refer to [30].

Inconsistency, also known as contradiction or conflict, refers to a conflict between the meanings of multiple statements. For example, having both sentences, "John is travelling to Toronto at 3:00 PM tomorrow" and "John is travelling to Montreal at 3:00 PM tomorrow", leads to inconsistency. For further details on inconsistency, refer to [28].

Despite increasing research addressing uncertainty in ontology-based systems, most studies focus on specific types of uncertainty in isolation, concentrating on particular formalisms and modelling approaches. While some works, like [7], offer partial classifications of ontological uncertainty covering aspects such as incompleteness, vagueness, inconsistency, and management formalisms, to date, however, no comprehensive, visual taxonomy exists that integrates major uncertainty types, their ontological locations, and their corresponding management formalisms. Foundational insights remain scattered across the literature but lack synthesis into a unified, visual framework. To address this gap, we developed the taxonomy presented in Figure 1 to order the large set of works covered in the literature in the period of 2010-2024. The purpose of the taxonomy, which is grounded in an extensive review of the literature, is to serve as a conceptual roadmap that clarifies how various uncertainty types commonly arise and are addressed within ontological frameworks.

This taxonomy classifies uncertainty into two main types based on where it arises in domain ontologies: *concept uncertainty* and *information uncertainty*. Conceptual uncertainty has three key carriers: First *Semantic ambiguity*, where a concept's meaning depends on the context (e.g., *"Apple"* as fruit or technology company). Second *Attribute of concept*, where uncertainty occurs in concept definitions, especially when vague or fuzzy attributes characterize concepts (e.g., the concept *"NoisyArea"* defined by the attribute *"HighNoiseLevel"*, which has no strict cutoff). Third, *Relationship between concepts*, such as uncertainty in the subsumption relation (e.g., Intern $\sqsubseteq$ Employee, might be uncertain). Although classical DL cannot inherently capture such uncertainties, these carriers can arise when modelling a domain using DL-based TBox axioms. For example, semantic ambiguity may appear in the TBox axiom "$Apple \sqsubseteq \top$", with context determining the intended interpretation. Uncertain attributes can occur in data property restrictions, like $\text{NoisyArea} \equiv \text{Area} \sqcap \exists\, \text{HighNoiseLevel.} \geq h$, where h is context-dependent. *HighNoiseLevel* is a data property (in DL) akin to an attribute in database ontologies or relational-based models, which is used to define the concept *"NoisyArea"*. The third carrier can occur in the subsumption TBox axioms. For example: Son $\sqsubseteq$ Child, which may not hold universally.

Information uncertainty addresses the challenge of mapping instances to their appropriate concepts or relations. For example, linking the instance *"John"* to the concept *"Student"* or relating *"John"* to the *"AcL"* company via the relation *"work_at"* involves uncertainties in associating individuals with the correct concepts or relationships. These types of information uncertainty cannot be modelled directly by DL-based ABox assertions. For instance, the assertion "$John : Student$" indicates that *"John"* is a member of the *"Student"* concept, but this membership might be uncertain or context-dependent. Similarly, the relation assertion "$(John, AcL) : work\_at$" may also be uncertain. Since classical DL does not inherently represent such uncertainty, these TBox and ABox axioms need to be extended by suitable formalisms to incorporate certainty or belief degrees that can express and support the uncertainty in these axioms.

These uncertainties stem from incomplete, imprecise, vague, or inconsistent information, as shown by arrows in our taxonomy. Every uncertainty carrier in ontology might be caused by one or more types of uncertainty. To handle these different types of uncertainty, one or a combination of formalisms can be employed. In particular, *incomplete information* is often managed using probability theory, DST, or possibility theory. *Imprecise* and *vague information* are commonly addressed through rough set theory or fuzzy set theory. *Inconsistency* is typically addressed using DST, paraconsistent logic, probability theory, or possibility theory. These formalisms are depicted by dashed arrows in the taxonomy, while dotted arrows represent subcategories within some of these formalisms. Overall, this taxonomy provides a comprehensive framework for identifying different types of uncertainty and applying the most appropriate formalism to manage and resolve them within an ontological system.

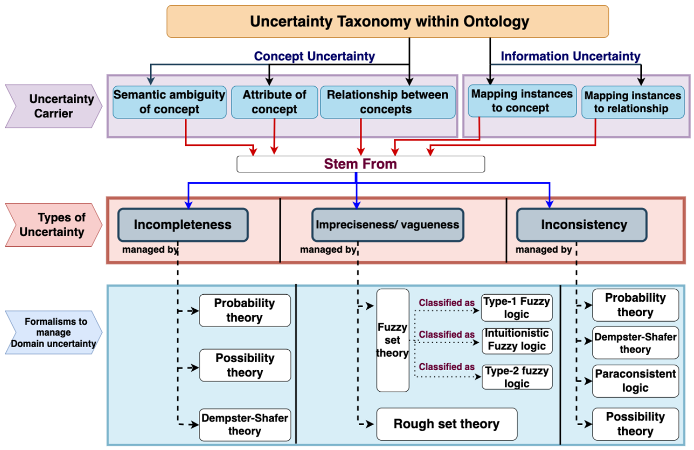


Fig. 1. Ontological Uncertainty Taxonomy.

2.2 Theoretical Background on Formalisms to Handle Uncertainty

This subsection briefly introduces the six formalisms used to handle uncertainty in domain ontologies: Fuzzy set theory, probability theory, DST, possibility theory, rough set theory, and paraconsistent logic.

2.2.1 Fuzzy Set Theory and Fuzzy Logics. The concept of fuzzy set theory, developed by L. A. Zadeh [32], extends the classical set theory to handle partial truth. In fuzzy set theory, the truth value of a fuzzy proposition x is assigned a real number between 0 and 1, determined by a membership function μ_F, defined as $\mu_F(x) : U \rightarrow [0, 1]$, where U is the universal set [33]. This function (i.e., μ_F) indicates the degree to which an object belongs to the set. A higher membership grade suggests that the object is more suitable for the set. A fuzzy set F can be defined based on μ_F as $F = \{(x, \mu_F(x)) \mid x \in U\}$, where F represents the fuzzy set, which is a collection of elements and their membership degrees, while μ_F represents the membership function that assigns degrees of membership to elements of the fuzzy set F [33, 34].

Fuzzy logic, based on fuzzy set theory, is a type of many-valued logic. There are two widely accepted interpretations of fuzzy logic: the broad view and the narrow view. The broad view applies fuzzy logical connectives and concepts from fuzzy set theory to develop techniques for "approximate reasoning," and it has various applications, such as fuzzy controllers and fuzzy IF-THEN rules [35]. The narrow interpretation focuses on constructing deductive systems within the context of fuzzy logic, involving a comparative understanding of truth similar to classical mathematical logic, including propositional and predicate calculi, as well as considerations of axiomatization, (in)completeness, complexity, and related aspects. This interpretation is often referred to as Mathematical Fuzzy Logic (MFL) [35]. Both Type-1 and Type-2 fuzzy logic, defined below, can be studied from both of these viewpoints.

Type-1 fuzzy logic, which is based on crisp and static membership functions, is the most widely known type of fuzzy logic. Type-2 fuzzy logic, on the other hand, involves fuzzy membership grades and interval representations [36, 37]. In fuzzy logic, norm-based operators model uncertainty in a structured way. These are Triangular Norms (T-norms)

Table 1. Combination functions of various fuzzy logics

	Łukasiewicz logic	Gödel logic	Product logic	Zadeh logic
$a \otimes b$	max(a+b - 1, 0)	min(a, b)	$a \cdot b$	min(a, b)
$a \oplus b$	min(a+b, 1)	max (a, b)	a+b - $a \cdot b$	max(a, b)
$a \Rightarrow b$	min(1-a+b, 1)	$\begin{cases} 1, & \text{if } a \leq b, \\ b, & \text{otherwise} \end{cases}$	min(1, b/a)	max(1-a, b)
$\ominus a$	1 - a	$\begin{cases} 1, & \text{if } a = 0 \\ 0, & \text{otherwise} \end{cases}$	$\begin{cases} 1, & \text{if } a = 0 \\ 0, & \text{otherwise} \end{cases}$	1- a

(fuzzy intersections), Triangular Conorms (T-conorms) (fuzzy unions), also known as S-norms, fuzzy negations, and fuzzy implications [38]. A T-norm is a function $T : [0,1] \times [0,1] \rightarrow [0,1]$ satisfying commutativity, associativity, monotonicity, and identity ($T(x,1) = x$). Its dual, a T-conorm, has similar properties but with 0 as its neutral element. Fuzzy negation, denoted by $\neg$, is a function $N : [0,1] \rightarrow [0,1]$ satisfies $N(0) = 1, N(1) = 0$, and antitonicity. Fuzzy implications include S-implications (also called the Kleene-Dienes implications), defined using T-conorms and fuzzy negation $I_S(x,y) = S(\neg x, y)$ [39], and R-implications (also called the residual implications, or the residua) are defined as follows: $R(x,y) = sup\{z \in [0,1] \mid T(x,z) \leq y\}, \forall x, y \in [0,1]$, where T is a left-continuous T-norm [40]. For a thorough review of fuzzy-based implications, we refer the reader to [39]. Gödel logic (G), Product logic ($\sqcap$), Łukasiewicz logic (Ł), and Zadeh logic are examples of norm-based fuzzy logic [41]. For a comprehensive understanding of these logics, refer to [42]. A comparative overview of these logics is provided in Table 1.

Intuitionistic fuzzy sets, which are a generalization of fuzzy sets, specify the degrees of membership μ_F, the degree of non-membership V_F, and the degree of hesitation π_F for an element to a set. The hesitation degree measures the level of uncertainty in assigning the membership degree of an element to a fuzzy set [43].

2.2.2 Probability Theory. Probability theory is a branch of mathematics that provides a framework for understanding and reasoning about random events [44]. It deals with both quantitative and qualitative probabilities. Quantitative probability is expressed as numerical values ranging from 0 to 1, while qualitative probability uses ordinal language to describe the likelihood of an event.

In probability theory, a sample space S is defined as the set of all possible outcomes of a random experiment, which might be finite or infinite, depending on the nature of the experiment. For example, when rolling a six-sided die, the sample space is $S = \{1, 2, 3, 4, 5, 6\}$. Each individual result from a trial of the experiment is called an outcome. In the die rolling example, each face value (1 through 6) is an outcome. An event is any subset of the sample space, representing one or more outcomes. For instance, the event "rolling an even number" corresponds to the subset $A = \{2, 4, 6\}$.

A probability measure P assigns a numerical value to each event, representing the likelihood of that event occurring. This function must satisfy the following properties [44]: (1) $0 \leq P(A) \leq 1, \forall A \subseteq S$. (2) $P(S) = 1$, and (3) for any pairwise disjoint events $A_1, A_2, ..., A_n \subseteq S$, $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$. This is known as the axiom of (finite) additivity. A probability model consists of a nonempty set called the sample space S, a collection of events that are subsets of S, and a probability measure P assigning a probability between 0 and 1 to each event.

A random variable is a function from the sample space S to the set $\mathbb{R}$ of all real numbers, i.e., $X : S \longrightarrow \mathbb{R}$. It assigns a numerical value to each outcome in the sample space. A random variable can be either discrete or continuous. A discrete random variable takes values from a finite or countably infinite set (e.g., rolling a six-sided die). A continuous random variable takes values from an uncountable range or interval (e.g., the height of a person in a population).

The probability distribution of a random variable indicates the likelihood of each possible value it can take. This is characterized by specific functions: a Probability Mass Function (PMF) for discrete random variables, which assigns probabilities to each value the variable can take. For example, the PMF of rolling a die is $P(X = k) = \frac{1}{6}, k = 1, 2, .., 6$. A Probability Density Function (PDF) is used for continuous random variables, where probabilities are defined over intervals rather than single points. More detailed explanations and foundational concepts on probability distributions and their associated functions can be found in [44].

The conditional probability $P(A \mid B)$ expresses the probability of event A occurring given that B has occurred, and is defined as [44]:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \text{ provided } P(B) \neq 0.$$

From this, Bayes' Theorem is derived, which provides a way to express the conditional probability of A given B in terms of the reverse conditional probability $P(B \mid A)$:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Conditional probability allows us to compute the joint probability of $P(A \cap B)$, which represents the likelihood of two events A and B occurring together and expressed as: $P(A \cap B) = P(A)P(B \mid A)$.

An extension to the traditional First-Order Logic (FOL) [45] is Probabilistic First-Order Logic (PFOL) [46]. In PFOL, probabilities can be associated with logical statements and predicates, as opposed to FOL, which only deals with deterministic predicates. A probabilistic predicate can be expressed as $\forall x(\Phi(x) \rightarrow Pr(\Psi(x)) = p)$, where $\Phi(x)$ represents the premise predicate, and $\Psi(x)$ represents the consequent predicate. This statement means that for every x, if $\Phi(x)$ is true, then the probability "Pr" that $\Psi(x)$ it true is exactly p. Thus, p quantifies the likelihood of $\Psi(x)$ being true under the condition that $\Phi(x)$ holds. Similar to FOL, the grounding process is crucial in PFOL. Grounding involves generating ground predicates or ground atoms by instantiating variables with constants. This process relies on the Herbrand logic [47], which defines a universe of all possible ground terms constructed using the function symbols and constants of the language. For instance, in a language with constants a and b and a predicate $\Phi(x)$, ground predicates such as $\Phi(a)$ and $\Phi(b)$ can be formed. In PFOL, probabilities can be assigned to these ground predicates. For instance, one can formulate a probabilistic assertion like: "If $\Phi(a)$ is true, then $\Phi(b)$ is true with a probability p".

In [48], an analysis framework of PFOL is presented. This framework differentiates between Type-1 and Type-2 PFOL through statistical and epistemic probability, receptively. Type-1 PFOL deals with statistical probabilities and is based on objective frequencies or distributions of attributes. This type of probability considers the likelihood of a particular situation occurring within a population and is grounded in facts. On the other hand, Type-2 PFOL is concerned with epistemic probabilities, which represent the subjective degree of belief an individual has in various possible worlds or interpretations. Therefore, Type-1 PFOL is linked to statistical (objective) probability, while Type-2 PFOL is associated with epistemic (subjective) probability, highlighting the distinction between empirical observations and personal beliefs.

Probabilistic Graphical Models (PGMs) [49] are powerful frameworks for representing and reasoning about uncertain information in a structured way. These models combine concepts from graph theory and probability theory to provide a compact representation of complex systems involving uncertainty. Examples of these models include Dynamical Uncertainty Causal Graphs (DUCG) [50], Conditional Random Fields (CRF) [51], Markov Network (MN) [52], BN [53] and its extensions such as Multi-Entity Bayesian Networks (MEBN) [54], and Object Oriented Bayesian Networks (OOBN) [55]. These models are used to represent the probabilistic relationships between variables in the graph. Among these, BN and MN are PGMs most commonly used in conjunction with ontologies. For a thorough description of PGMs, refer to [49].

2.2.3 Dempster-Shafer Theory. The DST, also known as evidence theory or theory of belief functions, is a mathematical framework for modelling epistemic uncertainty [9]. It emerged as an alternative to traditional probabilistic theory and serves as a generalization of probability theory in a finite discrete space. One of the key features of DST is the combination of evidence obtained from multiple sources and the modelling of conflict among them [9]. The theory allows for a higher level of abstraction when interpreting evidence and provides a framework for explicitly addressing uncertainty and conflict in decision-making processes.

Mathematically, the representation of DST relies on three primitive functions: Basic Probability Assignment (BPA) or mass function (m), belief or lower bound probability function (Bel), and plausibility or upper bound probability function (Pl). The concept of mass value is crucial in the DST as it represents the belief allocated directly to a specific subset of the frame of discernment Θ, (i.e., the set of all possible outcomes or hypotheses). In DST, hypotheses are propositions or statements about the domain that may compromise one or more events. The mass function m assigns a value to each subset A of Θ and must satisfy specific conditions to ensure a valid distribution of belief. Firstly, m is defined as a mapping from the power set of Θ to the interval $[0, 1]$, expressed mathematically as $m : 2^{\Theta} \longrightarrow [0, 1]$. Secondly, the mass assigned to the empty set must be zero, which means that $m(\emptyset) = 0$. Lastly, the total mass distributed across all subsets of Θ must sum to 1, expressed as $\sum_{S \subseteq \Theta} m(S) = 1$. For any set $A \subseteq \Theta$, $m(A)$ denotes the measure of evidence of A. However, it does not provide information about the subsets of A. To analyze the measure of evidence for any subset of A, we need to obtain the mass value for the subset itself.

From $m(A)$ we can derive $Bel(A)$, and $Pl(A)$. $Bel(A) = \sum_{S \subseteq A} m(S)$ is the sum of all masses of A's proper subsets, and represents its lower bound probability value. $Pl(A) = \sum_{S \cap A \neq \emptyset} m(S)$ is the total mass of all the sets that intersect with set A, and represents its upper bound probability value. Interdependency is a key feature of the DST, allowing us to derive the other two values of $\{(m(A), Bel(A), Pl(A))\}$ given one of them. For instance, Belief can be used to derive Plausibility: $Pl(A) = 1 - Bel(A')$, where A' is the complement of set A. The probability value of set A falls in the interval between the upper and lower bound values. A single probability value $P(A)$ is obtained when the plausibility measure $Pl(A)$ and the belief measure $Bel(A)$ are identical, consistent with classical probability theory, where $Bel(A) = Pl(A) = P(A)$. For more detailed information about the basic measures of DST, readers are referred to [9, 56].

In DST, there are several rules for combining evidence from different independent sources to support specific events or sets of events. The original rule for this purpose is called Dempster's rule of combination or orthogonal sum. It has faced criticism for not considering conflicts between different sources and assuming full agreement among them [9, 56, 57]. Efforts towards developing modified rules have been made, such as Yager's rule [58], Inagaki's unified combination rule [59], Zhang's center combination rule [60], and Dubois, Prade's disjunctive pooling [61, 62], among others (e.g., [9, 57, 58]).

Key distinctions exist between probability theory and DST. In probability theory, probabilities are assigned to specific events, whereas in DST, beliefs are assigned to non-singleton and to empty sets of events. This allows evidence in DST to be linked to multiple or sets of events, carrying meaning at a higher level of abstraction without assumptions about the events within the set. Another difference is how ignorance is handled: in probability theory, the probability of an unknown event is one minus the sum of the probabilities of other events in the considered sample space. In DST, ignorance is explicitly represented by assigning a mass to the vacuous set, which includes all hypotheses with no specific evidence [63]. For instances, consider a frame of discernment $U = \{A, B, C\}$, representing three possible states: A, B, and C. Assigning a vacuous mass means $m(U) = 1$ and $m(\{A\}) = m(\{B\}) = m(\{C\}) = 0$, indicating complete uncertainty with no evidence favoring any outcome.

The integration of DST with a graph-based theory for modelling uncertain knowledge using belief functions is explored in [64]. The resulting graph, referred to as the Directed Evidential Network (DEVN), offers a unified graphical and numerical framework for representing uncertain knowledge through belief functions.

2.2.4 Possibility Theory. Possibility theory is a mathematical framework for understanding and reasoning about incomplete and inconsistent knowledge [8]. This theory is classified into two major branches: qualitative and quantitative possibility theory. Qualitative possibility theory provides an ordinal interpretation of the possibility of a proposition to occur, without explicitly providing a numerical value. For instance, a meteorologist might state that it is very possible it will rain tomorrow, possible that it will be cloudy, and less possible that it will be sunny, without indicating precise numerical measures. In quantitative possibility theory, a numerical value is used to indicate the degree of possibility of a proposition. For instance, the meteorologist could state that the possibility of rain is 0.9, cloudy weather is 0.6, and sunshine is 0.2, with the values lying within a normalized range from 0 to 1.

Possibility distribution represents the epistemic state of an agent; that is, the agent's knowledge about the actual state of the world. A possibility distribution is a function that assigns to each element in a set of states a plausibility degree drawn from a totally ordered scale. Let S denote the set of states of affairs, which might be a finite or an infinite set. The distribution is formally defined as a mapping $\pi : S \to L$, where L is a totally ordered scale of plausibility values, often instantiated as the real unit interval $[0, 1]$, but it can also take other forms, such as a finite chain or the set of non-negative integers. The value $\pi(x)$ expresses how plausible the state $x \in S$ is, specifically [8]:

- $\pi(x) = 0$ indicates that x is impossible.
- $\pi(x) = 1$ indicates that x is totally possible (i.e., totally plausible).

The larger the value of $\pi(x)$, the more plausible the state x is considered to be. Assuming S is exhaustive, there must exist at least one $x \in S$ such that $\pi(x) = 1$. Multiple states x in S may also simultaneously have a possibility degree of 1. From the possibility distribution π, the possibility measure Π and necessity measure N can be defined over any subset $X \subseteq S$ as follows [8]:

$$\Pi(X) = \sup_{x \in X} \pi(x) \text{ and } N(X) = \inf_{x \notin X} (1 - \pi(x)).$$

Equivalently, the necessity measure can be expressed as: $N(X) = 1 - \Pi(X')$, where X' is the complement of X in S. The possibility measure $\Pi(X)$ quantifies the extent to which the set X is plausible or consistent with the available knowledge. Its dual, the necessity measure $N(X)$, expresses the degree of certainty that X is implied by that knowledge.

The maxitivity axiom expresses the fundamental property of the possibility measures in possibility theory, expressed as $\Pi(A \cup B) = \max(\Pi(A), \Pi(B))$. In contrast, the necessity measure fulfills the dual axiom, expressed as $N(A \cap B) = \min(N(A), N(B))$.

The minimum specificity principle, one of the foundational concepts of possibility theory, posits that any state not known to be impossible should remain under consideration. In the context of possibility theory, a possibility distribution π is considered at least as specific as another distribution π' if, for every state x in the universe S, the relationship $\pi(x) \leq \pi'(x)$ holds. This ordering implies that π is at least as informative and restrictive as π'. The possibilistic framework also provides ways to express cases of incomplete knowledge. In the case of complete knowledge, for some specific state x_i, $\pi(x_i) = 1$ and $\pi(x) = 0$ for all other states $x \neq x_i$, signifying that only x_i is possible. In contrast, complete ignorance is represented by assigning $\pi(x) = 1$ for every x in S, indicating that all states are equally possible [8].

An important strength of possibilistic logic lies in its ability to effectively manage inconsistency [65, 66]. This is achieved through the concept of an *inconsistency level*, denoted as *Inc(K)*, which quantifies and localizes the degree of contradiction within a knowledge base. This level represents the highest certainty degree at which a contraindication can be derived from the knowledge base K. Formally, the inconsistency level is expressed as $Inc(K) = max\{\alpha \mid K \vdash (\bot, \alpha)\}$. In this formula, the notation $K \vdash (\bot, \alpha)$ indicates that a contradiction ($\bot$) can be inferred from the knowledge base K with a certainty degree α. Thus, *Inc(K)* identifies the maximum level of certainty associated with contradictory information within K. If $Inc(K) = 0$, the knowledge base is consistent. However, when $Inc(K) \geq 0$, the possibilistic logic isolates inconsistency by excluding formulas with a certainty level less than or equal to *Inc(K).* The filtered subset, K_{cons}, is guaranteed to be consistent and serves as the foundation for subsequent reasoning.

The theory of possibility has various connections to other theories. Some (e.g., [67]) have suggested that it is a specific case of DST by equating the *necessity* and *possibility* measures of possibility theory to the *belief* and *plausibility* measures of DST. From a different perspective, Zadeh [67] has linked the possibility theory with fuzzy set theory, using possibility distributions to offer graded semantics for natural language statements. In his view, the possibility distribution of a set of alternatives is equivalent to a fuzzy set, and the membership function represents the degree of possibility. Looking at it from another angle, possibility theory is often linked to probability theory as an effective way to handle imprecise probability. In this interpretation, the probability is expressed as an interval using two measures (i.e., necessity and possibility) instead of a single point as in traditional probability theory. For a thorough analysis of possibility theory, its types, basic notions, and interpretations, we refer the reader to [8].

2.2.5 Rough Set Theory. Rough Set theory, also known as the theory of approximation [68], is commonly used when there is vagueness or imprecision in the data. It is applicable when, based on the available information, an object cannot be definitively categorized as a member of a set or its complement.

In an information system $S = (U, A)$, where U is a finite, non-empty set of objects called the universe, and A is a set of all attributes, we assign a set V_a for each attribute $a \in A$, representing the possible values of a. V_a is referred to as the domain of a. For any subset B of A, we can define a binary relation $I(B)$ on U, called the indiscernible relation [68]. This relation is an equivalence relation that divides U into a set of equivalent classes, where items within the same equivalence class cannot be distinguished based on the considered attributes. This relation is defined as follows:

$$\forall (x, y) \in U^2 \Big((x, y) \in I(B) \iff \big(\forall b \in B \cdot b(x) = b(y)\big) \Big),$$

where $b(x)$ represents the value of attribute b for the element x [68, 69]. $U/I(B)$, or simply U/B, represents the family of equivalence classes of $I(B)$, or the partition specified by B. These equivalence classes are referred to as B-elementary sets or B-granules. An equivalence class of $I(B)$, (a block of the partition U/B), containing x, is represented by $B(x)$, or $[x]_B$. If (x, y) belongs to $I(B)$, we shall state that x and y are B-indiscernible (indiscernible with regard to B).

Given the information system $S = (U, A)$, where $X \subseteq U$, and $B \subseteq A$, and based on the indiscernible relation $I(B)$, two distinct crisp sets of X can be defined: $B_*(X)$ and $B^*(X)$. These are referred to as the B-lower and B-upper approximations of X, respectively, and are defined as follows:

$$B_*(X) = \bigcup_{x \in U} \{B(x) : B(x) \subseteq X\}, \text{ and } B^*(X) = \bigcup_{x \in U} \{B(x) : B(x) \cap X \neq \emptyset\}.$$

The lower approximation or the positive region is the set of objects that can be classified with full certainty as elements of X. The upper approximation, also known as the negative region, indicates the set of objects that possibly belong to the target set X. The boundary region, $(BN)_B(X)$, is the difference between the upper approximation and lower approximation, expressed as $(BN)_B(X) = B^*(X) - B_*(X)$, which consists of the objects that cannot be classified as belonging or not belonging to the target set X. If the boundary region $(BN)_B(X) = \emptyset$, then X is considered crisp (exact) in relation to B. On the other hand, if $(BN)_B(X) \neq \emptyset$, then X is considered rough (inexact) in relation to B. The target set X is called a rough set if and only if the boundary region is not empty, that is, if $B^*(X) \neq B_*(X)$. Otherwise, the target set is considered to be crisp [68, 69].

2.2.6 Paraconsistent Logic. The core principle of paraconsistent logic is invalidating the *principle of explosion*, which states that from a contradiction (expressed as $A, \neg A$), any arbitrary statement Q can be inferred. In classical logic, this principle (i.e., $A, \neg A \vdash Q$) means that if both proposition A and its negation $\neg A$ are true, then any statement Q logically follows, regardless of its relevance to A. Paraconsistent logic is therefore a type of logic that is designed to avoid deriving a trivial conclusion from inconsistent premises [11]. Several types of logic exhibit paraconsistency, including many-valued paraconsistent logic, distance-based paraconsistent logic, and argumentation-based paraconsistent logic. Many-valued paraconsistent logics introduce special connective operators, adjust certain inference rules, and define additional truth values. This usually requires sacrificing some classical logic principles and inference rules. Examples of a well-known many-valued paraconsistent logics include Kleene three-valued logic [70], the logic of Paradox [71], Belnap four-valued logic [72], Quasi-classical logic [73], and Nelson four-valued logic [74]. A comprehensive comparison between these logics is provided in Table 2. Another class of paraconsistent logic is argumentation-based paraconsistent logic, which refers to a type of logical system that combines principles from argumentation theory with paraconsistent logic, as elucidated in detail in [75]. A third class of paraconsistent logic is distance-based paraconsistent logic, which was initially defined for a propositional language with a two-valued interpretation $\Lambda^2 = \{T, F\}$ (similar to classical logic). However, it can be naturally extended to many-valued logic. Pseudo-distance and aggregation functions play a role in determining the overall logical behaviour within the framework of distance-based logic, which is thoroughly explained in [76].

3 Methodology

In this section, we outline the methodology used in this review. This paper is based on the systematic review guidelines provided by Kitchenham and Charters [82]. The systematic review process involves the following steps: (1) Identification of the guiding questions, as detailed in subsection 3.1. (2) Clear articulation of the research strategy, found in subsection 3.2. (3) Establishment of reliable inclusion and exclusion criteria, outlined in subsection 3.3. (4) Provision of details regarding data collection, found in subsection 3.4 and data analysis, as discussed in subsection 3.5.

3.1 Guiding Questions

Our review aims to address various aspects of uncertainty modelling in domain ontologies by addressing several key guiding questions. These include the formalisms used to capture uncertainty, whether the paper is foundational or application-oriented, and the motivations behind modelling uncertainty, as well as the domain areas. We also explore the sources and types of uncertainty, as well as the location of uncertainty within the domain ontologies, based on our developed taxonomy. In addition, we examine how uncertainty is captured through formalisms, whether decidable reasoning procedures are employed, and the role of Machine Learning or Natural Language Processing

Table 2. Comparison of Key Types of Many-Valued Paraconsistent Logics

Logic type	Key features	Semantics & structure	Ref.
Kleene three-valued logic	- Truth values: Undetermined U (gap), True (T), False (F) - Defines two interpretations: strong (K_s) treats U as T, and weak (K_w) treats U as F - No tautologies	- Truth ordering: $F \leq_t U \leq_t T$ - knowledge ordering: $U \leq_k T$ and $U \leq_k F$ - Pre-bilattice structure based on these orderings.	[70, 77]
The logic of Paradox	- Truth values: True (T), False (F), Paradoxical P (glut) - Paradoxical value P indicating both true and false - Contains tautologies.	- Same truth tables as Kleene's (K_s).	[71]
Belnap's four-valued logic	- Truth values: True (T), False (F), None (N), Both (B) - No tautologies - Both logic of Paradox and the Kleene's three-valued logic are special cases: - Logic of Paradox omits "ignorance" (U) - Kleene's three-valued logic omits "contradiction" (P).	- Distributive bi-lattice structure $\langle B, \leq_t, \leq_k \rangle$ known as FOUR.	[72, 78]
Quasi-Classical logic	- Retains Belnap's four values - Restricts certain proof rules, such as disjunction introduction, from combining with decomposition rules like resolution - Logic weaker than classical logic - Connectives behave classically - No tautologies.	- Two semantics: strong and weak satisfaction.	[73, 79]
Nelson's four-valued logic	- Retains Belnap's four values. - Introduces strong negation ($\sim$) to express explicit falsehood - Allows independent interpretation of p and $\sim p$, enabling both to be true simultaneously - Contains tautologies.	- N4 semantics is represented by N4-lattices.	[74, 80, 81]

(NLP) in modelling uncertainty. Finally, we look at the tools and languages (e.g., Extensible Markup Language (XML), Semantic Web Rule Language (SWRL)) supporting these approaches, and highlight any given evaluations of the proposed approaches. These guiding questions serve as structured classification criteria commonly used in systematic literature reviews to guide data extraction and comparative analysis. Overall, this review paper addresses the following guiding questions:

- GQ1: What are the formalisms used to capture uncertainty?
- GQ2: Is the paper foundational or an application?
- GQ3-A: What are the given motivations for modelling uncertainty? GQ3-B: What is the specific problem under investigation (i.e., the domain area)?
- GQ4-A: What is the source of the uncertainty? GQ4-B: What are the specific types of uncertainty (e.g., impreciseness, incompleteness)?
- GQ5: How can the uncertainty raised in the ontology be classified based on our developed taxonomy (Figure 1)?
- GQ6: How is the formalism used to capture uncertainty?
- GQ7: Does the paper indicate that it uses a decidable reasoning procedure?
- GQ8: Does the paper contain any Machine Learning approach to help model uncertainty?
- GQ9: Does the approach to manage uncertainty involves NLP?

- GQ10-A: Is the approach supported by languages? GQ10-B: Is the approach supported by tools (e.g., using existing tools or building a new supporting tool)?
- GQ11: What is the given evaluation of the modelling approach (if any) (e.g., case study, experiment)?

3.2 Search Process

We conducted our search using the *Engineering Village* portal, which provides access to the *Compendex* and *Inspec* databases. Given the breadth of this area, we focused our search on the period between 2010 and 2024 to ensure comprehensive coverage of recent literature while maintaining a manageable scope across the full spectrum of relevant studies. This timeframe allows us to delimit our review to the most recent advancements in modelling and reasoning about uncertainty in ontologies. Using specific keywords, our initial search yielded "*519*" articles. This search was performed using a structured query with carefully selected keywords to target relevant articles. The search keywords are presented in Listing 1.

Listing 1. Search keywords

```
((uncertainty OR uncertain OR vague OR inconsistent OR imprecise) AND (Bayesian
Network OR rough sets OR Dempster-Shafer OR fuzzy logic OR possibility OR
Possibilistic OR probabilistic OR  paraconsistent OR fuzzy sets OR probability)
 AND (modeling OR reasoning OR modelling OR representation OR handling) AND
ontology OR ontologies) AND (2024 OR 2023 OR 2022 OR 2021 OR 2020 OR 2019 OR
 2018 OR 2017 OR 2016 OR 2015 OR 2014 OR 2013 OR 2012 OR 2011 OR 2010) NOT
("IoT" OR "Internet of things" OR "Temporal" OR "data fusion" OR "information
fusion" OR "evolution reasoning" OR "ontology matching" OR "Business Model" OR
 "ontology alignment" OR "ontologies alignment" OR "mining"))
```

The search terms were within the subject, title, and abstract of the document to ensure relevancy. These studies must be published as book chapters, journal articles, or reference papers. The "NOT" clause excludes articles on topics such as IoT, fusion, alignment, and temporal reasoning. These areas often involve uncertainty related to data processing, integration, or temporal aspects rather than uncertainty inherent to domain knowledge within ontologies. Our review focuses specifically on uncertainty that arises from the domain conceptualization itself, rather than uncertainties stemming from ontological alignment or system integration tasks.

These articles were then subjected to inclusion and exclusion criteria explicitly explained in subsection 3.3. We removed *167* duplicate papers and excluded *68* more based on our assessment of their abstracts, which left us with a selection of *284* papers. In addition, we carried out a comprehensive literature review using the snowballing technique, which involved examining referenced papers published between 2010 and 2024 to form further layers of literature. This process added *278* papers to our initial set, resulting in a total of *562* reviewed papers. After reviewing the papers, we end up with a total of *117* modelling studies. The paper selection process is depicted in Figure 2.

3.3 Inclusion and Exclusion Criteria

The inclusion criteria for this review focus on selecting articles that address uncertainty modelling through a variety of formal formalisms. Specifically, we include studies that utilize possibility theory, DST, rough set theory, paraconsistent

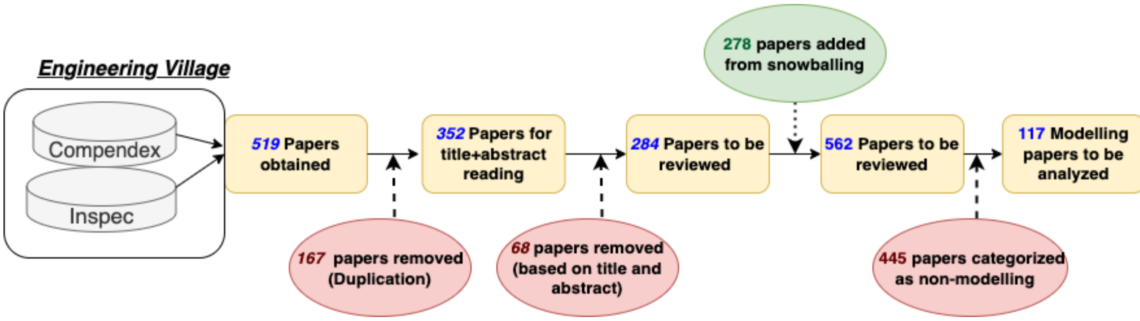


Fig. 2. Paper Selection Process

logic, probability theory, fuzzy set theory, or any combination of these frameworks. We strictly review papers that match our keyword search, as provided previously in subsection 3.2, or those identified through a snowballing process.

We apply several exclusion criteria to ensure the relevance and quality of the papers included in the review. First, we eliminate duplicate articles and limit the scope of the papers written in English. We excluded marginal articles that, while focusing on ontologies and related aspects, do not address uncertainty modelling within domain ontologies. We also exclude papers lacking sufficient technical details on the uncertainty modelling aspect, as well as older versions of papers for which more comprehensive and updated versions exist. Finally, articles with restricted access, where no full version is available online, are also excluded from our review.

3.4 Data Collection

The papers under review are categorized into nine general categories based on the formalism used to handle uncertainty. The categories are the following: *possibility theory, DST, rough set theory, paraconsistent logic, probability theory, fuzzy set theory, hybrid, other, and review.* In Table 3, we present a clear classification of the papers and the total number of papers for each formalism. The first group is derived from the initial phase of our research keywords. The second group is obtained through snowballing of the first group. The third group results from a further round of snowballing, this time applied to the second group.

Table 3. Total Numbers of Reviewed Papers Categorized by Formalism

Formalism	Initial phase	First snowballing	Second snowballing	Total
Possibilitic	16	2	3	21
DST	14	8	4	26
Rough set	8	3	0	11
Paraconsistent	10	9	1	20
Probabilistic	130	56	34	220
Fuzzy	79	69	76	224
Hybrid	17	4	1	22
Other	4	5	0	9
Review	6	3	0	9
	284	159	119	**562**

After reviewing the papers, we categorized each classification from Table 3 into four main groups based on the scope of the papers. The first group, *modelling papers*, includes studies focused on modelling or both modelling and reasoning, further divided into two subclasses: DL-based and non-DL-based. The second group, *reasoning papers*, includes studies on reasoners, reasoning algorithms, and optimization techniques. The third group, *application papers*, directly addresses GQ2 by identifying application-oriented studies. The final group, *excluded papers*, encompasses those excluded based on the specified exclusion criterion. Tabulated representations of these categorizations, Table 13 to Table 20 in the Supplementary Materials, provide clear answers to GQ1 and GQ2 of our study. Specifically, the tables divide studies by formalisms and highlight applications-oriented works.

3.5 Data Analysis

We conducted a comprehensive analysis of the guiding questions across all selected papers, that is, those studies that focus on modelling or involve both modelling and reasoning, a total of 117 papers. The data for each paper are presented, analyzed, and tabulated, aligned with the guiding questions (GQs) provided in subsection 3.1. The analyzed data are:

- The formalisms employed to model uncertainty within the ontological frameworks (addressing GQ1);
- The orientation of the paper, whether foundational or application-oriented (addressing GQ2);
- Motivations for modelling uncertainty (addressing GQ3-A) and the domain area (addressing GQ3-B);
- Sources of uncertainty (addressing GQ4-A) and types of uncertainty (addressing GQ4-B);
- The location of uncertainty in the ontology based on the developed taxonomy (addressing GQ5);
- The approach applied to model uncertainty within the ontological framework (addressing GQ6);
- Decidable reasoning procedures (addressing GQ7);
- Involvement of Machine Learning or NLP approaches (addressing GQ8 and GQ9);
- Supporting languages or tools (addressing GQ10-A-B);
- Evaluation of the approaches (addressing GQ11).

4 RESULTS

In this section, we present the review results organized according to the guiding questions (GQ1-GQ11) provided earlier. Each aspect of uncertainty modelling is explored in a dedicated subsection, categorized based on the formalism used (GQ1). Additionally, all selected papers are foundational in orientation (GQ2). In subsection 4.1, we examine the motivations for modelling uncertainty in domain ontologies, along with the specific domain areas where it is applied (GQ3-A and GQ3-B). Next, subsection 4.2 presents the types of uncertainty being modelled and identifies their sources (GQ4-A and GQ4-B). Following this, subsection 4.3 identifies where the uncertainty is depicted within the ontological framework (GQ5). Next, subsection 4.4 specifies how these formalisms are applied in the ontology (GQ6), while subsection 4.5 highlights any reasoning tasks provided in the studies (GQ7). The role of Machine Learning and NLP is examined in subsection 4.6, which identifies cases where these techniques are used to support uncertainty modelling (GQ8 and GQ9). In subsection 4.7 we present the languages and tools that facilitate the modelling process (GQ10-A and GQ10-B). Finally, subsection 4.8 provides an overview of the evaluation methodologies used in the reviewed papers (GQ11). The Supplementary Materials include two tree-like structures: the first tree, Figure 3, classifies papers based on GQ4-A, GQ4-B, and GQ5, while the second, Figure 4, organizes them according to GQ7-GQ11.

4.1 Motivations and Domain Areas for Uncertainty Modelling in Ontology

The motivations for uncertainty modelling in domain ontologies, summarized in Table 4, fall into two broad categories: knowledge representation (43 papers) and reasoning (74 papers). Although these aspects often overlap, papers tend to emphasize one over the other. For analytical clarity, we classified them accordingly, while fully acknowledging that the two are frequently intertwined. *Knowledge representation* primarily focuses on enhancing the expressiveness of ontologies to better capture incomplete, ambiguous, imprecise, or inconsistent domain knowledge. *Reasoning* aims primarily to perform inference tasks in the presence of uncertainty, such as uncertain query answering or classification. We note that the separation is not rigid: many papers contribute to both aspects. However, for thematic synthesis, this classification helps to highlight the dominant focus of each work.

Table 4. Motivations for Uncertainty Modelling within Ontological Frameworks Categorized by Formalism

Formalism	Motivation	
	Reasoning	Knowledge representation
Possibilistic	[83, 84, 85, 86, 87, 88]	[89, 90]
DST	[91, 92, 93, 94, 95, 96]	
Rough set	[97, 98]	[99, 100]
Paraconsistent	[78, 101, 102, 103, 104, 105, 106, 107, 108]	[109]
Probabilistic	[110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124]	[125, 126, 127, 128, 129, 130, 131, 132]
Fuzzy	[133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164]	[165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191]
Hybrid	[192, 193]	[194, 195]
Other	[196, 197]	[198]

Table 5 outlines the domain areas where uncertainty modelling in domain ontologies has been applied, as identified in the reviewed papers. The applications cover a wide range of fields, including information retrieval, Intrusion Detection System (IDS), Decision Support System (DSS), Case-Based Reasoning (CBR), and the semantic web. The analyzed papers not cited in Table 5 present general-purpose applications, highlighting foundational modelling or reasoning approaches that can be adopted across multiple domains.

Table 5. Domain Area for Uncertainty Modelling within Ontological Frameworks Categorized by Formalism

Formalism	Domain Area											
	Semantic web	Data fusion	Medical domain	Geo-spatial	Robotics	Information retrieval	IDS	Ontology learning	CBR	Science	DSS	Gaming
Possibilistic	[84, 85, 88]		[89]	[90]		[83]						
DST	[92]	[93, 94]				[96]					[91]	
Rough set			[97]					[99]				
Paraconsistent	[78, 102, 103, 107]		[101]									
Probabilistic	[110, 114, 120, 123, 127, 129]		[113, 115, 121, 131]		[125, 126]	[116, 118, 119]		[130]			[132]	
Fuzzy	[134, 145, 146, 147, 151, 172, 180, 186, 188]		[136, 137, 144, 164]	[159]		[148, 171]	[143, 163, 167]	[152, 153, 165, 174, 175, 176, 182, 183, 185, 187]	[133]	[149, 156, 166]	[138, 162, 168, 169, 170, 179, 178]	[139, 140, 141, 142]
Hybrid	[193, 194, 195]											
Other		[197]										

4.2 Sources and Types of Uncertainty

In our review of selected papers, we identified a range of sources of uncertainty addressed across various applications. These sources include sensor limitations, conflicting information from different sources, measurement errors, and model limitations. On top of that, contextual factors contribute to uncertainty, as they often influence how information is interpreted. Lastly, some papers did not specify a particular source of uncertainty, so we identified these as undetermined. Some papers addressed multiple sources of uncertainty, resulting in their being cited more than once. This aspect is summarized and presented in Table 6.

Table 6. Sources of Uncertainty Categorized by Formalism

Formalism	Sources of Uncertainty					
	Sensors	Contradictory sources	Measurements errors	Model limitations	Context	Undetermined
Possibilistic	[85, 89]	[84, 85, 86]	[90]	[87]	[83, 88]	
DST	[91, 93]	[93, 94]	[93]		[92, 95, 96]	
Rough set			[100]	[97]	[98, 99]	
Paraconsistent		[78, 101, 103]		[104, 105, 106, 107, 108]	[102, 109]	
Probabilistic	[117, 125]	[110, 120]	[125]	[125]	[111, 112, 113, 114, 115, 116, 118, 121, 127, 128, 129, 130, 131]	[119, 122, 123, 124, 126, 132]
Fuzzy					[134, 133, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 149, 150, 152, 153, 154, 155, 158, 159, 160, 161, 164, 165, 166, 167, 168, 170, 171, 172, 173, 175, 176, 177, 178, 180, 181, 182, 183, 184, 185, 188, 189, 190, 191]	[148, 151, 156, 157, 162, 163, 169, 174, 179, 186, 187]
Hybrid					[193, 195]	[192, 194]
Other		[197]			[196, 198]	

The reviewed papers identified several distinct types of uncertainty, including incomplete information, as well as vague and imprecise information, which were often considered together. Inconsistent information is also commonly noted. Some studies address multiple types collectively. These are categorized as "Several types", while other studies do not specify a particular type of uncertainty; these are categorized as "Undetermined". This aspect is summarized and presented in Table 7.

Table 7. Types of Uncertainty Categorized by Formalism

Formalism	Types of Uncertainty				
	Incomplete information	Vagueness and Imprecision	Inconsistency	Several types	Undetermined
Possibilistic	[85]		[83, 84, 86, 88]	[87, 89, 90]	
DST	[92]		[95]	[91, 93, 94, 96]	
Rough set			[99]	[97, 98, 100]	
Paraconsistent			[78, 103, 104, 105, 106, 107, 108, 109]	[101, 102]	
Probabilistic				[110, 121]	[111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132]
Fuzzy		[133, 134, 135, 136, 137, 139, 140, 141, 142, 143, 144, 148, 149, 151, 155, 156, 157, 158, 159, 160, 161, 163, 165, 166, 167, 168, 169, 170, 171, 173, 174, 175, 176, 177, 178, 180, 181, 182, 183, 184, 185, 186, 188, 189, 190, 191]		[150, 154]	[138, 145, 146, 147, 152, 162, 153, 164, 172, 179, 187]
Hybrid		[194]		[192, 193, 195]	
Other		[196]		[197, 198]	

4.3 Location of Uncertainty within Ontology

In the reviewed papers, we found that uncertainty in domain ontologies might arise at several points, namely, in the relationship between concepts, the properties or attributes associated with concepts, the semantic ambiguity of concepts, the mapping of instances to concepts, and the mapping of instances to relationships. Some papers focus on a single location of uncertainty, while others address multiple areas. Consequently, we have categorized these papers as "Several locations". This aspect is summarized and presented in Table 8.

Table 8. Location of Uncertainty Categorized by Formalism

Formalism	Locations of Uncertainty					
	Relationship between concepts	Attribute of concept	Semantic ambiguity of concept	Mapping instance to concept	Mapping instances to relationship	Several locations
Possibilistic		[89]				[83, 84, 85, 86, 87, 88, 90]
DST			[91]	[92, 94]		[93, 95, 96]
Rough set		[98, 99]	[97]			[100]
Paraconsistent			[101, 109]			[78, 102, 103, 104, 105, 106, 107, 108]
Probabilistic	[113, 115, 121, 126]	[111, 124, 129]	[130]			[110, 112, 114, 116, 117, 118, 119, 120, 122, 123, 125, 127, 128, 131, 132]
Fuzzy	[188]	[136, 164, 165]		[134, 135, 138, 139, 140, 141, 142, 143, 144, 150, 162, 163]	[168, 169, 170, 171, 179]	[133, 137, 145, 146, 147, 148, 149, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 166, 167, 172, 173, 174, 175, 176, 177, 178, 180, 181, 182, 183, 184, 185, 186, 187, 189, 190, 191]
Hybrid						[192, 193, 194, 195]
Other				[196, 197, 198]		

4.4 Modelling Approach

In this subsection, we present the results of how different approaches are used to model uncertainty, categorized by the adopted formalism. Each sub-subsection is dedicated to an analysis of papers based on each formalism. We divide these approaches into two main classes: DL-based approaches, which represent the majority of the papers, and non-DL-based approaches. Specifically, for DL-based approaches, we examine whether the approach extends only the TBox, only the ABox, or the entire knowledge base, or if it operates at the language level by adding annotations to support uncertainty modelling. In addition, we analyze and present papers that adopt non-DL based approaches, exploring their contributions to uncertainty modelling.

4.4.1 Possibilistic Papers. In possibilistic DL-based approaches, uncertainty is addressed by incorporating necessity or possibility degrees into the axioms of the ontologies. This can involve ABox assertions, in [85], TBox axioms in [87], or both TBox and ABox axioms, demonstrated in [83, 84, 88]. The study of [86], goes further by incorporating a possibility distribution for each interpretation in addition to the axioms. The work of [90] operates at the language level by introducing annotations to Web Ontology Language 2 (OWL2) to enable support for possibilistic ontologies. An approach [89], not based on DL, assigns possibility degrees directly to ontology concepts.

4.4.2 Dempster-Shafer Papers. In DST papers, uncertainty is managed by incorporating belief values with various components of the ontology. In [92], the authors introduce a method to convert an ontology into a Terminological Decision Tree (TDT), where each node corresponds to a concept in the ontology and belief values are assigned to each node to capture uncertainty. Another study [95], based on DL, introduces belief values to both ABox and TBox axioms to form an extended knowledge base, while [91] adds annotation "hasMass" to Web Ontology Language (OWL) to support belief-based concepts. In addition, a non-specific DST-based ontology, which can be added as an upper ontology to OWL is developed in [93, 94], incorporating the *Uncertain_Concept* to facilitate DST-based reasoning. Lastly, paper [96] focuses on subjective DL-Lite by adding belief values to ABox assertions.

4.4.3 Rough Set Papers. In rough set DL-based papers, rough concepts are introduced by generating upper and lower approximations of the concepts. These approximations are derived based on the indiscernibility relation, which is defined by the values of the set of attributes characterizing the concept [97, 99, 100]. In [98], the author focuses on the language level by proposing r-OWL, an extension of OWL with added annotations designed to support the representation and reasoning of rough ontologies.

4.4.4 Paraconsistent Papers. Paraconsistent logics are commonly used to manage inconsistencies within DL-based ontologies, with different approaches leveraging various types of paraconsistent semantics. Belnap's four-valued paraconsistent logic, as applied in [78, 101], associates to each DL-concept or relation two subsets: one representing positive membership and the other representing negative membership. The truth values $\{T, F, B, U\}$ are used to capture the membership of instance to this concept/relation. Specifically, T (True) indicates that an instance belongs to the positive membership set, while F (False) indicates membership in the negative subset. B (Both) represents cases where the instance simultaneously belongs to both subsets. Finally, U (Unknown) represents cases where an instance does not belong to either subset. On the other hand, paper [109] introduces a dual interpretation structure, $\langle \Delta^{PI}, \cdot^{I+}, \cdot^{I-} \rangle$, designed to support Nelson's four-valued paraconsistent logic. In this framework, Δ^{PI} is a non-empty domain, while $\cdot^{I+}$ and $\cdot^{I-}$ are interpretation functions that assign two subsets, A^{I+} and A^{I-}, to each concept and relation within Δ^{PI}. These subsets represent the positive (A^{I+}) and negative (A^{I-}) interpretations of the concepts or relations. The framework also defines a paraconsistent negation as $(\neg C)^{I+} := C^{I-}$. Papers [104, 108] inherits the semantics of the quasi-classical paraconsistent logic, which defines two interpretations: weak and strong. The weak interpretation is a reformulation of Belnap's four-valued interpretation, and the strong interpretation redefines the interpretations of the disjunction of concepts and conjunction of concepts of the weak interpretation to validate some inferences rules. The DL fragment $\mathcal{SROIQ}$ is extended in [103] with Kleene's three-valued paraconsistent logic, introducing truth values $\{T, F, I\}$, where I represents indeterminate, while [107] presents a variant of Kleene's three-valued logic with $\{T, F, B\}$, where B denotes both true and false, extending the DL fragment $\mathcal{ALC}$ with paradoxical paraconsistent logic. From a different perspective, paper [102] studies a number of different paraconsistent semantics for $\mathcal{SROIQ}$ such as Belnap's, and Kleene's three-valued paraconsistent logics. Argumentation-based paraconsistent DL is presented in [105], while distance-based paraconsistent DL is introduced in [106].

4.4.5 Probabilistic Papers. Probabilistic approaches to uncertainty modelling in domain ontologies can be broadly divided into two main directions: extending domain ontologies and their languages with probabilistic axioms and leveraging PGMs. The first direction focuses on embedding probabilities, either epistemic or statistical, into ontological structures. For instance, statistical probabilities have been added to both TBox axioms and ABox assertions of DL-based ontologies, as demonstrated in [122], while epistemic probabilities have been applied to ABox assertions in [125, 131], entire knowledge base axioms in [114, 119], and relational axioms in [126]. Probabilistic knowledge bases further extend this direction by combining a classical TBox with probability assertions based on discrete or continuous distributions [117, 121], and naïve Bayes classifiers have been employed to calculate probabilities for all relations in the ontology in [115].

These classifiers are probabilistic models based on Bayes' theorem, assuming strong independence among features [199]. Another approach [118], does not specify whether the probabilities used are statistical or epistemic, but simply attaches probabilities to ABox assertions. In contrast, log-linear description logic adopts a different perspective by representing an ontology as a pair consisting of a deterministic CBox, C^D, and an uncertain CBox, C^U. The uncertain CBox is defined as $C^U = \{(c, w_c) \mid$ c is an EL++ axiom and w_c a real valued weight for c$\}$ [124].

Distribution semantics [200] extends logical systems by assigning probabilities to axioms, enabling reasoning over uncertainty in ontologies. It forms the basis of the *DISPONTE* probabilistic framework, which incorporates probabilistic axioms into domain ontologies for uncertain knowledge reasoning [120, 123].

Further, probabilistic extensions to ontology languages enhance uncertainty modelling. For example, annotated OWL adds probabilistic annotations to facilitate reasoning [111, 116], and PR-OWL introduces constructs like *"definesUncertaintyOf"* to explicitly define probabilistic concepts [129]. In addition, the paper [132] extends PR-OWL by introducing the PR-OWL decision language, an advancement achieved by integrating the Multi-Entity Decision Graph (MEDG) framework into the existing structure. This integration enables effective decision-making within domain ontologies, even in the presence of domain uncertainty.

The second direction employs PGMss such as BNs and Markov Logic Network (MLN)s to represent probabilistic ontologies. BNs are integrated into ontology frameworks to provide a probabilistic axiomatic foundation [112, 113, 127, 128, 130], while MLNs extend ontology languages such as *Datalog+/–* with probabilistic semantics [110].

4.4.6 Fuzzy Papers. Fuzzy ontological approaches enhance traditional ontology frameworks by incorporating membership degrees into instances, relationships, concepts, or combinations of these elements. Several types of fuzzy logics, including T-norms-based, type-1, and interval type-2 fuzzy logics, have been employed to support these enhancements. These approaches have enabled representations of uncertainty in ontological models, as detailed below.

T-norms-based fuzzy logics have been extensively used to extend DL frameworks. The paper [160] introduces the *fALCN-DL*, which incorporates fuzzy concepts, fuzzy roles, and fuzzy interpretations. The work of [154] presents a tractable fuzzy extension of *EL++* using Gödel T-norms-based logic. Other contribution includes the extension of *SROIQ-DL* with a combination of Gödel and Zadeh logics [173]. The development of fuzzy *ALCH-DL* is introduced in [181], and the introduction of fuzzy *ALC-DL* is found in [189, 191].

Type-1 fuzzy logic is frequently used to extend several DLs by integrating fuzzy membership functions into various ontology components. The works of [150, 162] explore mapping instances to concepts, while [179] focuses on relationships between instances of concepts. In addition, studies of [133, 148, 177] extend the entire ontological structure by incorporating fuzzy membership functions, enabling diverse applications across domains. At the language level, the study of [190] extends OWL2 by integrating fuzzy annotations. This includes fuzzy concepts, fuzzy nominal, fuzzy relations, fuzzy axioms, fuzzy modifiers, and fuzzy datatype annotations.

Interval type-2 fuzzy logic has also been adopted to extend several DLs. The development of *IFALCN*, an extension of *ALCN* is introduced in [184]. In addition, fuzzy *ALC* is found in [156], and *Gf-EL++*, a generalized fuzzy extension of *EL++* is proposed in [186]. A comprehensive paper that defines type-0, type-1, interval and type-2 fuzzy DLs is [161].

Other approaches include methodologies for fuzzy ontology construction, as proposed in [165, 166, 176], and the introduction of Fuzz-Onto, a meta-ontology for representing fuzzy concepts, fuzzy relationships, and fuzzy properties [149]. In addition, techniques that integrate fuzzy logic with Formal Concept Analysis (FCA) to create Fuzzy Formal Concept Analysis (FFCA) for ontology construction are discussed in [174, 175]. On the other hand, approaches that combine fuzzy-DL with fuzzy Horn logic rules to construct hybrid knowledge bases are presented in [151, 158, 188].

Some studies address the priority or importance of properties in determining membership degrees highlighting the role of property importance and priority in membership calculations [134, 135]. On the other hand, paper [180] defines and analyzes fuzzy comparison cuts, a method for representing fuzzy knowledge about comparisons between membership degrees by annotating concepts and roles with comparison expressions. The work of [159] introduces intuitionistic fuzzy sets to represent both membership and non-membership functions in a spatial domain ontology. The study of [157] extends the well-known fuzzy *f-SHIN* DL, creating the *L-SHIN* framework by using certainty lattices. Paper [155] enhances *ALC* DL by incorporating multi-valued semantics based on certainty lattices.

Techniques for translating structures like XML, Unified Modelling Language (UML), Fuzzy Enhanced Entity Relationship (FEER), and Fuzzy Object-Oriented Database (FOOD) into fuzzy ontologies through predefined translation rules have been proposed in the studies [145, 146, 147, 152, 153, 172, 178, 182, 183, 185, 187].

One group of authors has contributed papers on constructing interval type-2 fuzzy-DL for use in various domains, namely IDS and DSS [167, 168]. They have also explored various applications of fuzzy-DL in different contexts, as seen in [169, 170, 171]. Another group of researchers has focused on constructing fuzzy ontologies and inference systems for domains such as gaming, malware detections, scheduling systems, and diet recommendations [139, 141]. In some of their publications, they incorporate type-2 fuzzy logic and a fuzzy markup language, as demonstrated in [136, 137, 138, 140, 142, 143, 144, 163, 164].

4.4.7 Hybrid Papers. Hybrid approaches leverage combinations of two formalisms to address uncertainty in domain ontologies. In [194], a fuzzy *ALC(D)* is extended by assigning belief values to each fuzzy assertion in the ABox, with this extension also applying to non-fuzzy assertions. In a different approach, described in [193], a new ontology language Belief-Augmented OWL (BOWL), which integrates OWL-DL with the Belief-Augmented Frames (BAF) [201]. In this model, each fact about individuals (i.e., each ABox assertion), which may be fuzzy, is augmented with Belief-Augmented Frames (BAF), a pair comprising belief and disbelief measures. A fuzzy rough extension of the *SROIQ(D)* is proposed in [195], where fuzzy-DL relations are augmented with rough approximation sets, which include tight and loose lower and upper approximations. Similarly, paper [192] develops upper and lower approximations of a concept based on two similarity relations defined over paraconsistent sets, rather than fuzzy sets.

4.4.8 Other Papers. Other approaches include those that do not adopt any of the primary formalisms identified in our search keywords, namely, probability theory, possibility theory, DST, rough set theory, fuzzy logic, and paraconsistent logic. We identified three additional formalisms used to manage uncertainty within domain ontologies: truth gap theory, Defeasible Logic Programs (DeLP), and soft set theory. In the following, we provide a brief introduction to each formalism, followed by a discussion of the approaches that adopt these formalisms.

A truth gap refers to a convex interval or ordered sequences of values within a specific domain where the satisfaction of a membership assertion to a vague concept cannot be definitely determined. In other words, when the value of an attribute (e.g., price, size, etc.) falls within this interval, the membership of the concept is indeterminate: it is neither definitely true nor definitively false [202]. The work of [196] introduces a framework for addressing vagueness in domain ontologies by leveraging truth gaps to represent indeterminate concepts within OWL2. In this framework, assigning instances to vague concepts based on specific attributes can result in three possible cases: tt for "definitely true", ff for "definitely false", and numerical ranges (truth gaps) when the membership of an individual a to concept C is indeterminate.

It is worth noting that DeLP represents a formalism for reasoning that incorporates defeasible reasoning, where conclusions can be retracted or modified in light of new evidence or exceptions [203]. Unlike classical logic, where conclusions are final, DeLP supports non-monotonic reasoning, which means that adding new information can alter or invalidate previous conclusions. In the work of [197], the δ-ontology framework extends DL ontologies by interpreting them as DeLP. It partitions the DL TBox T into two distinct sets: a strict terminology (T_S), containing non-defeasible axioms, and a defeasible terminology (T_D), containing defeasible axioms. The ontology is represented as a tuple $\Sigma = (T_S, T_D, A)$, where A is the ABox. The framework uses two functions, T_Π and T_Δ, to translate strict and defeasible rules in DeLP, receptively. The partitioning allows for reasoning with potentially inconsistent ontologies by maintaining strict rules that must hold universally and defeasible rules that can be retracted or modified based on new information.

Soft set theory is one of the parameterized theories introduced to handle uncertainty and address challenges faced by other uncertainty management frameworks, such as the difficulty in defining membership functions in fuzzy set theory [204]. Mathematically, given a finite sample space U, and a set of parameters E, a soft set (over U) is defined as a pair (F, E), where F maps each parameter $\epsilon \in E$ to a subset of U. In other words, a soft set can be viewed as a parameterized family of subsets of U. These sets $F(\epsilon)$ may overlap, intersect, be disjoint or even be empty, depending on the relationship between parameters and the sample space elements. In [198], the authors integrate soft set theory with DLs, where DL concepts act as parameters for the soft set. The approach leverages a DL interpretation $I = (\Delta_I, \cdot^I)$, which consists of a domain interpretation Δ_I and an interpretation function $\cdot^I$. The interpretation function maps each atomic concept or role to a subset of the domain Δ_I. A soft set (I, M) is defined as a pair, where $I = (\Delta_I, \cdot^I)$ is a model of DL knowledge base, and M is a set of DL concepts. The mapping $I \mid M : M \Rightarrow P(\Delta_I)$ assigns each concept C in M to a subset of Δ_I, representing the set of C-elements (or approximate individuals) in the soft set. For instance, a soft set might include approximations like (Expensive Hotels, $\{h_1, h_4\}$), where the first part of the pair is a concept (i.e., Expensive Hotels), and the second part is a set of approximate values, "instances", (i.e., $\{h_1, h_4\}$) corresponding to hotels that are considered "approximately" expensive.

4.5 Reasoning

As presented previously in the graphs of subsection 3.4, some of the papers focus on modelling and reasoning simultaneously. In particular, 81 papers out of 117 reviewed involve reasoning tasks. These reasoning tasks encompass a variety of tasks, including subsumption, satisfiability, instance checking, entailment, and checking knowledge base inconsistency. Some papers address more than one reasoning task, and as a result, they are cited multiple times. Others cover all the reasoning tasks outlined. This information is summarized and tabulated in Table 9.

4.6 Machine Learning or NLP Approaches

This subsection focuses specifically on papers that employ Machine Learning or NLP techniques as part of their methodology to handle or support uncertainty modelling within the ontological framework. While Machine Learning or NLP are commonly used in ontology learning and construction [205, 206], our aim here is not to address such uses, but to highlight cases where these techniques directly contribute to managing uncertainty. Among the 117 reviewed papers, we find that only a few papers utilized either Machine Learning or NLP-based approaches to support uncertainty modelling within ontological framework papers. Paper [91] applies DST as the formalism and uses Hidden Markov Model (HMM), a Machine Learning approach, to compare it with their proposed method. Similarly, paper [92] integrates DST with Terminological Random Forest (TRF), an ensemble of TDTs, as a Machine Learning approach. This enhanced DST was applied to infer class membership for instances. Paper [115] presents a probabilistic study, leveraging a Naïve Bayes classifier, another Machine Learning approach, to calculate probability values. On the other hand, paper [176]

Table 9. Reasoning Tasks Categorized by Formalism

Formalism	Reasoning task						
	Subsumption	Satisfiability	Consistency	Instance checking	Entailment	Query answering	All reasoning tasks
Possibilistic	[86, 90]	[88]	[83, 85]	[90]			
DST			[91, 95]	[92, 93, 96]		[96]	
Rough set	[98]					[97]	
Paraconsistent			[104, 105]		[107, 108]	[102]	
Probabilistic	[113]		[118, 122]	[116]	[121, 128, 132]	[110, 115, 117, 118, 119, 120, 123, 124]	[111, 112, 114, 125, 127, 129, 131]
Fuzzy	[186]	[155]	[180, 184]	[150]	[158, 188]	[148, 158]	[133, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 152, 153, 154, 156, 157, 159, 160, 161, 162, 163, 164, 173, 178, 181, 182, 183, 187, 189, 191]
Hybrid	[195]	[195]		[193]		[192]	
Other				[196, 197]			[198]

adopts a fuzzy approach that relies on similarity measures, an NLP method, to identify fuzzy similar relations, where the final similarity degree derived from the given formula was encoded as the degree of membership.

4.7 Languages and Tools

In this survey, we identified several languages used to support the modelling and reasoning of uncertainty in ontological frameworks in the 117 reviewed papers. These include both ontology representation languages (such as OWL, OWL2 variants, and UML) and rule-based languages (Fuzzy Markup Language (FML) and SWRL). Representation languages are employed to formally define concepts and relationships within ontologies, providing the logical foundation for uncertain knowledge modelling. Rule-based languages, on the other hand, are used to express inference rules or fuzzy mechanisms that operate over these ontologies. Table 10 summarizes the use of these languages under each uncertainty formalism. Some papers utilized multiple languages, leading to their inclusion in the analysis more than once. Conversely, papers that did not specify a language were excluded from Table 10. General-purpose implementation, programming languages, or query languages such as Java, C++, Python, Structured Query Language (SQL), and XML are excluded from the table, as they are not primarily designed for ontology modelling or rule-based reasoning.

Table 10. Ontology Representation and Rule Languages Used for Uncertainty Modelling, Categorized by Formalism

Formalism	Language						
	OWL	OWL2	OWL2-QL	OWL2-EL	OWL2-RL	UML	FML	SWRL
Possibilistic	[88]	[90]	[83]	[86]				
DST	[91, 93]	[94]					[93]	
Rough set	[97]	[98]						
Paraconsistent	[78, 108]	[102, 101, 103, 106]						
Probabilistic	[111, 115, 116, 119, 129, 130, 132]	[121]	[118]	[124]				
Fuzzy	[143, 145, 146, 147, 149, 150, 152, 153, 155, 156, 162, 167, 168, 169, 170, 171, 172, 174, 175, 176, 177, 179, 182, 183, 185, 187, 188]	[161, 173, 190]			[151]	[165, 176, 178]	[138, 140, 141, 143, 164]	[162, 175]
Hybrid	[193, 194]	[150, 195]						
Other		[196]						

Throughout the 117 papers analyzed, a wide range of tools were utilized to support uncertainty modelling in ontological frameworks. These included the use of algorithms implemented through the OWL API, alongside various existing DL-based reasoners, such as Pellet, Fact++, and Hermit reasoners. In addition, *DeLorean*, a fuzzy description logic reasoner, and *Fuzzy-DL*, another fuzzy reasoner, were utilized in several papers. *Pronto*, a probabilistic ontology reasoner, was also applied. Table 11 highlights the extensive range of existing tools leveraged to address uncertainty in ontological systems.

Table 11. Tools Used for Uncertainty Modelling in Ontologies Categorized by Formalism

Formalism	Tool						
	OWL API	Existing DL reasoners	FuzzyDL	DeLorean	Pronto	Protégé	Jena framework
Possibilistic	[88]	[88]				[90]	
DST	[91]	[91]					[93]
Rough set		[97, 98]				[97, 98]	
Paraconsistent	[78, 108]	[78, 101, 102, 103, 107, 108, 109]				[101]	
Probabilistic		[120]			[121]		[129]
Fuzzy	[151, 179, 190]	[178]	[167, 168, 188, 190]	[153, 169, 173, 182, 183, 190]		[143, 156, 162, 167, 168, 170, 171, 174, 190]	
Hybrid			[195]	[195]			
Other		[198]					

In addition to the tools summarized in Table 11, many studies introduced specialized tools or prototypes. Within the possibilistic category, [90] presents a custom possibilistic reasoner, while [88] develops *PossDL*, a possibilistic DL reasoner supporting inconsistency handling. In the paraconsistent category, the work of [108] developed *PROSE*, a paraconsistent reasoner. For probabilistic approaches, the study of [115] proposes an optimized Naive Bayes classifier using AdaBoost; [124] develops *ELOG*, a reasoner for EL++ ontologies. Paper [130] proposes *PrOntoLearn* system for probabilistic ontology learning; and [132] develops MEDG. The *BUNDLE* reasoner in [120] supports *DISPONTE* ontologies by integrating *PELLET* for OWL reasoning and *ProbLog* [207] for probabilistic *Prolog* reasoning. The latter is a well-known logic programming language that is commonly used for implementing reasoning systems, often utilized in rule-based reasoning frameworks and ontology-based systems [208]. *BORN*, a Bayesian DL reasoner, is used by [113], building on earlier work [209]. In the fuzzy category, a variety of tools were introduced. These include a custom fuzzy-OWL reasoner [149], the DL-MEDIA prototype [148], and automated ontology construction tools such as *FXML2FOnto* [182] (from fuzzy XML), *FEER2FOnto* in [152, 172] (from FEER models), *FRDB2Onto* in [153] (from Fuzzy Relational Database (FRDB) in MySQL), *FRDB2DL* in [145] (constructing f-ALCNI KBs), and *FOOD2FOWL* (from FOOD models and database instances). The *eHealthcare* application [162] classifies individuals based on fuzzy rules. Finally, in the hybrid category, the work of [193] implements a reasoner for BOWL using Constraint Logic Programming (CLP) paradigm.

4.8 Evaluation of the Approaches

In the selected papers, two types of evaluations might be found: use case demonstrations (i.e., a simplified, hypothetical example used to demonstrate the approach) and real-world experiments (i.e., a practical test conducted in a real environment to evaluate performance and effectiveness). In Table 12, we provide a comprehensive list of all papers that include an evaluation of their proposed modelling approaches, categorized according to the formalisms employed. Any paper not included in Table 12 does not present an evaluation.

Table 12. Evaluation Types of the Proposed Approaches Categorized by Formalism

Formalism	Evaluation Type	
	Use Case	Experiment
Possibilistic	[87, 90]	[88, 89]
DST	—	[91, 92, 95]
Rough set	[98, 99]	[97]
Paraconsistent	—	[78, 101, 105, 106, 108]
Probabilistic	[132]	[111, 115, 116, 119, 120, 124, 125, 126, 130, 177]
Fuzzy	[134, 135, 152, 153, 168, 169, 189]	[133, 136, 137, 138, 139, 141, 142, 143, 144, 145, 147, 148, 149, 150, 151, 156, 162, 163, 164, 166, 167, 171, 175, 176, 178, 179, 182, 185, 190]
Hybrid	[195]	[193, 194]
Other	[197]	—

5 RELATED WORK

In this section, we provide an overview of previous reviews conducted between 2008 and 2024 that focus on uncertainty modelling within domain ontologies. These papers are categorized based on the discussed formalism to uncertainty modelling in ontologies, specifically into fuzzy and probabilistic reviews. Notably, only one review paper covers both formalisms. It is noteworthy that fuzzy ontology reviews are more numerous compared to probabilistic ontology reviews. In the following, we provide a summary of these reviews.

The first category focuses on fuzzy-DL reviews. The literature on fuzzy-DL presents several review and survey papers that examine extensions and applications of fuzzy-DL for managing vagueness. The works of [30, 31, 41] provide descriptive overviews summarizing key fuzzy modelling approaches, applications, and reasoning methods, along with a comprehensive list of fuzzy-DL reasoners. Review [210] explores fuzzy extensions of OWL and their equivalent fuzzy-DL, discussing semantics derived from OWL's syntactic constructs and presenting a serialization method for abstract syntax into Resource Description Framework (RDF)/XML. The paper also introduces an approach for reducing fuzzy OWL ontology entailment to fuzzy-DL satisfiability. Similarly, [15] focuses on a prototypical fuzzy-DL based on classical *ALC*, discussing various fuzzy semantics such as Zadeh semantics, Gödel, Łukasiewicz, and product T-norms, and their impact on reasoning complexity and open challenges.

Another category focuses on fuzzy ontologies and is not restricted to DL-based literature. The paper [211] provides a concise review highlighting two main approaches to fuzzy ontologies: one modifies existing ontology languages to incorporate fuzzy representation, while the other develops dedicated fuzzy extensions. Another work [212] offers a brief subjective review exploring fuzzy ontology generation frameworks, tools, languages, applications, aggregation operators, and plug-ins, along with recommendations for new tools and languages to automate fuzzy ontology development and reasoning. The review of [16] presents a comprehensive review focusing on two perspectives: the extensions of classical ontologies with fuzzy logic, addressing both representation and reasoning, and the challenges associated with fuzzy ontologies: mapping, integration, storage, and applications. Review [17] provides an extensive review that discusses fuzzy ontology from multiple aspects: definitions, applications, representations, development approaches, and evaluation approaches, comparing various fuzzification methods and highlighting shortcomings in existing approaches. Paper [213] offers a brief review defining the concepts of vagueness (fuzziness), and uncertainty, exploring the mathematical foundations of fuzzy logic and fuzzy set, and discussing the structure and representation tools in fuzzy ontologies. On the other hand, paper [214] conducts a Systematic Mapping Study (SMS) exploring the contributions of fuzzy logic and ontology in the information systems, providing a general overview of relevant topics, future directions, and countries contributing to this field. Lastly, the work of [215] provides a detailed subjective review discussing type-2 fuzzy logic and its shortcomings in the context of ontology. This paper also offers a comprehensive list of type-2 fuzzy logic applications within the semantic web. Paper [216] explores the current advancements in representing and reasoning with fuzzy knowledge within RDF, OWL2 family, and rule languages. It further demonstrates how these frameworks can be extended to encompass annotation domains, enabling support for temporal and provenance information.

A different category covers Bayesian ontologies, focusing on the integration of probabilistic and logical reasoning. Paper [19] conducts a systematic literature review addressing three key research questions: the motivations, factors, and techniques for combining logical and probabilistic reasoning in knowledge bases, particularly using BN. The paper examines the foundational principles of integrating probabilistic with logical reasoning in ontologies and highlights prominent techniques and approaches. The work of [18] provides a brief comparative review discussing several approaches that merge ontologies with Bayesian Networks, analyzing their limitations, strengths, and primary purposes. Lastly, paper [14] conducts a comparative study of four Bayesian-based ontology approaches, concluding with a matrix that summarizes the results. Their findings suggest that BayesOWL is the most effective approach based on the provided metrics: complexity, accuracy, ease of implementation, availability of reasoning mechanisms, and supporting tools.

A final category addresses general reviews on managing uncertainty in expressive DL and ontology languages that combine more than one formalism. We found only one review [20], which provides a descriptive overview of the main approaches for managing uncertainty and vagueness in DL. The review highlights available methods for probabilistic, possibilistic, and fuzzy ontologies. Although this review was published in 2008 and falls outside the scope period of our review, it is one of the unique reviews in this domain, and we include it in our analysis.

6 DISCUSSION

This systematic review maps the landscape of uncertainty modelling in domain ontologies by analyzing 117 research papers through multidimensional aspects, focusing on a set of guiding questions. This review presents existing works and reveals emerging trends and critical gaps.

A central finding drawn from GQ3-A is that reasoning is the main motivation for integrating uncertainty into ontologies. As shown in Table 4, 74 of the 117 analyzed papers emphasize reasoning over knowledge representation, highlighting a strong focus on enabling uncertain inference. GQ3-B shows that uncertainty-aware ontologies span diverse application domains such as the semantic web, medical informatics, and robotics, as shown in Table 5. Fuzzy logic and probabilistic-based approaches dominate the field due to their wide applicability across diverse domains. Additional application-driven studies are listed in the Supplementary Materials in the *Application* category of each table.

From (GQ4-A, Table 6) we observe that context is the most frequently cited source of uncertainty in fuzzy-based studies, and (GQ4-B, Table 7) point out that vagueness and imprecision are predominantly modelled with fuzzy logic, which is consistent with the role of fuzzy logic in modelling domain-dependent and context-based concepts. Contradictory information is a dominant source of uncertainty in works employing possibilistic logic, DST, and paraconsistent logic, reflecting their intended use in reasoning with inconsistent information. Notably, many papers address multiple types of uncertainty, reinforcing the observation that uncertainty in ontologies is inherently multi-faceted. Regarding the locations of uncertainty (GQ5, Table 8), most studies address several components of ontologies simultaneously, such as concept definitions, relations, and instance data, suggesting that uncertainty is typically distributed and interconnected within knowledge structures. These findings are reflected in the structure of the proposed taxonomy (Figure 1).

The analysis of reasoning tasks (GQ7, Table 9) reveals that many formalisms, including possibilistic, rough set, and hybrid-based approaches, support only a narrow range of reasoning capabilities, typically limited to one or two tasks. In contrast, fuzzy-based studies exhibit a dominant trend toward providing comprehensive reasoning systems that cover all major reasoning tasks. It is worth noting that recent developments in typicality reasoning within DLs have opened promising directions for advanced ontology-based reasoning. Recent work by [217] and earlier studies (e.g., [218, 219]) introduce probabilistic typicality axioms (e.g., $p :: T(C) \sqsubseteq D$) grounded in distribution semantics, allowing uncertainty to be captured within prototypical knowledge. The *PEAR* tool [220] supports reasoning in these enriched settings. In parallel, the works of [221, 222] extend typicality-based DLs with fuzzy logic. While these approaches represent important advancements in integrating prototypical reasoning with uncertainty handling in DLs, it is important to emphasize that typicality-based reasoning addresses a different class of problems than those targeted in our work. Specifically, typicality focuses on modelling default knowledge, capturing what is generally true for most instances of a concept, while allowing for exceptions. This is conceptually distinct from the types of classical uncertainty addressed in our survey, which include vagueness, incompleteness, and inconsistent information. Although similar formalisms, such as fuzzy logic and probability theory, may be used in both domains, they serve different semantic purposes: in typicality, they help characterize degrees of typicality, whereas in uncertainty modelling, they are employed to quantify missing, imprecise, or conflicting information in the domain knowledge.

Notably, Machine Learning and NLP techniques remain underutilized in the context of uncertainty modelling within ontology-based frameworks, featuring in only a small fraction of the surveyed literature (GQ8-9). This limited presence reflects a broader research gap: while formal uncertainty models are well-established, the potential benefit of combining them with Machine Learning and NLP methods is not yet fully realized.

In terms of language usage (GQ10-A, Table 10), uncertainty modelling is largely built upon semantic web standards, especially OWL and its profiles, reinforcing the foundational role of description logic-based standards in ontology-based systems. (GQ10-B, Table 11) also highlights the development of dedicated tools, such as FuzzyDL and DeLorean for fuzzy logic, Pronto for probabilistic ontologies, and PossDL for possibilistic reasoning. In contrast, paraconsistent approaches often reduce reasoning to classical DL fragments, enabling reuse of standard DL-based reasoners like Pellet, HermiT, or FaCT++. This difference highlights a divergence in tool development strategies: while some formalisms require bespoke reasoning engines due to their non-classical semantics, others can be reduced to the classical DL infrastructure, leveraging established reasoning support without the need for new tools.

Finally, evaluation practices assessed through (GQ11, Table 12) reveal that only 56% of modelling studies include any form of empirical validation. This points to a significant gap in assessing real-world performance and emphasizes the need for more rigorous evaluations to ensure the practical viability of uncertainty-aware ontology systems.

7 CONCLUSION

This review provides a systematic analysis of uncertainty modelling in domain ontologies, categorizing approaches primarily based on their modelling formalisms. A comprehensive taxonomy was developed to identify specific locations and types of uncertainty within ontological structures. In addition to this, we presented reasoning methods that complement these modelling approaches. The review also examined the tools and languages used for uncertainty modelling, highlighting the integration of Machine Learning and Natural Language Processing techniques to support these efforts. The findings establish a solid foundation for advancing research, highlighting opportunities to refine current methods and address gaps identified in the ontological systems. Future research could include a review that focuses on the reasoning aspects, including complexity, tractability, and decidability, to address computational challenges and improve the effectiveness of reasoning in uncertain ontology frameworks.

References

[1] Shaker El-Sappagh, Francesco Franda, Farman Ali, and Kyung-Sup Kwak. 2018. SNOMED CT standard ontology based on the ontology for general medical science. *BMC Med Inform Decis Mak*, 18, 1, 76. doi: https://doi.org/10.1186/s12911-018-0651-5.

[2] Shivani Jain, KR Seeja, and Rajni Jindal. 2021. A fuzzy ontology framework in information retrieval using semantic query expansion. *International Journal of Information Management Data Insights*, 1, 1, 100009. doi: https://doi.org/10.1016/j.jjimei.2021.100009.

[3] Lisheng He, Wenjia Joyce Zhao, and Sudeep Bhatia. 2022. An ontology of decision models. *Psychological Review*, 129, 1, 49. doi: 10.1037/rev0000231.

[4] RV Karthik and Sannasi Ganapathy. 2021. A fuzzy recommendation system for predicting the customers interests using sentiment analysis and ontology in e-commerce. *Applied Soft Computing*, 108, 107396. doi: https://doi.org/10.1016/j.asoc.2021.107396.

[5] 2018. *Ontology. Encyclopedia of Database Systems*. Springer New York, New York, NY, 2574–2576. doi: 10.1007/978-1-4614-8265-9_1318.

[6] M. Fareh. 2019. Modeling incomplete knowledge of semantic web using bayesian networks. *Applied Artificial Intelligence*, 33, 11, 1022–1034. doi: https://doi.org/10.1080/08839514.2019.1661578.

[7] Sanjay Kumar Anand and Suresh Kumar. 2022. Uncertainty analysis in ontology-based knowledge representation. *New Generation Computing*, 40, 1, 339–376. doi: https://doi.org/10.1007/s00354-022-00162-6.

[8] 2015. *Possibility theory and its applications: where do we stand? Springer Handbook of Computational Intelligence*. Springer Berlin Heidelberg, Berlin, Heidelberg, 31–60. doi: 10.1007/978-3-662-43505-2_3.

[9] Kari Sentz and Scott Ferson. 2002. *Combination of evidence in Dempster-Shafer theory*. Volume 4015. Sandia National Laboratories Albuquerque. doi: 10.2172/800792.

[10] 2020. *Overview of rough set theory*. Springer International Publishing, 7–60. doi: https://doi.org/10.1007/978-3-030-29566-0_2.

[11] 2002. *Paraconsistent logic. Handbook of Philosophical Logic*. Springer Netherlands, Dordrecht, 287–393. doi: 10.1007/978-94-017-0460-1_4.

[12] Radha G Laha and Vijay K Rohatgi. 2020. *Probability theory*. Courier Dover Publications.

[13] Hans-Jürgen Zimmermann. 2011. *Fuzzy set theory—and its applications*. Springer Science & Business Media. doi: https://doi.org/10.1007/978-94-010-0646-0.

[14] Foni Agus Setiawan, Wahyu Catur Wibowo, and Novita Br Ginting. 2015. Handling uncertainty in ontology construction based on bayesian approaches: a comparative study. In *International Conference on Soft Computing, Intelligence Systems, and Information Technology*. Springer, 234–246. doi: https://doi.org/10.1007/978-3-662-46742-8_22.

[15] Stefan Borgwardt and Rafael Peñaloza. 2017. Fuzzy description logics – a survey. In *Scalable Uncertainty Management*. Springer International Publishing, Cham, 31–45. doi: https://doi.org/10.1007/978-3-319-67582-4_3.

[16] Fu Zhang, Jingwei Cheng, and Zongmin Ma. 2016. A survey on fuzzy ontologies for the semantic web. *The Knowledge Engineering Review*, 31, 3, 278–321. doi: 10.1017/S0269888916000059.

[17] Zahra Riahi Samani and Mehrnoush Shamsfard. 2018. The state of the art in developing fuzzy ontologies- a survey. *arXiv preprint arXiv:1805.02290*. doi: 10.48550/arXiv.1805.02290.

[18] Asma S Larik and Sajjad Haider. 2010. Efforts to blend ontology with bayesian networks: an overview. In *2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*. Volume 2. IEEE. doi: 10.1109/ICACTE.2010.5579483.

[19] Foni Agus Setiawan, Eko K. Budiardjo, T. Basaruddin, and Siti Aminah. 2017. A systematic literature review on combining ontology with bayesian network to support logical and probabilistic reasoning. In *Proceedings of the 2017 International Conference on Software and e-Business*. Association for Computing Machinery, 1–12. doi: 10.1145/3178212.3178223.

[20] T. Lukasiewicz and U. Straccia. 2008. Managing uncertainty and vagueness in description logics for the semantic web. *Journal of Web Semantics*, 6, 4, 291–308. doi: 10.1007/978-3-540-85658-0_2.

[21] Sundeep Samson, James A. Reneke, and Margaret M. Wiecek. 2009. A review of different perspectives on uncertainty and risk and an alternative modeling paradigm. *Reliability Engineering & System Safety*, 94, 2, 558–567. issn: 0951-8320. doi: https://doi.org/10.1016/j.ress.2008.06.004.

[22] William L. Oberkampf, Sharon M. DeLand, Brian M. Rutherford, Kathleen V. Diegert, and Kenneth F. Alvin. 2002. Error and uncertainty in modeling and simulation. *Reliability Engineering & System Safety*, 75, 3, 333–357. issn: 0951-8320. doi: https://doi.org/10.1016/S0951-8320(01)00120-X.

[23] 2021. *Types of uncertainty. Mastering Uncertainty in Mechanical Engineering*. Springer International Publishing, Cham, 25–42. isbn: 978-3-030-78354-9. doi: 10.1007/978-3-030-78354-9_2.

[24] Bo Sun, Jerry Huang, Nicolas Christianson, Mohammad Hajiesmaili, Adam Wierman, and Raouf Boutaba. 2024. Online algorithms with uncertainty-quantified predictions. In *Proceedings of the 41st International Conference on Machine Learning* (ICML'24). JMLR.org, Vienna, Austria.

[25] 2018. *Experimentation, errors, and uncertainty. Experimentation, Validation, and Uncertainty Analysis for Engineers*. John Wiley & Sons, Ltd. Chapter 1, 1–32. isbn: 9781119417989. doi: https://doi.org/10.1002/9781119417989.ch1.

[26] 2005. *Introduction. Uncertainty and Information*. John Wiley & Sons, Ltd. Chapter 1, 1–25. isbn: 9780471755579. doi: https://doi.org/10.1002/0471755575.ch1.

[27] 1997. *Imperfect information: imprecision and uncertainty. Uncertainty Management in Information Systems: From Needs to Solutions*. Springer US, Boston, MA, 225–254. ISBN: 978-1-4615-6245-0. DOI: 10.1007/978-1-4615-6245-0_8.

[28] Mihai Florea, Anne-Laure Jousselme, Dominic Grenier, and Éloi Bossé. 2002. An unified approach to the fusion of imperfect data? *Proceedings of SPIE - The International Society for Optical Engineering*, 4731, (March 2002). DOI: 10.1117/12.458372.

[29] Umberto Straccia. 2013. *Foundations of Fuzzy Logic and Semantic Web Languages*. Chapman & Hall/CRC. DOI: https://doi.org/10.1201/b15460.

[30] Z. M. Ma, Fu Zhang, Hailong Wang, and Li Yan. 2013. An overview of fuzzy description logics for the semantic web. *The Knowledge Engineering Review*, 28, 1, 1–34. DOI: 10.1017/S0269888912000306.

[31] 2017. *From fuzzy to annotated semantic web languages. Reasoning Web: Logical Foundation of Knowledge Graph Construction and Query Answering*. Springer International Publishing, Cham, 203–240. DOI: 10.1007/978-3-319-49493-7_6.

[32] Lotfi Aliasker Zadeh. 1965. Fuzzy sets. *Information and control*, 8, 3, 338–353. DOI: 10.2307/2272014.

[33] Timothy J. Ross. 2010. *Fuzzy Logic with Engineering Applications*. (3rd edition). John Wiley and Sons. DOI: 10.1002/9781119994374.

[34] Abdul Quaiyum Ansari. 1998. The basics of fuzzy logic: a tutorial review. *Computer Education -Stafford- Computer Education Group*, 88, 5–9.

[35] Petr Hájek. 2006. What is mathematical fuzzy logic. *Fuzzy sets and systems*, 157, 5, 597–603. DOI: https://doi.org/10.1016/j.fss.2005.10.004.

[36] Lotfi A Zadeh. 1975. The concept of a linguistic variable and its application to approximate reasoning-III. *Information sciences*, 9, 1, 43–80. DOI: 10.1016/0020-0255(75)90036-5.

[37] 2017. *Type-2 fuzzy sets. Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions, 2nd Edition*. Springer International Publishing, Cham, 259–306. DOI: 10.1007/978-3-319-51370-6_6.

[38] M. M. Gupta and J. Qi. 1991. Theory of T-norms and fuzzy inference methods. *Fuzzy Sets and Systems*, 40, 3, 431–450. DOI: https://doi.org/10.1016/0165-0114(91)90171-L.

[39] M. Mas, M. Monserrat, J. Torrens, and E. Trillas. 2007. A survey on fuzzy implication functions. *IEEE Transactions on Fuzzy Systems*, 15, 6, 1107–1121. DOI: 10.1109/TFUZZ.2007.896304.

[40] Shi, Yun. 2009. *A Deep Study of Fuzzy Implications*. PhD thesis. Ghent University, V, 147. https://api.semanticscholar.org/CorpusID:125023124.

[41] 2015. *All about fuzzy description logics and applications. Reasoning Web. Web Logic Rules: Tutorial Lectures*. Springer International Publishing, Cham, 1–31. DOI: 10.1007/978-3-319-21768-0_1.

[42] Petr Hájek. 1998. *Metamathematics of Fuzzy Logic*. Kluwer Academic Publishers, Dordrecht, Boston and London. DOI: https://doi.org/10.1007/978-94-011-5300-3.

[43] Krassimir T Atanassov and G Gargov. 2017. *Intuitionistic fuzzy logics*. Springer. DOI: https://doi.org/10.1007/978-3-319-48953-7.

[44] M.J. Evans and J.S. Rosenthal. 2004. *Probability and Statistics: The Science of Uncertainty*. W. H. Freeman. https://books.google.com.sa/books?id=hkWK8kFzXWIC.

[45] Jon Barwise. 1977. An introduction to first-order logic. In *Studies in Logic and the Foundations of Mathematics*. Volume 90. Elsevier, 5–46. DOI: https://doi.org/10.1016/S0049-237X(08)71097-8.

[46] Brigitte Jaumard, Alexandre Fortin, Md. Istiaque Shahriar, and Razia Sultana. 2006. First order probabilistic logic. In *NAFIPS 2006 - 2006 Annual Meeting of the North American Fuzzy Information Processing Society*, 341–346. DOI: 10.1109/NAFIPS.2006.365433.

[47] 2012. *Herbrand logic. Introduction to Logic*. Springer International Publishing, Cham, 53–79. DOI: 10.1007/978-3-031-01798-8_6.

[48] Joseph Y. Halpern. 1990. An analysis of first-order logics of probability. *Artificial Intelligence*, 46, 3, 311–350. ISSN: 0004-3702. DOI: https://doi.org/10.1016/0004-3702(90)90019-V.

[49] Franz Pernkopf, Robert Peharz, and Sebastian Tschiatschek. 2014. Introduction to probabilistic graphical models. In *Academic Press Library in Signal Processing*. Volume 1. Elsevier, 989–1064. DOI: https://doi.org/10.1016/B978-0-12-396502-8.00018-8.

[50] Qin Zhang and Quanying Yao. 2018. Dynamic uncertain causality graph for knowledge representation and reasoning: utilization of statistical data and domain knowledge in complex cases. *IEEE Transactions on Neural Networks and Learning Systems*, 29, 5, 1637–1651. DOI: 10.1109/TNNLS.2017.2673243.

[51] Charles Sutton and Andrew McCallum. 2012. An introduction to conditional random fields. *Foundations and Trends in Machine Learning*, 4, 4, 267–373. ISSN: 1935-8237. DOI: 10.1561/2200000013.

[52] Matthew Richardson and Pedro Domingos. 2006. Markov logic networks. *Machine Learning*, 62, 1, 107–136. https://doi.org/10.1007/s10994-006-5833-1.

[53] Richard E Neapolitan et al. 2003. *Learning Bayesian Networks*. Volume 38. Pearson Prentice Hall Upper Saddle River, NJ. DOI: 10.1145/1327942.1327961.

[54] Patnaik Patnaikuni, R Shrinivasan, and Sachin R Gengaje. 2017. Survey of multi entity bayesian networks (MEBN) and its applications in probabilistic reasoning. *International Journal of Advanced Research in Computer Science*, 8, 5, 2425–2429. DOI: https://doi.org/10.26483/ijarcs.v8i5.3954.

[55] Daphne Koller and Avi Pfeffer. 1997. Object-oriented bayesian networks. In *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 302–313. https://api.semanticscholar.org/CorpusID:1939888.

[56] Jeffrey A Barnett. 1991. Calculating dempster-shafer plausibility. *IEEE transactions on pattern analysis and machine intelligence*, 13, 6, 599–602. DOI: 10.1109/34.87345.

[57] Ronald R. Yager. 1987. On the dempster-shafer framework and new combination rules. *Information Sciences*, 41, 2, 93–137. DOI: https://doi.org/10.1016/0020-0255(87)90007-7.

[58] Bree R Mathon, Metin M Ozbek, and George F Pinder. 2010. Dempster–shafer theory applied to uncertainty surrounding permeability. *Mathematical Geosciences*, 42, 3, 293–307. doi: https://doi.org/10.1007/s11004-009-9246-0.

[59] T. Inagaki. 1991. Interdependence between safety-control policy and multiple-sensor schemes via dempster–shafer theory. *IEEE Transactions on Reliability*, 40, 2, 182–188. doi: 10.1109/24.87125.

[60] Lianwen Zhang. 1994. Representation, independence, and combination of evidence in the dempster–shafer theory. In *Advances in the Dempster-Shafer theory of evidence*. Wiley. https://api.semanticscholar.org/CorpusID:125654612.

[61] Didier Dubois and Henri Prade. 1986. A set-theoretic view on belief functions: logical operations and approximations by fuzzy sets. *Int. J. General Systems*, 12, 193–226. doi: 10.1080/03081078608934937.

[62] Didier Dubois and Henri Prade. 1992. On the combination of evidence in various mathematical frameworks. In *Reliability data collection and analysis*. Springer, 213–241. doi: https://doi.org/10.1007/978-94-011-2438-6_13.

[63] Jean Gordon and Edward H Shortliffe. 1984. The dempster-shafer theory of evidence. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, 3, 832-838, 3–4. doi: https://doi.org/10.1007/978-3-7908-1811-6_6.

[64] Amira Essaid and Boutheina Ben Yaghlane. 2009. BeliefOWL: an evidential representation in OWL ontology. In *URSW*, 77–80. https://api.semanticscholar.org/CorpusID:14808783.

[65] 1994. *Possibilistic logic. Handbook of Logic in Artificial Intelligence and Logic Programming (Vol. 3): Nonmonotonic Reasoning and Uncertain Reasoning*. Oxford University Press, Inc., USA, 439–513.

[66] Didier Dubois and Henri Prade. 2014. Possibilistic logic — an overview. In *Computational Logic*. Handbook of the History of Logic. Volume 9. North-Holland, 283–342. doi: https://doi.org/10.1016/B978-0-444-51624-4.50007-1.

[67] Andrew Alola, Mustafa Tunay, and Uju Alola. 2013. Analysis of possibility theory for reasoning under uncertainty. *International Journal of Statistics and Probability*, 2, (February 2013), 12. doi: 10.5539/IJSP.V2N2P12.

[68] Zdzisław Pawlak. 2002. Rough set theory and its applications. *Journal of Telecommunications and Information Technology*, 7–10. doi: 10.26636/jtit.2002.140.

[69] Qinghua Zhang, Qin Xie, and Guoyin Wang. 2016. A survey on rough set theory and its applications. *CAAI Transactions on Intelligence Technology*, 1, 4, 323–333. doi: https://doi.org/10.1016/j.trit.2016.11.001.

[70] Melvin Fitting. 1991. Kleene's logic, generalized. *Journal of Logic and Computation*, 1, (December 1991). doi: 10.1093/logcom/1.6.797.

[71] Graham Priest. 1979. The logic of paradox. *Journal of Philosophical Logic*, 8, 1, 219–241. doi: 10.1007/bf00258428.

[72] 1977. *A useful four-valued logic. Modern Uses of Multiple-Valued Logic*. Springer Netherlands, Dordrecht, 5–37. doi: 10.1007/978-94-010-1161-7_2.

[73] A Hunter. 2000. Reasoning with contradictory information using quasi-classical logic. *Journal of Logic and Computation*, 10, 5, 677–703. doi: 10.1093/logcom/10.5.677.

[74] David Nelson. 1949. Constructible falsity. *The Journal of Symbolic Logic*, 14, 1, 16–26. http://www.jstor.org/stable/2268973.

[75] Philippe Besnard and Anthony Hunter. 2001. A logic-based theory of deductive arguments. *Artificial Intelligence*, 128, 1, 203–235. doi: https://doi.org/10.1016/S0004-3702(01)00071-6.

[76] Ofer Arieli. 2008. Distance-based paraconsistent logics. *International Journal of Approximate Reasoning*, 48, 3, 766–783. doi: https://doi.org/10.1016/j.ijar.2007.07.002.

[77] Yaroslav Petrukhin. 2016. Natural deduction for three-valued regular logics. *Logic and Logical Philosophy*, 26, (August 2016), 197–206. doi: 10.12775/LLP.2016.025.

[78] Frederick Maier, Yue Ma, and Pascal Hitzler. 2012. Paraconsistent OWL and related logics. *Semantic Web*, 4, (May 2012). doi: 10.3233/SW-2012-0066.

[79] Philippe Besnard and Anthony Hunter. 1995. Quasi-classical logic: non-trivializable classical reasoning from inconsistent information. In *European Conference on Symbolic and Quantitative Approaches to Reasoning and Uncertainty*. Springer. (July 1995), 44–51. doi: 10.1007/3-540-60112-0_6.

[80] 2008. *Semantical study of paraconsistent nelson's logic*. Springer, (January 2008), 131–157. doi: 10.1007/978-1-4020-6867-6_8.

[81] Norihiro Kamide and Heinrich Wansing. 2012. Proof theory of nelson's paraconsistent logic: a uniform perspective. *Theoretical Computer Science*, 415, 1–38. issn: 0304-3975. doi: https://doi.org/10.1016/j.tcs.2011.11.001.

[82] Staffs Keele et al. 2007. Guidelines for performing systematic literature reviews in software engineering. Technical report. Technical report, ver. 2.3 ebse technical report. ebse.

[83] Salem Benferhat and Zied Bouraoui. 2017. Min-based possibilistic DL-Lite. *Journal of Logic and Computation*, 27, 1, 261–297. doi: 10.1093/logcom/exv014.

[84] Salem Benferhat, Zied Bouraoui, and Karim Tabia. 2014. On the revision of prioritized DL-Lite knowledge bases. In *Scalable Uncertainty Management*. Springer International Publishing, Cham, 22–36. doi: https://doi.org/10.1007/978-3-319-11508-5_3.

[85] Khaoula Boutouhami, Salem Benferhat, Faiza Khellaf, and Farid Nouioua. 2017. Uncertain lightweight ontologies in a product-based possibility theory framework. *International Journal of Approximate Reasoning*, 88, 237–258. doi: https://doi.org/10.1016/j.ijar.2017.05.012.

[86] Rym Mohamed, Zied Loukil, and Zied Bouraoui. 2018. Qualitative-based possibilistic EL ontology. In *PRIMA 2018: Principles and Practice of Multi-Agent Systems*. Springer International Publishing, Cham, 552–559. doi: https://doi.org/10.1007/978-3-030-03098-8_41.

[87] ShengTao Sun. 2013. A novel semantic quantitative description method based on possibilistic logic. *Journal of Intelligent and Fuzzy Systems*, 25, 931–940. doi: 10.3233/IFS-120695.

[88] Guilin Qi, Qiu Ji, Jeff Z. Pan, and Jianfeng Du. 2011. Extending description logics with uncertainty reasoning in possibilistic logic. *International Journal of Intelligent Systems*, 26, 4, 353–381. doi: https://doi.org/10.1007/978-3-540-75256-1_72.

[89] Yosra Ben Salem, Rihab Idoudi, Karim Saheb Ettabaa, Kamel Hamrouni, and Basel Soleiman. 2018. Mammographie image based possibilistic ontological representation. In *2018 4th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, 1–6. doi: 10.1109/ATSIP.2018.8364475.

[90] Bal-Bourai Safia and Mokhtari Aicha. 2014. Poss-OWL 2: possibilistic extension of OWL 2 for an uncertain geographic ontology. *Procedia Computer Science*, 35, 407–416. doi: https://doi.org/10.1016/j.procs.2014.08.121.

[91] Mohd Halim Mohd Noor, Zoran Salcic, and Kevin I-Kai Wang. 2016. Enhancing ontological reasoning with uncertainty handling for activity recognition. *Knowledge-Based Systems*, 114, 47–60. doi: https://doi.org/10.1016/j.knosys.2016.09.028.

[92] Giuseppe Rizzo, Nicola Fanizzi, Claudia d'Amato, and Floriana Esposito. 2018. Approximate classification with web ontologies through evidential terminological trees and forests. *International Journal of Approximate Reasoning*, 92, 340–362. doi: https://doi.org/10.1016/j.ijar.2017.10.019.

[93] Bellenger Amandine and Gatepaille Sylvain. 2011. Uncertainty in ontologies: dempster-shafer theory for data fusion applications. *ArXiv*, abs/1106.3876. doi: https://doi.org/10.48550/arXiv.1106.3876.

[94] Amandine Bellenger, Xavier Lerouvreur, Habib Abdulrab, and Jean-Philippe Kotowicz. 2012. Semantic beliefs fusion. In *Advances on Computational Intelligence*. Springer Berlin Heidelberg, Berlin, Heidelberg, 491–500. doi: https://doi.org/10.1007/978-3-642-31709-5_50.

[95] Murat Sensoy, Achille Fokoue, Jeff Z. Pan, Timothy J. Norman, Yuqing Tang, Nir Oren, and Katia Sycara. 2013. Reasoning about uncertain information and conflict resolution through trust revision. In *Proceedings of the 2013 International Conference on Autonomous Agents and multi-agent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 837–844. https://api.semanticscholar.org/CorpusID:1109051.

[96] Jhonatan Garcia, Jeff Z. Pan, Achille Fokoue, Katia Sycara, Yuqing Tang, and Federico Cerutti. 2015. Handling uncertainty: an extension of DL-lite with subjective logic. English. *CEUR Workshop Proceedings*, 1350. issn: 1613-0073. https://api.semanticscholar.org/CorpusID:13893357.

[97] C. Maria Keet. 2010. Ontology engineering with rough concepts and instances. In *Proceedings of the 17th International Conference on Knowledge Engineering and Management by the Masses*. Springer-Verlag, Berlin, Heidelberg, 503–513. doi: https://doi.org/10.1007/978-3-642-16438-5_40.

[98] C. Maria Keet. 2011. Rough subsumption reasoning with rOWL. In *Proceedings of the South African Institute of Computer Scientists and Information Technologists Conference on Knowledge, Innovation and Leadership in a Diverse, Multidisciplinary Environment*. Association for Computing Machinery, New York, NY, USA, 133–140. doi: 10.1145/2072221.2072237.

[99] HongLi Chen and ShanGuo Lv. 2010. Study on ontology model based on rough set. In *2010 Third International Symposium on Intelligent Information Technology and Security Informatics*. IEEE, 105–108. doi: 10.1109/IITSI.2010.169.

[100] A.F. Kana and Babatunde Akinkunmi. 2016. Modeling uncertainty in ontologies using rough set. *International Journal of Intelligent Systems and Applications*, 8, (April 2016), 49–59. doi: 10.5815/ijisa.2016.04.06.

[101] Cristian Cocos, Fahim T. Imam, and Wendy MacCaull. 2012. Ontology merging and reasoning using paraconsistent logics. *Int. J. Knowl. Based Organ.*, 2, 35–51. https://api.semanticscholar.org/CorpusID:8365711.

[102] Linh Anh Nguyen. 2010. Paraconsistent and approximate semantics for the OWL 2 web ontology language. In *Rough Sets and Current Trends in Computing*. Springer Berlin Heidelberg, Berlin, Heidelberg, 710–720. doi: https://doi.org/10.1007/978-3-642-13529-3_75.

[103] 2012. *Paraconsistent reasoning for semantic web agents. Transactions on Computational Collective Intelligence VI*. Springer Berlin Heidelberg, Berlin, Heidelberg, 36–55. doi: 10.1007/978-3-642-29356-6_2.

[104] Zhang Xiaowang and Zuoquan Lin. 2012. Quasi-classical description logic. *Journal of Multiple-Valued Logic and Soft Computing*, 18, (January 2012). https://api.semanticscholar.org/CorpusID:12577174.

[105] Zhang Xiaowang and Zuoquan Lin. 2013. An argumentation framework for description logic ontology reasoning and management. *Journal of Intelligent Information Systems*, 40, (June 2013). doi: 10.1007/s10844-012-0230-7.

[106] Xiaowang Zhang, Kewen Wang, Zhe Wang, Yue Ma, Guilin Qi, and Zhiyong Feng. 2017. A distance-based framework for inconsistency-tolerant reasoning and inconsistency measurement in DL-Lite. *International Journal of Approximate Reasoning*, 89, 58–79. doi: https://doi.org/10.1016/j.ijar.2016.08.003.

[107] Xiaowang Zhang, Zuoquan Lin, and Kewen Wang. 2010. Towards a paradoxical description logic for the semantic web. In *Foundations of Information and Knowledge Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, 306–325. doi: 10.1007/978-3-642-11829-6_20.

[108] Xiaowang Zhang, Guohui Xiao, Zuoquan Lin, and Jan Van den Bussche. 2014. Inconsistency-tolerant reasoning with OWL DL. *International Journal of Approximate Reasoning*, 55, 2, 557–584. doi: https://doi.org/10.1016/j.ijar.2013.09.005.

[109] Norihiro Kamide. 2010. Paraconsistent description logics revisited. In *23rd International Workshop on Description Logics DL2010*, 197. https://api.semanticscholar.org/CorpusID:2253734.

[110] Georg Gottlob, Thomas Lukasiewicz, and Gerardo I. Simari. 2011. Conjunctive query answering in probabilistic datalog+/- ontologies. In *Web Reasoning and Rule Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, 77–92. doi: https://doi.org/10.1007/978-3-642-23580-1_7.

[111] Abdul-Wahid Mohammed, Yang Xu, and Ming Liu. 2016. Knowledge-oriented semantics modelling towards uncertainty reasoning. *SpringerPlus*, 5, 1, 706. doi: 10.1186/s40064-016-2331-1. https://doi.org/10.1186/s40064-016-2331-1.

[112] Leonard Botha, Thomas Meyer, and Rafael Peñaloza. 2019. A bayesian extension of the description logic ALC. In *European Conference on Logics in Artificial Intelligence*. Springer, 339–354. doi: https://doi.org/10.1007/978-3-030-19570-0_22.

[113] Rafael Peñaloza Nyssen İsmail İlkan Ceylan. 2017. The bayesian ontology language bel. *Journal of Automated Reasoning*, 58, 1, 67–95. https://doi.org/10.1007/s10817-016-9386-0.

[114] Eugene Santos and Jacob C Jurmain. 2011. Bayesian knowledge-driven ontologies: intuitive uncertainty reasoning for semantic networks. In *2011 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, 856–863. DOI: 10.1109/ICSMC.2011.6083717.

[115] Ying Shen, Yaliang Li, Hai-Tao Zheng, Buzhou Tang, and Min Yang. 2019. Enhancing ontology-driven diagnostic reasoning with a symptom-dependency-aware naïve bayes classifier. *BMC Bioinformatics*, 20, 1, 330. https://doi.org/10.1186/s12859-019-2924-0.

[116] Salvatore F. Pileggi. 2015. An individual-centric probabilistic extension for OWL: modelling the uncertainness. *Procedia Computer Science*, 51, 1742–1751. ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2015.05.378.

[117] Franz Baader, Patrick Koopmann, and Anni-Yasmin Turhan. 2017. Using ontologies to query probabilistic numerical data. In *Frontiers of Combining Systems*. Springer International Publishing, Cham, 77–94. DOI: https://doi.org/10.1007/978-3-319-66167-4_5.

[118] Jean Christoph Jung and Carsten Lutz. 2012. Ontology-based access to probabilistic data with OWL QL. In *The Semantic Web – ISWC 2012*. Springer Berlin Heidelberg, Berlin, Heidelberg, 182–197. DOI: https://doi.org/10.1007/978-3-642-35176-1_12.

[119] Liu Yuan. 2018. Combining concept learning and probabilistic information retrieval model to understand user's searching intent in OWL knowledge base. In *Knowledge Management and Acquisition for Intelligent Systems*. Springer International Publishing, Cham, 76–89. DOI: https://doi.org/10.1007/978-3-319-97289-3_6.

[120] Fabrizio Riguzzi, Elena Bellodi, Evelina Lamma, and Riccardo Zese. 2015. Probabilistic description logics under the distribution semantics. *Semantic Web*, 6, (June 2015), 477–501. DOI: 10.3233/SW-140154.

[121] Janna Hastings, Ludger Jansen, Christoph Steinbeck, and Stefan Schulz. 2011. Modelling threshold phenomena in OWL: metabolite concentrations as evidence for disorders. *CEUR Workshop Proceedings*, 796, (January 2011). https://api.semanticscholar.org/CorpusID:4434717.

[122] Raghav Ramachandran, Guilin Qi, Kewen Wang, Junhu Wang, and John Thornton. 2012. Probabilistic reasoning in DL-Lite. In *PRICAI 2012: Trends in Artificial Intelligence*. Springer Berlin Heidelberg, Berlin, Heidelberg, 480–491. DOI: https://doi.org/10.1007/978-3-642-32695-0_43.

[123] Fabrizio Riguzzi, Elena Bellodi, and Evelina Lamma. 2012. Probabilistic ontologies in datalog+/-. *CEUR Workshop Proceedings*, 857, (January 2012), 221–235. https://api.semanticscholar.org/CorpusID:9768762.

[124] Mathias Niepert, Jan Noessner, and Heiner Stuckenschmidt. 2011. Log-linear description logics. In *IJCAI*, 2153–2158. DOI: 10.5591/978-1-57735-516-8/IJCAI11-359.

[125] Jose-Raul Ruiz-Sarmiento, Cipriano Galindo, and Javier Gonzalez-Jimenez. 2017. Building multiversal semantic maps for mobile robot operation. *Knowledge-Based Systems*, 119, 257–272. DOI: https://doi.org/10.1016/j.knosys.2016.12.016.

[126] José-Raúl Ruiz-Sarmiento, Cipriano Galindo, and Javier Gonzalez-Jimenez. 2015. Exploiting semantic knowledge for robot object recognition. *Knowledge-Based Systems*, 86, 131–142. ISSN: 0950-7051. DOI: https://doi.org/10.1016/j.knosys.2015.05.032.

[127] 2020. *Introduction to probabilistic ontologies. Reasoning Web. Declarative Artificial Intelligence: Tutorial Lectures*. Springer International Publishing, Cham, 1–35. DOI: 10.1007/978-3-030-60067-9_1.

[128] Ismail Ilkan Ceylan and Rafael Peñaloza. 2015. Dynamic bayesian ontology languages. *ArXiv*, abs/1506.08030. https://api.semanticscholar.org/CorpusID:9650356.

[129] Rommel N. Carvalho, Kathryn B. Laskey, and Paulo C.G. Costa. 2017. PR-OWL – a language for defining probabilistic ontologies. *International Journal of Approximate Reasoning*, 91, 56–79. ISSN: 0888-613X. DOI: https://doi.org/10.1016/j.ijar.2017.08.011.

[130] Saminda Abeyruwan, Ubbo Visser, Vance Lemmon, and Stephan Schürer. 2013. PrOntoLearn: unsupervised lexico-semantic ontology generation using probabilistic methods. In *Uncertainty Reasoning for the Semantic Web II*. Springer Berlin Heidelberg, Berlin, Heidelberg, 217–236. DOI: https://doi.org/10.1007/978-3-642-35975-0_12.

[131] Víctor Gutiérrez-Basulto, Jean Christoph Jung, Carsten Lutz, and Lutz Schröder. 2017. Probabilistic description logics for subjective uncertainty. *Journal of Artificial Intelligence Research*, 58, 1–66. DOI: https://doi.org/10.1613/jair.5222.

[132] Shou Matsumoto, Kathryn B. Laskey, and Paulo C. G. Costa. 2016. PR-OWL decision: toward reusable ontology language for decision making under uncertainty. In *Semantic Technologies for Intelligence, Defense, and Security*. https://api.semanticscholar.org/CorpusID:16129757.

[133] Panos Alexopoulos, Manolis Wallace, Konstantinos Kafentzis, and Dimitris Askounis. 2010. Utilizing imprecise knowledge in ontology-based cbr systems by means of fuzzy algebra. *International Journal of Fuzzy Systems*, 12, 1, 1. DOI: 10.30000/ijfs.201003.0001.

[134] Ching-man Au Yeung and Ho-Fung Leung. 2010. A formal model of ontology for handling fuzzy membership and typicality of instances. *The Computer Journal*, 53, 3, 316–341. DOI: 10.1093/comjnl/bxn060.

[135] Yi Cai and Ho-fung Leung. 2011. Formalizing object membership in fuzzy ontology with property importance and property priority. In *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)*. IEEE, 1719–1726. DOI: 10.1109/FUZZY.2011.6007511.

[136] Chang-Shing Lee, Mei-Hui Wang, and Hani Hagras. 2010. A type-2 fuzzy ontology and its application to personal diabetic-diet recommendation. *IEEE Transactions on Fuzzy Systems*, 18, 2, 374–395. DOI: 10.1109/TFUZZ.2010.2042454.

[137] Mei-Hui Wang, Chang-Shing Lee, Zhi-Wei Chen, Chi-Fang Lo, Su-E Kuo, Hui-Ching Kuo, and Hui-Hua Cheng. 2010. Property and application of fuzzy ontology for dietary assessment. In *International Conference on Fuzzy Systems*, 1–8. DOI: 10.1109/FUZZY.2010.5584619.

[138] 2013. *A type-2 FML-based meeting scheduling support system. On the Power of Fuzzy Markup Language*. Springer Berlin Heidelberg, Berlin, Heidelberg, 169–187. DOI: 10.1007/978-3-642-35488-5_10.

[139] Chang-Shing Lee, Mei-Hui Wang, Shi-Jim Yen, YL Chen, Cheng-Wei Chou, Guillaume Chaslot, Jean-Baptiste Hoock, Arpad Rimmel, and Hassen Doghmen. 2010. An ontology-based fuzzy inference system for computer go applications. *International Journal of Fuzzy Systems*, 12, (June 2010). DOI: 10.30000/IJFS.201006.0002.

[140] Chang-Shing Lee, Mei-Hui Wang, Zhi-Rong Yan, Yu-Jen Chen, Hassen Doghmen, and Olivier Teytaud. 2010. FML-based type-2 fuzzy ontology for computer go knowledge representation. In *2010 International Conference on System Science and Engineering*, 63–68. DOI: 10.1109/ICSSE.2010.5551703.

[141] Chang-Shing Lee, Mei-Hui Wang, Yu-Jen Chen, Hani Hagras, Meng-Jhen Wu, and Olivier Teytaud. 2012. Genetic fuzzy markup language for game of NoGo. *Knowledge-Based Systems*, 34, 64–80. ISSN: 0950-7051. DOI: https://doi.org/10.1016/j.knosys.2012.02.006.

[142] Chang-Shing Lee, Mei-Hui Wang, Meng-Jhen Wu, Olivier Teytaud, and Shi-Jim Yen. 2015. T2FS-based adaptive linguistic assessment system for semantic analysis and human performance evaluation on game of go. *IEEE Transactions on Fuzzy Systems*, 23, 2, 400–420. DOI: 10.1109/TFUZZ.2014.2312989.

[143] Hsien-De Huang, Chang-Shing Lee, Mei-Hui Wang, and Hung-Yu Kao. 2014. IT2FS-based ontology with soft-computing mechanism for malware behavior analysis. *Soft Computing*, 18, 2, 267–284. DOI: 10.1007/s00500-013-1056-0.

[144] 2013. *Type-2 fuzzy set and fuzzy ontology for diet application. Advances in Type-2 Fuzzy Sets and Systems: Theory and Applications*. Springer, New York, NY, 237–256. DOI: 10.1007/978-1-4614-6666-6_15.

[145] Zong Min Ma, Fu Zhang, Li Yan, and Jingwei Cheng. 2011. Extracting knowledge from fuzzy relational databases with description logic. *Integrated Computer-Aided Engineering*, 18, 2, 181–200. DOI: 10.3233/ICA-2011-0366.

[146] Fu Zhang, Z.M. Ma, Li Yan, and Yu Wang. 2012. A description logic approach for representing and reasoning on fuzzy object-oriented database models. *Fuzzy Sets and Systems*, 186, 1, 1–25. Theme: Logic and computer science. ISSN: 0165-0114. DOI: https://doi.org/10.1016/j.fss.2011.04.010.

[147] Fu Zhang, ZM Ma, and Xu Chen. 2015. Formalizing fuzzy object-oriented database models using fuzzy ontologies. *J. Intell. Fuzzy Syst.*, 29, 4, (October 2015), 1407–1420. DOI: 10.3233/IFS-151532.

[148] Umberto Straccia. 2010. An ontology mediated multimedia information retrieval system. In *2010 40th IEEE International Symposium on Multiple-Valued Logic*, 319–324. DOI: 10.1109/ISMVL.2010.65.

[149] Cristiane Yaguinuma, Marilde TP Santos, Heloisa A Camargo, Maria C Nicoletti, and Tatiane M Nogueira. 2012. Fuzz-onto: a meta-ontology for representing fuzzy elements and supporting fuzzy classification rules. In *2012 12th International Conference on Intelligent Systems Design and Applications (ISDA)*. IEEE, 166–171. DOI: 10.1109/ISDA.2012.6416531.

[150] Asma Djellal and Zizette Boufaida. 2017. Individual classification: an ontological fuzzy based approach. *Informatica*, 41, 2. https://api.semanticscholar.org/CorpusID:36990380.

[151] Giorgos Stoilos, Tassos Venetis, and Giorgos Stamou. 2015. A fuzzy extension to the OWL 2 RL ontology language. *The Computer Journal*, 58, 11, 2956–2971. DOI: 10.1093/comjnl/bxv028.

[152] Fu Zhang, Z.M. Ma, Li Yan, and Jingwei Cheng. 2013. Construction of fuzzy OWL ontologies from fuzzy EER models: a semantics-preserving approach. *Fuzzy Sets and Systems*, 229, 1–32. ISSN: 0165-0114. DOI: https://doi.org/10.1016/j.fss.2013.03.016.

[153] Fu Zhang, Li Yan, and Z. M. Ma. 2012. Reasoning of fuzzy relational databases with fuzzy ontologies. *International Journal of Intelligent Systems*, 27, 6, 613–634. DOI: https://doi.org/10.1002/int.21540. https://onlinelibrary.wiley.com/doi/abs/10.1002/int.21540.

[154] Theofilos Mailis, Giorgos Stoilos, Nikolaos Simou, Giorgos Stamou, and Stefanos Kollias. 2012. Tractable reasoning with vague knowledge using fuzzy $\mathcal{EL}$++. *Journal of Intelligent Information Systems*, 39, (October 2012). DOI: 10.1007/s10844-012-0195-6.

[155] Stefan Borgwardt and Rafael Peñaloza. 2011. Description logics over lattices with multi-valued ontologies. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*. AAAI Press, Barcelona, Spain, 768–773. DOI: 10.5591/978-1-57735-516-8/IJCAI11-135.

[156] Ruixuan Li, Kunmei Wen, Xiwu Gu, Yuhua Li, Xiaolin Sun, and Bing Li. 2011. Type-2 fuzzy description logic. *Frontiers of Computer Science in China*, 5, 2, 205–215. DOI: 10.1007/s11704-011-0109-8.

[157] Yuncheng Jiang, Yong Tang, Ju Wang, Peimin Deng, and Suqin Tang. 2010. Expressive fuzzy description logics over lattices. *Knowledge-Based Systems*, 23, 2, 150–161. ISSN: 0950-7051. DOI: https://doi.org/10.1016/j.knosys.2009.11.002.

[158] Theofilos Mailis, Giorgos Stoilos, and Giorgos Stamou. 2010. Expressive reasoning with horn rules and fuzzy description logics. *Knowledge and Information Systems*, 25, 1, 105–136. DOI: 10.1007/s10115-010-0304-1.

[159] Céline Hudelot, Jamal Atif, and Isabelle Bloch. 2010. Integrating bipolar fuzzy mathematical morphology in description logics for spatial reasoning. In *Proceedings of the 2010 Conference on ECAI 2010*. IOS Press, NLD, 497–502. DOI: 10.3233/978-1-60750-606-5-497.

[160] Jidi Zhao and Harold Boley. 2010. Knowledge representation and reasoning in norm-parameterized fuzzy description logics. In *Canadian Semantic Web*. Springer US, Boston, MA, 27–53. DOI: 10.1007/978-1-4419-7335-1_2.

[161] Fernando Bobillo and Umberto Straccia. 2017. Generalizing type-2 fuzzy ontologies and type-2 fuzzy description logics. *International Journal of Approximate Reasoning*, 87, 40–66. DOI: https://doi.org/10.1016/j.ijar.2017.04.012.

[162] ASM Kayes, Wenny Rahayu, Tharam Dillon, Elizabeth Chang, and Jun Han. 2019. Context-aware access control with imprecise context characterization for cloud-based data resources. *Future Generation Computer Systems*, 93, 237–255. DOI: https://doi.org/10.1016/j.future.2018.10.036.

[163] Hsien-De Huang, Chang-Shing Lee, Hani Hagras, and Hung-Yu Kao. 2012. TWMAN+: a type-2 fuzzy ontology model for malware behavior analysis. In *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2821–2826. DOI: 10.1109/ICSMC.2012.6378176.

[164] Chang-Shing Lee, Mei-Hui Wang, and Shun-Teng Lan. 2015. Adaptive personalized diet linguistic recommendation mechanism based on type-2 fuzzy sets and genetic fuzzy markup language. *IEEE Transactions on Fuzzy Systems*, 23, 5, 1777–1802. DOI: 10.1109/TFUZZ.2014.2379256.

[165] Hanen Ghorbel, Afef Bahri, and Rafik Bouaziz. 2010. Fuzzy ontologies building method: fuzzy OntoMethodology. In *2010 Annual Meeting of the North American Fuzzy Information Processing Society*, 1–8. DOI: 10.1109/NAFIPS.2010.5548211.

[166] Liu Yang, Zhi Hu, and Jun Long. 2016. Expert selection service system by fuzzy ontology modelling. *International Journal of Computational Science and Engineering*, 12, (January 2016), 124. doi: 10.1504/IJCSE.2016.076215.

[167] Robin Wikström and József Mezei. 2014. Intrusion detection with type-2 fuzzy ontologies and similarity measures. *Intelligent Methods for Cyber Warfare*, 151–172. doi: https://doi.org/10.1007/978-3-319-08624-8_7.

[168] József Mezei, Robin Wikström, and Christer Carlsson. 2015. Aggregating linguistic expert knowledge in type-2 fuzzy ontologies. *Applied Soft Computing*, 35, 911–920. issn: 1568-4946. doi: https://doi.org/10.1016/j.asoc.2015.03.023.

[169] Ignacio Pérez, Robin Wikström, József Mezei, Christer Carlsson, and Enrique Herrera-Viedma. 2013. A new consensus model for group decision making using fuzzy ontology. *Soft Computing*, 17, 9, 1617–1627. doi: 10.1007/s00500-012-0975-5.

[170] Christer Carlsson, Matteo Brunelli, and József Mezei. 2012. Decision making with a fuzzy ontology. *Soft Computing*, 16, 7, 1143–1152. doi: 10.1007/s00500-011-0789-x.

[171] Marek Reformat, Christer Carlsson, József Mezei, and Matteo Brunelli. 2013. Fuzzy ontology used for knowledge mobilization. *International Journal of Intelligent Systems*, 28, (January 2013). doi: 10.1002/int.21574.

[172] Yanhui Lv. 2010. Constructing fuzzy ontology based on fuzzy EER model. In *2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery*. Volume 3, 1109–1113. doi: 10.1109/FSKD.2010.5569169.

[173] Fernando Bobillo, Miguel Delgado, Juan Gómez-Romero, and Umberto Straccia. 2012. Joining gödel and zadeh fuzzy logics in fuzzy description logics. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 20, 04, 475–508. doi: 10.1142/S0218488512500249.

[174] Amira Aloui, Alaa Ayadi, and Amel Grissa-Touzi. 2014. A semi-automatic method to fuzzy-ontology design by using clustering and formal concept analysis. *DBKDA 2014*, 2, 27. https://api.semanticscholar.org/CorpusID:251884737.

[175] J Jelsteen, D Evangelin, J Alice Pushparani, and J Nelson Samuel Jebastin. 2013. Ontology learning process using fuzzy formal concept analysis. *International Journal of Engineering Trends and Technology (IJETT)*, 4, 2, 148–152. https://api.semanticscholar.org/CorpusID:212519659.

[176] Arturo Urquizo Baldárrago, Marilde TP Santos, and Antonio F do Prado. 2012. UPFON: unified process for building fuzzy ontology. In *2012 9th International Conference on Fuzzy Systems and Knowledge Discovery*. IEEE, 617–622. doi: 10.1109/FSKD.2012.6234174.

[177] Mounir Hemam, Meriem Djezzar, and Zianou Ahmed Seghir. 2018. Multi-viewpoints ontological knowledge representation: a fuzzy description logics based approach. In *Proceedings of the Fourth International Conference on Engineering & MIS 2018* (ICEMIS '18). Association for Computing Machinery, New York, NY, USA. doi: 10.1145/3234698.3234761. https://doi.org/10.1145/3234698.3234761.

[178] Qing-long Wang, Zhi-xue Wang, Ting-ting Zhang, and Wei-xing Zhu. 2017. A quality requirements model and verification approach for system of systems based on description logic. *Frontiers of Information Technology & Electronic Engineering*, 18, 3, 346–361. doi: 10.1631/FITEE.1500309.

[179] C. Martinez-Cruz, C. Porcel, J. Bernabé-Moreno, and E. Herrera-Viedma. 2015. A model to represent users trust in recommender systems using ontologies and fuzzy linguistic modeling. *Information Sciences*, 311, 102–118. doi: https://doi.org/10.1016/j.ins.2015.03.013.

[180] DaZhou Kang, BaoWen Xu, and YanHui Li. 2013. A fuzzy extension for SHOIQ based on comparisons between degrees of membership. *Science China Information Sciences*, 56, 8, 1–11. doi: 10.1007/s11432-011-4423-4.

[181] Fernando Bobillo and Umberto Straccia. 2013. Finite fuzzy description logics and crisp representations. In *Uncertainty Reasoning for the Semantic Web II*. Springer Berlin Heidelberg, Berlin, Heidelberg, 99–118. doi: https://doi.org/10.1007/978-3-642-35975-0_6.

[182] Fu Zhang, Z.M. Ma, and Li Yan. 2013. Construction of fuzzy ontologies from fuzzy XML models. *Knowledge-Based Systems*, 42, 20–39. issn: 0950-7051. doi: https://doi.org/10.1016/j.knosys.2012.12.015.

[183] Fu Zhang and Zongmin Ma. 2013. Construction of fuzzy ontologies from fuzzy UML models. *Int. J. Computational Intelligence Systems*, 6, (May 2013), 442–472. doi: 10.1080/18756891.2013.780735.

[184] Linlin Huo, Jihong Ouyang, Wei Sun, and Dayou Liu. 2012. Interval-valued fuzzy description logic IFALCN. *Journal of Computational Information Systems*, 8, (April 2012). doi: 10.1109/ICICISYS.2010.5658500.

[185] Fu Zhang, Z. M. Ma, Gaofeng Fan, and Xing Wang. 2010. Automatic fuzzy semantic web ontology learning from fuzzy object-oriented database model. In *Database and Expert Systems Applications*. Springer Berlin Heidelberg, Berlin, Heidelberg, 16–30. doi: 10.1007/978-3-642-15364-8_2.

[186] Afef Bahri, Rafik Bouziz, and Faïez Gargouri. 2010. A generalized fuzzy extension of EL++. In *2010 Annual Meeting of the North American Fuzzy Information Processing Society*, 1–6. doi: 10.1109/NAFIPS.2010.5548202.

[187] Zong Min Ma, Fu Zhang, Li Yan, and Yanhui Lv. 2010. Formal semantics-preserving translation from fuzzy ER model to fuzzy OWL DL ontology. *Web Intelligence and Agent Systems: An International Journal*, 8, 4, 397–412. doi: 10.3233/WIA-2010-0199.

[188] Huizhang Shen, Jidi Zhao, and Jing Dong. 2010. Building fuzzy hybrid horn logic rules on top of ontologies. In *2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery*. Volume 3, 1105–1108. doi: 10.1109/FSKD.2010.5569173.

[189] Àngel Garcı´a-Cerdaña, Eva Armengol, and Francesc Esteva. 2010. Fuzzy description logics and t-norm based fuzzy logics. *International Journal of Approximate Reasoning*, 51, 6, 632–655. doi: https://doi.org/10.1016/j.ijar.2010.01.001.

[190] Fernando Bobillo and Umberto Straccia. 2011. Fuzzy ontology representation using OWL2. *International Journal of Approximate Reasoning*, 52, 7, 1073–1094. doi: https://doi.org/10.1016/j.ijar.2011.05.003.

[191] Marco Cerami, Àngel Garcia-Cerdaña, and Francesc Esteva. 2010. From classical description logic to n-graded fuzzy description logic. In *International Conference on Fuzzy Systems*, 1–8. doi: 10.1109/FUZZY.2010.5584114.

[192] Henrique Viana, Joao Alcântara, and ATC Martins. 2011. Paraconsistent rough description logic. In *Proc. of the 2011 Int. Workshop on Description Logics (DL'11)*. Volume 745. Citeseer, 587–597. https://api.semanticscholar.org/CorpusID:18957701.

[193] Jin Song Dong, Yuzhang Feng, Yuan-Fang Li, Colin Keng-Yan Tan, Bimlesh Wadhwa, and Hai H. Wang. 2015. BOWL: augmenting the semantic web with beliefs. *Innovations in Systems and Software Engineering*, 11, 3, 203–215. https://doi.org/10.1007/s11334-015-0243-9.

[194] L. Karanikola and I. Karali. 2016. Dempster-shafer logical model for fuzzy description logics. In *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE, Piscataway, NJ, USA. DOI: 10.1109/SSCI.2016.7849960.

[195] Fernando Bobillo and Umberto Straccia. 2012. Generalized fuzzy rough description logics. *Information Sciences*, 189, 43–62. ISSN: 0020-0255. DOI: https://doi.org/10.1016/j.ins.2011.10.002. https://www.sciencedirect.com/science/article/pii/S0020025511005342.

[196] Mustapha Bourahla. 2015. Exact reasoning over imprecise ontologies. In *Computer Science and Its Applications*. Springer International Publishing, Cham, 355–366. DOI: https://doi.org/10.1007/978-3-319-19578-0_29.

[197] Carlos Iván Chesñevar Sergio Alejandro Gómez and Guillermo Ricardo Simari. 2010. Reasoning with inconsistent ontologies through argumentation. *Applied Artificial Intelligence*, 24, 1-2, 102–148. DOI: 10.1080/08839510903448692.

[198] Yuncheng Jiang, Yong Tang, Qimai Chen, Ju Wang, and Suqin Tang. 2010. Extending soft sets with description logics. *Computers and Mathematics with Applications*, 59, 6, 2087–2096. ISSN: 0898-1221. DOI: https://doi.org/10.1016/j.camwa.2009.12.014.

[199] Kevin P Murphy et al. 2006. Naive bayes classifiers. *University of British Columbia*, 18, 60, 1–8.

[200] SATO Taisuke. 1995. A statistical learning method for logic programs with distribution semantics. In *Proceedings of ICLP*. Citeseer, 715–729.

[201] Colin Keng-Yan Tan. 2003. *Belief Augmented Frames*. PhD thesis. National University of Singapore. https://api.semanticscholar.org/CorpusID:16273645.

[202] James R. Shaw. 2014. What is a truth-value gap? *Linguistics and Philosophy*, 37, 6, 503–534. DOI: 10.1007/s10988-014-9160-x.

[203] Alejandro J. García and Guillermo R. Simari. 2004. Defeasible logic programming: an argumentative approach. *Theory Pract. Log. Program.*, 4, 2, 95–138. DOI: 10.1017/S1471068403001674.

[204] D. Molodtsov. 1999. Soft set theory—first results. *Computers and Mathematics with Applications*, 37, 4, 19–31. ISSN: 0898-1221. DOI: https://doi.org/10.1016/S0898-1221(99)00056-5.

[205] Sarah Ghidalia, Ouassila Labbani Narsis, Aurélie Bertaux, and Christophe Nicolle. 2024. Combining machine learning and ontology: a systematic literature review. *arXiv preprint arXiv:2401.07744*.

[206] Krithikha Saravanan and Velammal Bhagavathiappan. 2024. Innovative agricultural ontology construction using NLP methodologies and graph neural network. *Engineering Science and Technology, an International Journal*, 52, 101675. DOI: https://doi.org/10.1016/j.jestch.2024.101675.

[207] Luc De Raedt, Angelika Kimmig, and Hannu Toivonen. 2007. Problog: a probabilistic prolog and its application in link discovery. In *Proceedings of the 20th International Joint Conference on Artifical Intelligence* (IJCAI'07). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2468–2473.

[208] Max Bramer. 2014. *Logic Programming with Prolog*. (2nd edition). Springer Publishing Company, Incorporated. ISBN: 144715486X.

[209] İsmail İlkan Ceylan and Rafael Peñaloza Julian Mendez. 2015. The bayesian ontology reasoner is BORN! In *Proceedings of the 4th International Workshop(ORE 2015)*. Volume 1387. CEUR Workshop Proceedings, 8–14. https://api.semanticscholar.org/CorpusID:15369492.

[210] Giorgos Stoilos, Giorgos Stamou, and Jeff Z Pan. 2010. Fuzzy extensions of OWL: logical properties and reduction to fuzzy description logics. *International journal of approximate reasoning*, 51, 6, 656–679. DOI: https://doi.org/10.1016/j.ijar.2010.01.005.

[211] V. V. Cross. 2014. Fuzzy ontologies: the state of the art. In *2014 IEEE Conference on Norbert Wiener in the 21st Century (21CW)*. IEEE, Boston, MA, USA, 1–8. DOI: 10.1109/NORBERT.2014.6893867.

[212] Nilavu Devadoss and Sivakumar Ramakrishnan. 2015. Knowledge representation using fuzzy ontologies – a review. *International Journal of Computer Science and Information Technologies*, 6, 5, 4304–4308. https://api.semanticscholar.org/CorpusID:17660365.

[213] A.A. Kindo, G. Kaladzavi, S. Malo, G. Camara, T.M.Y. Tapsoba, and Kolyang. 2020. Fuzzy logic approach for knowledge modeling in an ontology: a review. In *2nd International Conference SCCIC 2020*, 1–8. DOI: 10.1109/SCCIC51516.2020.9377335.

[214] Diana Kalibatiene and Jolanta Miliauskaitė. 2021. A systematic mapping with bibliometric analysis on information systems using ontology and fuzzy logic. *Applied Sciences*, 11, 7. DOI: https://doi.org/10.3390/app11073003.

[215] I. Qasim, M. Alam, S. Khan, A.W. Khan, K.M. Malik, M. Saleem, and S.A.C. Bukhari. 2020. A comprehensive review of type-2 fuzzy ontology. *Artificial Intelligence Review*, 53, 2, 1187–1206. DOI: https://doi.org/10.1007/s10462-019-09693-9.

[216] Umberto Straccia. 2017. Fuzzy semantic web languages and beyond. In *Advances in Artificial Intelligence: From Theory to Practice*. Springer International Publishing, Cham, 3–8. DOI: https://doi.org/10.1007/978-3-319-60042-0_1.

[217] Antonio Lieto and Gian Luca Pozzato. 2024. Two semantic interpretations of probabilities in description logics of typicality. *Logic Journal of the IGPL*, (July 2024), jzae086. DOI: 10.1093/jigpal/jzae086.

[218] Antonio Lieto and Gian Luca Pozzato. 2020. A description logic framework for commonsense conceptual combination integrating typicality, probabilities and cognitive heuristics. *Journal of Experimental & Theoretical Artificial Intelligence*, 32, 5, 769–804. DOI: 10.1080/0952813X.2019.1672799.

[219] Gian Luca Pozzato. 2019. Typicalities and probabilities of exceptions in nonmotonic description logics. *International Journal of Approximate Reasoning*, 107, 81–100. ISSN: 0888-613X. DOI: https://doi.org/10.1016/j.ijar.2019.02.003.

[220] Gian Luca Pozzato, Gabriele Soriano, et al. 2019. PEAR: a tool for reasoning about scenarios and probabilities in description logics of typicality. In *CEUR WORKSHOP PROCEEDINGS*. Volume 2396. CEUR-WS, 147–156.

[221] Laura Giordano. 2021. On the KLM properties of a fuzzy DL with typicality. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*. Springer International Publishing, Cham, 557–571.

[222] Laura Giordano et al. 2022. From weighted conditionals with typicality to a gradual argumentation semantics and back. In *NMR*, 127–138.

8 Supplementary Materials

In the Appendix, we present tables (Table 13—Table 20) that systematically classify the reviewed uncertainty modelling studies. We classify the papers by their focus: modelling, modelling combined with reasoning, specific reasoning algorithms or reasoners, and application-based studies. The first two classes (i.e., modelling and modelling combined with reasoning) are further divided into DL, and non-DL based papers. In some cases, additional subcategories are introduced based on specific branches of the formalism. For instance, within the probabilistic classification, we distinguish between BN-based and MLN-based studies.

These tables include a dedicated section that details the excluded references, organized into distinct subdivisions: studies not found (i.e., no accessible version available), older versions of expanded papers, works lacking sufficient technical details on modelling, and marginal cases. Marginal studies are further divided into subcategories, including those where classical ontologies serve as input to uncertainty reasoning, merging approaches, ontology repair, and other peripheral contributions. This structured classification provides a systematic and detailed overview of both the included and excluded works, emphasizing the diversity of methodologies and scopes in uncertainty modelling across the reviewed literature, categorized by the adopted formalism.

Table 13. Possibilistic Ontology Papers Classified based on Scope

Scope	Sub-scope	Sub sub-scope	References
Only modelling	DL-based	–	[84, 87]
	Non-DL-based	–	[89]
Modelling and reasoning	DL-based	–	[83, 85, 86, 88, 90]
	Non-DL-based	–	–
Reasoning algorithms and reasoners	–	–	[223, 224]
Application	–	–	–
Excluded	Not found	–	–
	Old version	–	[225, 226, 227, 228, 229, 230]
	Short on technical details	–	[231]
	Marginal	Classical ontology is input to uncertainty reasoning	[232, 233]
		Merging approach	[234]
		Ontology repair	[235]

Table 14. DST Ontology Papers Classified based on Scope

Scope	Sub-scope	Sub sub-scope	References
Only modelling	DL-based	–	[94]
	Non-DL-based	–	–
Modelling and reasoning	DL-based	–	[91, 92, 93, 95, 96]
	Non-DL-based	–	–
Reasoning algorithms and reasoners	–	–	[236]
Application	–	–	–
Excluded	Not found	–	–
	Old version	–	[237, 238, 239, 240, 241, 242, 243, 244]
	Short on technical details	–	[245, 246]
	Marginal	No uncertainty modelling involved	[247, 248, 249, 250]
		Merging approach	[251]
		No ontology involved	[252]
		Classical ontology is input to uncertainty reasoning	[253, 254, 255]

Table 15. Rough Ontology Papers Classified based on Scope

Scope	Sub-scope	Sub sub-scope	References
Only modelling	DL-based	—	[100]
	Non-DL-based	—	[99]
Modelling and reasoning	DL-based	—	[97, 98]
	Non-DL-based	—	—
Reasoning algorithms and reasoners	—	—	[256]
Application	—	—	—
Excluded	Not found	—	[257]
	Old version	—	[258, 259]
	Short on technical details	—	[260, 261]
	Marginal	Granular approach	[262]

Table 16. Paraconsistent Ontology Papers Classified based on Scope

Scope	Sub-scope	Sub sub-scope	References
Only modelling	DL-based	Multi-valued paraconsistent logic	[78, 101, 103, 109]
		Distance-based paraconsistent logic	[106]
	Non-DL-based	—	—
Modelling and reasoning	DL-based	Multi-valued paraconsistent logic	[102, 104, 107, 108]
		Argumentation-based paraconsistent	[105]
	Non-DL-based	—	—
Reasoning algorithms and reasoners	—	—	[263, 264]
Application	—	—	—
Excluded	Not found	—	—
	Old version	—	[265, 266, 267, 268, 269, 270]
	Short on technical details	—	—
	Marginal	Ontology integration process	[271, 272]

Table 17. Hybrid Ontology Papers Classified based on Scope

Scope	Sub-scope	Sub sub-scope	References
Only modelling	DL-based	DST + fuzzy	[194]
	Non-DL-based	—	—
Modelling and reasoning	DL-based	Paraconsistent + rough set	[192]
		Rough + fuzzy	[195]
		BAF + fuzzy	[193]
	Non-DL-based	—	—
Reasoning algorithms and reasoners	—	—	[273]
Application	—	—	[274, 275, 276]
Excluded	Not found	—	[277, 278, 279, 280]
	Old version	—	—
	Short on technical details	—	[281, 282, 283, 284]
	Marginal	Classical ontology is input to uncertainty reasoning	[285, 286, 287]
		Not working with ontology	[288, 289, 290]

Table 18. Probabilistic Ontology Papers Classified based on Scope

Scope	Sub-scope	Sub sub-scope	References
Only modelling	DL-based	BN	[130]
	Non-DL-based	CRF	[126]
Modelling and reasoning	DL-based	Probabilistic logic	[115, 116, 118, 121]
		Herbrand set	[122, 124]
		Discrete and continuous distribution	[111, 117]
		Type-2 PFOL (i.e., epistemic)	[119, 120, 131]
		Type-1 and type-2 PFOL	[123]
		BN	[112, 113, 114, 127, 128]
		MEBN	[129]
		MLN	[110]
		Multi-Entity Decision Graph	[132]
	Non-DL-based	CRF	[125]
Reasoning algorithms and reasoners	–	–	[291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 209, 313, 314, 315, 316, 317]
Application	–	–	[318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362]
Excluded	Not found	–	[363, 364, 365, 366, 367, 368, 369, 370]
	Old version	–	[371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386]
	Short on technical details	–	[387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399]
	Marginal	Classical ontology is input to uncertainty reasoning	[400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433]
		No uncertainty modelling involved	[434, 435, 436, 437, 438, 439, 440, 441, 442]
		Ontology repair	[443, 444]
		Probabilistic RDF	[445, 446, 447, 448]
		Software engineering methodological process	[449, 450, 451, 452, 453, 454, 455, 456, 457]
		Translate the probabilistic DL into probabilistic first-order logic	[291, 458]
		Probabilistic knowledge base (not ontology)	[459, 460, 461]
		Translate ontology into BN or its extensions	[462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478]
		About preferences	[479, 480, 481, 482, 483]
		Ontology learning process	[484, 485, 486]

Table 19. Fuzzy Ontology Papers Classified based on Scope

Scope	Sub-scope	Sub sub-scope	References
Only modelling	DL-based	Transfer(UML, EER,..etc) to fuzzy ontology	[147, 172, 185]
		Linguist modelling	[159]
		Fuzzy with horn logic	[158, 188]
		Interval-type2 linguistic fuzzy logic	[168]
		Interval-type2 fuzzy logic	[167]
	Non-DL-based	Fuzzy set theory(Zadeh)	[165]
		Fuzzy set Theory+weighted properties	[134, 135]
		Linguistic modelling	[166]
Modelling and reasoning	DL-based	Transfer(UML, EER,..etc) to fuzzy ontology	[145, 146, 152, 153, 178, 182, 183, 187]
		Linguist modelling	[170, 171, 179]
		T-norm-based logic	[190]
		T-norm-based with horn logic	[151]
		Fuzzy set theory(Zadeh)	[149, 169, 174, 175, 176, 177]
		Intuitionistic fuzzy sets	[159]
		Fuzzy with horn logic	[158, 188]
		Finite chain	[181]
	Non-DL-based	Fuzzy set theory(Zadeh)	[133]
		Type-2 fuzzy logic	[136, 137, 138, 139, 140, 141, 142, 144, 164, 163]
Reasoning algorithms and reasoners	–	–	[487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515]
Application	–	–	[516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571]
Excluded	Not found	–	[572, 573, 574, 575, 576, 577, 578]
	Old version	–	[579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601]
	Short on technical details	–	[602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616]
	Marginal	Inductive logic programming	[617, 618]
		Fuzzy UML/relational database	[619, 620, 621, 622]
		Adding additional features to fuzzy-ontology	[623, 624, 625, 626, 627, 628, 629, 630]
		Natural language processing	[631]
		Software engineering methodological process	[632, 633, 634, 635, 636, 637]
		Ontology integration	[638]
		No fuzzy ontology modelling involved	[639, 640, 641, 642, 643, 644]
		Fuzzy markup language	[645, 646]

Table 20. Other Ontology Papers Classified based on Scope

Scope	Sub-scope	Sub sub-scope	References
Only modelling	DL-based	–	–
	Non-DL-based	–	–
Modelling and reasoning	DL-based	Soft set theory	[198]
		Vagueness theory	[196]
		Augmentation theory	[197]
	Non-DL-based	–	–
Reasoning algorithms and reasoners	–	–	[647, 648]
Application	–	–	–
Excluded	Not found	–	–
	Old version	–	[649]
	Short on technical details	–	–
	Marginal	Ontology integration	[650, 651]
		Ontology classification	[652]

In the following, we present a visual classification of the reviewed studies based on the guiding questions related to uncertainty in ontologies. The first tree, presented in Figure 3, illustrates the primary sources of uncertainty identified across the literature, such as sensor errors, contradictory sources, and contextual factors, which addressing question (GQ4-A). The types of uncertainty encountered (GQ4-B) include incomplete information, vagueness, and inconsistency. Finally, highlights the various locations where uncertainty manifests within ontological structures (GQ5), such as relationships between concepts, attributes, and semantic ambiguities. This hierarchical representation facilitates understanding of how different aspects of uncertainty are addressed in existing research. The second tree, presented in Figure 4, organizes the analyzed studies according to key guiding questions related to reasoning tasks, methodological involvement, and supporting technologies. It begins with GQ7, which classifies studies based on the specific reasoning tasks they address, including subsumption, satisfiability, consistency checking, instance checking, entailment, query answering, and those covering all reasoning tasks comprehensively. Following this, GQ8 identifies studies that incorporate Machine Learning techniques, while GQ9 highlights the involvement of NLP methods. The tree then branches into GQ10-A, categorizing the supported ontology languages, such as OWL, OWL2 profiles, UML, FML, and SWRL, reflecting the diverse modelling formalisms employed. Finally, GQ10-B lists the existing tools and reasoners that underpin these approaches, including OWL API, DL reasoners, FuzzyDL, DeLorean, Pronto, Protégé, and the Jena framework. Additionally, we address GQ11 by classifying the types of evaluation provided in the studies, such as use cases and experimental validations. This hierarchical classification illustrates the multi-faceted nature of uncertainty modelling research, spanning reasoning capabilities, methodological integrations, tool support, and evaluation practices.

- GQ4-A: Sources of uncertainty
  - Sensors — [85, 89, 91, 93, 117, 125]
  - Contradictory Sources — [84, 85, 86, 93, 94, 78, 101, 103, 110, 120, 197]
  - Measurement Errors — [90, 93, 100, 125]
  - Model limitations — [87, 97, 104, 105, 106, 107, 108, 125]
  - Context — [83, 88, 92, 95, 96, 98, 99, 102, 109, 111, 112, 113, 114, 115, 116, 118, 121, 127, 128, 129, 130, 131, 134, 133, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 149, 150, 152, 153, 154, 155, 158, 159, 160, 161, 164, 165, 166, 167, 168, 170, 171, 172, 173, 175, 176, 177, 178, 180, 181, 182, 183, 184, 185, 188, 189, 190, 191, 193, 195, 196, 198]
  - Undetermined — [119, 122, 123, 124, 126, 132, 148, 151, 156, 157, 162, 163, 169, 174, 179, 186, 187, 192, 194]

- GQ4-B: Types of uncertainty
  - Incomplete Information — [85, 92]
  - Vagueness and Imprecision — [133, 134, 135, 136, 137, 139, 140, 141, 142, 143, 144, 148, 149, 151, 155, 156, 157, 158, 159, 160, 161, 163, 165, 166, 167, 168, 169, 170, 171, 173, 174, 175, 176, 177, 178, 180, 181, 182, 183, 184, 185, 186, 188, 189, 190, 191, 194, 196]
  - Inconsistency — [83, 84, 86, 88, 95, 99, 78, 103, 104, 105, 106, 107, 108, 109]
  - Several types — [87, 89, 90, 91, 93, 94, 96, 97, 98, 100, 101, 102, 110, 121, 150, 154, 192, 193, 195, 197, 198]
  - Undetermined — [111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 138, 145, 146, 147, 152, 162, 153, 164, 172, 179, 187]

- GQ5: Locations of uncertainty
  - Relationship between concepts — [113, 115, 121, 126, 188]
  - Attribute of concept — [89, 98, 99, 111, 124, 129, 136, 164, 165]
  - Semantic ambiguity of concept — [91, 97, 101, 109, 130]
  - Mapping instance to concept — [92, 94, 134, 135, 138, 139, 140, 141, 142, 143, 144, 150, 162, 163, 196, 197, 198]
  - Mapping instances to relationships — [168, 169, 170, 171, 179]
  - Several locations — [83, 84, 85, 86, 87, 88, 90, 93, 95, 96, 100, 78, 102, 103, 104, 105, 106, 107, 108, 110, 112, 114, 116, 117, 118, 119, 120, 122, 123, 125, 127, 128, 131, 132, 133, 137, 145, 146, 147, 148, 149, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 166, 167, 172, 173, 174, 175, 176, 177, 178, 180, 181, 182, 183, 184, 185, 186, 187, 189, 190, 191, 192, 193, 194, 195]

Fig. 3. Visualization and Classification of Reviewed Studies Based on Guiding Questions 4-A, B and 5.

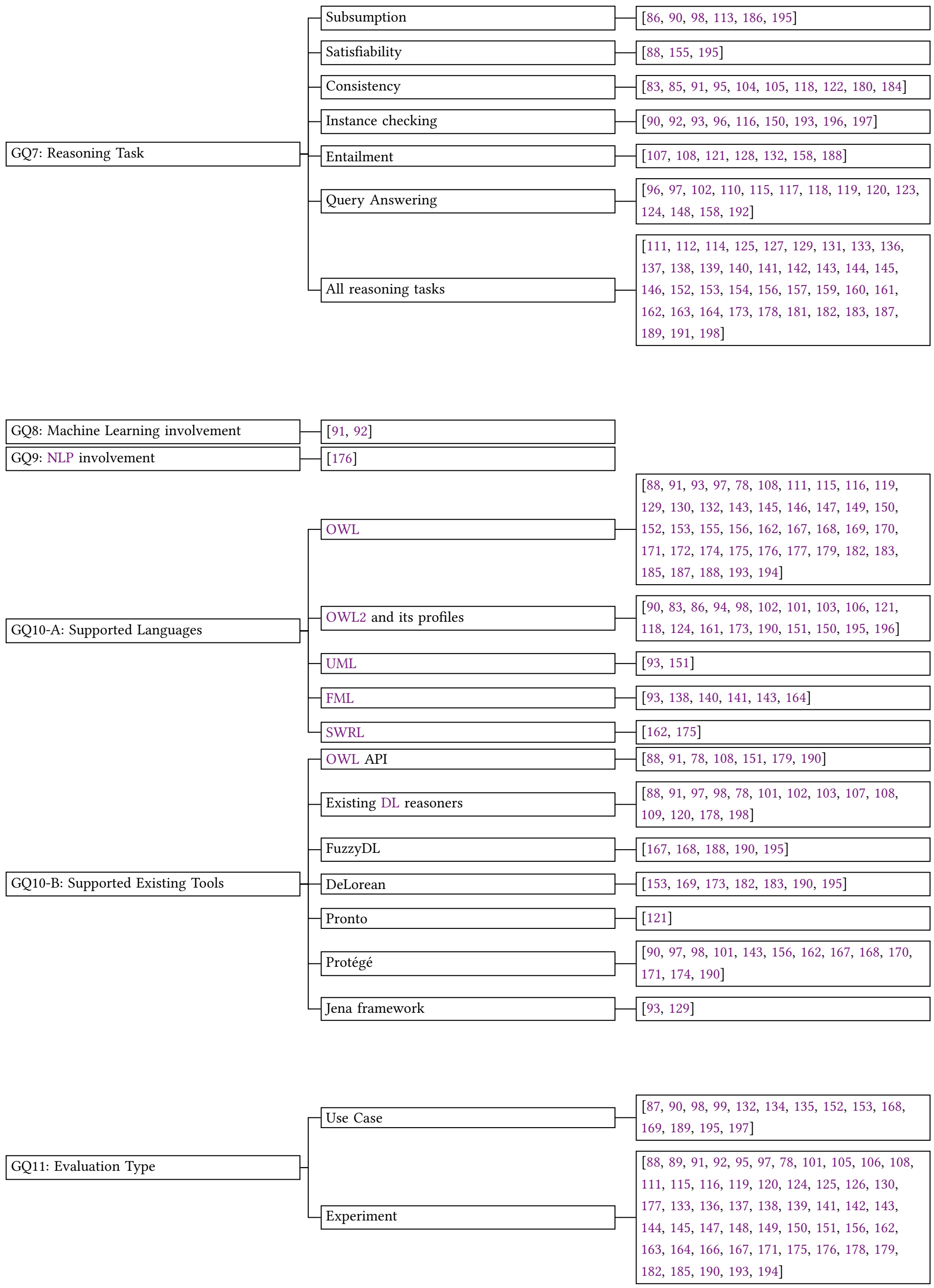


Fig. 4. Visualization and Classification of Reviewed Studies Based on Guiding Questions 7 to 11.

Excluded Readings

[223] Guilin Qi, Qiu Ji, Jeff Z. Pan, and Jianfeng Du. 2010. PossDL — a possibilistic DL reasoner for uncertainty reasoning and inconsistency handling. In *The Semantic Web: Research and Applications*. Lora Aroyo, Grigoris Antoniou, Eero Hyvönen, Annette ten Teije, Heiner Stuckenschmidt, Liliana Cabral, and Tania Tudorache, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 416–420. DOI: https://doi.org/10.1007/978-3-642-13489-0_35.

[224] Khaoula Boutouhami and Ibrahim Hussein Musa. 2019. An argumentation-based approach for computing inconsistency degree in possibilistic lightweight ontologies. *Journal of Physics: Conference Series*, 1229, 1, (May 2019), 012081. DOI: 10.1088/1742-6596/1229/1/012081.

[225] Salem Benferhat, Khaoula Boutouhami, Faiza Khellaf, and Farid Nouioua. 2017. Representing lightweight ontologies in a product-based possibility theory framework. In *Soft Methods for Data Science*. Springer International Publishing, Cham, 45–52. DOI: https://doi.org/10.1007/978-3-319-42972-4_6.

[226] Yosra Ben Salem, Rihab Idoudi, Karim Saheb Ettabaa, Kamel Hamrouni, and Basel Solaiman. 2017. Ontology based possibilistic reasoning for breast cancer aided diagnosis. In *Information Systems*. Marinos Themistocleous and Vincenzo Morabito, editors. Springer International Publishing, Cham, 353–366. DOI: https://doi.org/10.1007/978-3-319-65930-5_29.

[227] Safia Bal Bourai, Aicha Mokhtari, and Faiza Khellaf. 2013. Poss–SROIQ (D): possibilistic description logic extension toward an uncertain geographic ontology. *New Trends in Databases and Information Systems. AISC*, 241, 277–286. DOI: https://doi.org/10.1007/978-3-319-01863-8_30.

[228] Sihem Belabbes and Salem Benferhat. 2020. On dealing with conflicting, uncertain and partially ordered ontologies. In *The Thirty-Third International Flairs Conference*. https://api.semanticscholar.org/CorpusID:219319131.

[229] Sihem Belabbes and Salem Benferhat. 2022. Computing a possibility theory repair for partially preordered inconsistent ontologies. *IEEE Transactions on Fuzzy Systems*, 30, 8, 3237–3246. DOI: 10.1109/TFUZZ.2021.3107776.

[230] Salem Benferhat and Zied Bouraoui. 2013. Possibilistic DL-Lite. In *Scalable Uncertainty Management*. Springer Berlin Heidelberg, Berlin, Heidelberg, 346–359. DOI: https://doi.org/10.1007/978-3-642-40381-1_27.

[231] Albeiro Espinal, Yannis Haralambous, Dominique Bedart, and John Puentes. 2022. An ontology-based possibilistic framework for extracting relevant terms from job advertisements. In *2022 International Conference on Fuzzy Computation Theory and Applications (FCTA)*. DOI: 10.5220/0011521700003332.

[232] Amel Ben Othmane, Andrea Tettamanzi, Serena Villata, Nhan Le Thanh, and Michel Buffa. 2016. A multi-context framework for modeling an agent-based recommender system. In *International Conference on Agents and Artificial Intelligence*. Volume 2. SCITEPRESS, 31–41. DOI: 10.5220/0005686500310041.

[233] Fatiha Saïs, Rallou Thomopoulos, and Sébastien Destercke. 2010. Ontology-driven possibilistic reference fusion. In *On the Move to Meaningful Internet Systems, OTM 2010*. Robert Meersman, Tharam Dillon, and Pilar Herrero, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 1079–1096. DOI: https://doi.org/10.1007/978-3-642-16949-6_30.

[234] Salem Benferhat, Zied Bouraoui, Sylvain Lagrue, and Julien Rossit. 2014. Min-based assertional merging approach for prioritized DL-Lite knowledge bases. In *Scalable Uncertainty Management*. Umberto Straccia and Andrea Calì, editors. Springer International Publishing, Cham, 8–21. DOI: https://doi.org/10.1007/978-3-319-11508-5_2.

[235] Sihem Belabbes and Salem Benferhat. 2022. Characterizing the possibilistic repair for inconsistent partially ordered assertions. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems*. Springer International Publishing, Cham, 652–666. DOI: https://doi.org/10.1007/978-3-031-08974-9_52.

[236] Loukia Karanikola, Isambo Karali, and Sally McClean. 2013. Uncertainty reasoning for the semantic web based on dempster-shafer model. In *IISA 2013*, 1–4. DOI: 10.1109/IISA.2013.6623687.

[237] Emma Sloan, Marguerite McDaniel, William Nick, James Mayes, and Albert Esterline. 2017. Marshalling situation-based evidence in identity cases. In *SoutheastCon 2017*, 1–6. DOI: 10.1109/SECON.2017.7925318.

[238] Janelle Mason, Kofi Kyei, Hannah Foster, and Albert Esterline. 2018. A framework for identity: dempster-shafer theory the flow and combination of evidence. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, 1700–1706. DOI: 10.1109/SSCI.2018.8628738.

[239] Giuseppe Rizzo, Nicola Fanizzi, Claudia d'Amato, and Floriana Esposito. 2011. Prediction of class and property assertions on OWL ontologies through evidence combination. In *Proceedings of the International Conference on Web Intelligence, Mining and Semantics* (WIMS '11). Association for Computing Machinery, New York, NY, USA. ISBN: 9781450301480. DOI: 10.1145/1988688.1988741. https://doi.org/10.1145/1988688.1988741.

[240] Giuseppe Rizzo, Claudia d'Amato, Nicola Fanizzi, and Floriana Esposito. 2013. Assertion prediction with ontologies through evidence combination. In *Uncertainty Reasoning for the Semantic Web II*. Fernando Bobillo, Paulo C. G. Costa, Claudia d'Amato, Nicola Fanizzi, Kathryn B. Laskey, Kenneth J. Laskey, Thomas Lukasiewicz, Matthias Nickles, and Michael Pool, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 282–299. DOI: https://doi.org/10.1007/978-3-642-35975-0_15.

[241] Giuseppe Rizzo, Claudia d'Amato, Nicola Fanizzi, and Floriana Esposito. 2014. Towards evidence-based terminological decision trees. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems*. Springer International Publishing, Cham, 36–45. DOI: https://doi.org/10.1007/978-3-319-08795-5_5.

[242] Giuseppe Rizzo, Claudia d'Amato, and Nicola Fanizzi. 2015. On the effectiveness of evidence-based terminological decision trees. In *Foundations of Intelligent Systems*. Springer International Publishing, Cham, 139–149. DOI: https://doi.org/10.1007/978-3-319-25252-0_15.

[243] Giuseppe Rizzo, Claudia d'Amato, Nicola Fanizzi, and Floriana Esposito. 2015. Inductive classification through evidence-based models and their ensembles. In *The Semantic Web. Latest Advances and New Domains*. Springer International Publishing, Cham, 418–433. DOI: https://doi.org/10.1007/978-3-319-18818-8_26.

[244] Amandine Bellenger, Sylvain Gatepaille, Habib Abdulrab, and Jean-Philippe Kotowicz. 2011. An Evidential Approach for Modeling and Reasoning on Uncertainty in Semantic Applications. *7th International Workshop on Uncertainty Reasoning for the Semantic Web*, 27–38. https://api.semanticscholar.org/CorpusID:17173199.

[245] Ling Zhao and Guanyu Li. 2018. Abnomal activity identification method based on ontology and dempster-shafer theory. In *2018 10th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*. Volume 01, 289–292. DOI: 10.1109/IHMSC.2018.00073.

[246] Dongmei Li, Youfang Lin, Houkuan Huang, Shudong Hao, and Jianxin Wang. 2014. Dempster-shafer inconsistency values. *Chinese Journal of Electronics*, 23, (April 2014), 227–231. https://api.semanticscholar.org/CorpusID:117952908.

[247] Marguerite McDaniel, Emma Sloan, Siobahn Day, James Mayes, Albert Esterline, Kaushik Roy, and William Nick. 2017. Situation-based ontologies for a computational framework for identity focusing on crime scenes. In *2017 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, 1–7. DOI: 10.1109/COGSIMA.2017.7929579.

[248] Albert Esterline, William Nick, and Janelle Mason. 2018. Situations, information, and evidence (late breaking report). In *2018 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, 171–173. DOI: 10.1109/COGSIMA.2018.8423970.

[249] Janelle MASON, Kofi KYEI, Darrion LONG, Hannah FOSTER, William NICK, James MAYES, and Albert ESTERLINE. 2018. A computational framework for identity based on situation theory. In *International Symposium on Affective Science and Engineering ISASE2018*. Japan Society of Kansei Engineering, 1–7. DOI: https://doi.org/10.5057/isase.2018-C000027.

[250] Janelle Mason, Prosenjit Chatterjee, Kaushik Roy, and Albert Esterline. 2019. A framework to identify people in unstructured environments incorporating biometrics. In *Security, Privacy, and Anonymity in Computation, Communication, and Storage*. Guojun Wang, Jun Feng, Md Zakirul Alam Bhuiyan, and Rongxing Lu, editors. Springer International Publishing, Cham, 65–75. DOI: 10.1007/978-3-030-24900-7_5.

[251] Guangquan Xu, Xiaohong Li, Zhiyong Feng, and Yingyuan Xiao. 2010. Ontology driven formal description of swift trust and its computational model. In *2010 Second World Congress on Software Engineering*. Volume 2, 126–129. DOI: 10.1109/WCSE.2010.111.

[252] Vasanth Sarathy and Matthias Scheutz. 2018. A logic-based computational framework for inferring cognitive affordances. *IEEE Transactions on Cognitive and Developmental Systems*, 10, 1, 26–43. DOI: 10.1109/TCDS.2016.2615326.

[253] William Nick, Yenny Dominguez, and Albert C. Esterline. 2016. Situations and evidence for identity using dempster-shafer theory. In *Midwest Artificial Intelligence and Cognitive Science Conference*. https://api.semanticscholar.org/CorpusID:9999607.

[254] Emma Sloan, Marguerite McDaniel, William Nick, James Mayes, and Albert Esterline. 2016. Structure and evidence in identity cases. In *2016 IEEE Symposium Series on Computational Intelligence, SSCI 2016*. Athens, Greece, 1–8. DOI: 10.1109/SSCI.2016.7849916.

[255] Thomas Reineking, Niclas Schult, and Joana Hois. 2011. Combining statistical and symbolic reasoning for active scene categorization. In *Knowledge Discovery, Knowlege Engineering and Knowledge Management*. Ana Fred, Jan L. G. Dietz, Kecheng Liu, and Joaquim Filipe, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 262–275. DOI: https://doi.org/10.1007/978-3-642-19032-2_20.

[256] Tingting Zou, Shuai Lv, and Lei Liu. 2012. Rough description logic programs. *Journal of Computers*, 7, (November 2012). DOI: 10.4304/jcp.7.11.2719-2725.

[257] L. Liu, Tingting Zou, and S. Lv. 2011. Rough description logics with concrete domain. *ICIC Express Letters, Part B: Applications*, 2, (October 2011), 1143–1149. https://api.semanticscholar.org/CorpusID:118299923.

[258] A.F. Kana and Babatunde Akinkunmi. 2014. An algebra of ontologies approximation under uncertainty. *International Journal of Intelligence Science*, 4, (January 2014). DOI: 10.4236/ijis.2014.42007.

[259] C. Maria Keet. 2010. On the feasibility of description logic knowledge bases with rough concepts and vague instances. In *Proceedings of the 23rd International Workshop on Description Logics (DL 2010), Waterloo, Ontario, Canada, May 4-7, 2010* (CEUR Workshop Proceedings). Volker Haarslev, David Toman, and Grant E. Weddell, editors. Volume 573. CEUR-WS.org. https://ceur-ws.org/Vol-573/paper_9.pdf.

[260] Beniamino Murgante and Francesco Scorza. 2011. Ontology and spatial planning. In *Computational Science and Its Applications - ICCSA 2011*. Beniamino Murgante, Osvaldo Gervasi, Andrés Iglesias, David Taniar, and Bernady O. Apduhan, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 255–264. DOI: 10.1007/978-3-642-21887-3_20.

[261] Lian Yu, Shan Luo, and Arne Glenstrup. 2010. Rough sets based context-aware service discovery framework. In *2010 International Conference on Service Sciences*, 167–172. DOI: 10.1109/ICSS.2010.84.

[262] Nenad Dragun and Nikola Bogunović. 2019. Extending a communication service provider's operations management based on knowledge granules and aspects derived from imperfect information. *International Journal of Communication Systems*, 32, 14, e4076. DOI: https://doi.org/10.1002/dac.4076. https://onlinelibrary.wiley.com/doi/abs/10.1002/dac.4076.

[263] Zhang Xiaowang, Zhiyong Feng, Wenrui Wu, Xin Wang, and Guozheng Rao. 2018. PROSE: a plugin-based framework for paraconsistent reasoning on semantic web. *International Journal on Semantic Web and Information Systems*, 14, (January 2018), 121–139. DOI: 10.4018/IJSWIS.2018010105.

[264] Liping Zhou, Houkuan Huang, Guilin Qi, Yue Ma, Zhisheng Huang, and Youli Qu. 2012. Paraconsistent query answering over DL-Lite ontologies. *Web Intelligence and Agent Systems: An International Journal*, 10, 19–31. DOI: 10.3233/WIA-2012-0228.

[265] Linh Anh Nguyen and Andrzej Szałas. 2010. Three-valued paraconsistent reasoning for semantic web agents. In *Agent and Multi-Agent Systems: Technologies and Applications*. Springer Berlin Heidelberg, Berlin, Heidelberg, 152–162. DOI: https://doi.org/10.1007/978-3-642-13480-7_17.

[266] Wenrui Wu, Zhiyong Feng, Xiaowang Zhang, Xin Wang, and Guozheng Rao. 2016. PROSE: a plugin-based paraconsistent OWL reasoner. In *Semantic Technology*. Guilin Qi, Kouji Kozaki, Jeff Z. Pan, and Siwei Yu, editors. Springer International Publishing, Cham, 255–270. isbn: 978-3-319-31676-5.

[267] Xiaowang Zhang, Kewen Wang, Zhe Wang, Yue Ma, and Guilin Qi. 2015. A distance-based paraconsistent semantics for DL-Lite. In *Knowledge Science, Engineering and Management*. Songmao Zhang, Martin Wirsing, and Zili Zhang, editors. Springer International Publishing, Cham, 15–27. isbn: 978-3-319-25159-2. doi: 10.1007/978-3-319-25159-2_2.

[268] Xiaowang Zhang, Zhihu Zhang, Dai Xu, and Zuoquan Lin. 2010. Argumentation-based reasoning with inconsistent knowledge bases. In *Advances in Artificial Intelligence*. Atefeh Farzindar and Vlado Kešelj, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 87–99. doi: 10.1007/978-3-642-13059-5_11.

[269] Frederick Maier. 2010. Extending paraconsistent SROIQ. In *Web Reasoning and Rule Systems*. Pascal Hitzler and Thomas Lukasiewicz, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 118–132. isbn: 978-3-642-15918-3. doi: https://doi.org/10.1007/978-3-642-15918-3_10.

[270] Tobias Kaminski, Matthias Knorr, and João Leite. 2015. Well-founded paraconsistent semantics for hybrid theories composed of rules and ontologies. In *JOWO@IJCAI*. https://api.semanticscholar.org/CorpusID:5829935.

[271] Tobias Kaminski, Matthias Knorr, and João Leite. 2015. Efficient paraconsistent reasoning with ontologies and rules. In *Proceedings of the 24th International Conference on Artificial Intelligence* (IJCAI'15). AAAI Press, Buenos Aires, Argentina, 3098–3105. https://api.semanticscholar.org/CorpusID:15472244.

[272] Michael Fink. 2012. Paraconsistent hybrid theories. In *Thirteenth International Conference on the Principles of Knowledge Representation and Reasoning*. https://api.semanticscholar.org/CorpusID:18740800.

[273] D. Li, H. Wu, J. Gao, Z. Liu, L. Li, and Z. Zheng. 2019. Uncertain knowledge reasoning based on the fuzzy multi entity bayesian networks. *Computers, Materials and Continua*, 61, 1, 301–321. doi: https://doi.org/10.32604/cmc.2019.05953.

[274] Loukia Karanikola and Isambo Karali. 2016. A fuzzy logic approach for reasoning under uncertainty and vagueness - a matchmaking case study. In *2016 2nd International Conference on Information Management (ICIM)*, 52–56. doi: 10.1109/INFOMAN.2016.7477533.

[275] D. Nilavu and R. Sivakumar. 2015. Knowledge representation using type-2 fuzzy rough ontologies in ontology web language. *Fuzzy Information and Engineering*, 7, 1, 73–99. issn: 1616-8658. doi: https://doi.org/10.1016/j.fiae.2015.03.006.

[276] Ishak Riali, Messaouda Fareh, Mohamed Chakib Ibnaissa, and Mounir Bellil. 2023. A semantic-based approach for hepatitis c virus prediction and diagnosis using a fuzzy ontology and a fuzzy bayesian network. *Journal of Intelligent & Fuzzy Systems*, 44, 2381–2395. doi: 10.3233/jifs-213563.

[277] Zhi-Yun Zheng, Zhuo-Yun Liu, Lun Li, Dun Li, and Zhen-Fei Wang. 2016. A fuzzy MEBN ontology language based on OWL2. In *Fuzzy Systems and Data Mining II*. IOS Press, 71–80. doi: 10.3233/978-1-61499-722-1-71.

[278] Guanyu Li, Jiahong Yan, and Yanhan Liu. 2015. Research on the semantic web fuzzy rough ontology reasoner. *Journal of Computational Information Systems*, 11, 12, 4353 –4366. doi: 10.12733/jcis14643.

[279] Ishak Riali, Messaouda Fareh, and Hafida Bouarfa. 2019. Fuzzy Probabilistic Ontology Approach: A Hybrid Model for Handling Uncertain Knowledge in Ontologies. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 15, 4, (October 2019), 1–20. https://ideas.repec.org/a/igg/jswis0/v15y2019i4p1-20.html.

[280] Matteo Caglioni and Giovanni Fusco. 2014. Formal ontologies and uncertainty. in geographical knowledge. *TeMA - Journal of Land Use, Mobility and Environment*, (May 2014). doi: 10.6092/1970-9870/2530.

[281] Zhen Lei, Peizhi Cui, and Yanyan Huang. 2020. Tactical intention recognition based on fuzzy multi-entity bayesian network. In *2020 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS)*. IEEE, 965–968. doi: 10.1109/ICPICS50287.2020.9202082.

[282] Guan Yu Li, Yan Zhao, and Hai Yan Li. 2012. Imprecise ontology merging framework design. In *Advances in Science and Engineering II* (Applied Mechanics and Materials). Volume 135. Trans Tech Publications Ltd, (January 2012), 578–584. doi: 10.4028/www.scientific.net/AMM.135-136.578.

[283] Henrique Viana, Thiago Alves Rocha, João FL Alcântara, and Ana Teresa C Martins. 2010. A tractable paraconsistent fuzzy description logic. In *URSW*, 101–104. https://api.semanticscholar.org/CorpusID:10566596.

[284] Yinghui Huang and Guanyu Li. 2010. Imprecise semantic web ontology model. In *2010 International Conference on E-Business and E-Government*, 1426–1429. doi: 10.1109/ICEE.2010.362.

[285] Lucas Christian. 2012. Multi-criteria modelling and clustering of spatial information. *International Journal of Geographical Information Science*, 26, 10, 1897–1915. doi: 10.1080/13658816.2012.661056.

[286] Ishak Riali, Messaouda Fareh, and Hafida Bouarfa. 2019. A semantic approach for handling probabilistic knowledge of fuzzy ontologies. In *International Conference on Enterprise Information Systems*. https://api.semanticscholar.org/CorpusID:174804135.

[287] Maria Vargas-Vera and Miklos Nagy. 2014. Establishing agent trust for contradictory evidence by means of fuzzy voting model: an ontology mapping case study. *Computers in Human Behavior*, 30, 745–752. doi: https://doi.org/10.1016/j.chb.2013.10.032. https://www.sciencedirect.com/science/article/pii/S0747563213003828.

[288] Ahmad Adee, Peter Munk, Roman Gansch, and Peter Liggesmeyer. 2020. Uncertainty representation with extended evidential networks for modeling safety of the intended functionality (sotif). In *30th European Safety and Reliability Conference, ESREL 2020 and 15th Probabilistic Safety Assessment and Management Conference, PSAM 2020*. Venice, Virtual, Italy, 4148 –4155. doi: 10.3850/978-981-14-8593-0_5737-cd.

[289] Anna Formica. 2012. Semantic web search based on rough sets and fuzzy formal concept analysis. *Knowledge-Based Systems*, 26, 40–47. issn: 0950-7051. doi: https://doi.org/10.1016/j.knosys.2011.06.018.

[290] Tingting Zou, Yanpeng Qu, and Ansheng Deng. 2014. Tightly coupled fuzzy rough description logic programs under the answer set semantics for the semantic web. In *IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2014, Beijing, China, July 6-11, 2014*. IEEE, 1480–1487. DOI: 10.1109/FUZZ-IEEE.2014.6891558.

[291] Pavel Klinov and Bijan Parsia. 2013. Understanding a probabilistic description logic via connections to first-order logic of probability. In *Uncertainty Reasoning for the Semantic Web*. https://api.semanticscholar.org/CorpusID:23516431.

[292] Marcelo Finger, Renata Wassermann, and Fábio Gagliardi Cozman. 2011. Satisfiability in EL with sets of probabilistic aboxes. In *Proceedings of the 24th International Workshop on Description Logics (DL 2011), Barcelona, Spain, July 13-16, 2011* (CEUR Workshop Proceedings). Volume 745. CEUR-WS.org. https://api.semanticscholar.org/CorpusID:17868050.

[293] Rodrigo B. Polastro, Fabio G. Cozman, Felipe I. Takiyama, and Kate C. Revoredo. 2012. Computing inferences for credal ALC terminologies. In *Proceedings of the 8th International Conference on Uncertainty Reasoning for the Semantic Web - Volume 900* (URSW'12). CEUR-WS.org, Aachen, DEU, 94–97. https://api.semanticscholar.org/CorpusID:11219059.

[294] Jan Noessner and Mathias Niepert. 2011. ELOG: A probabilistic reasoner for OWL EL. In *Web Reasoning and Rule Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, 281–286. DOI: 10.1007/978-3-642-23580-1_25.

[295] Fabrizio Riguzzi, Elena Bellodi, Evelina Lamma, and Riccardo Zese. 2012. Semantics and inference for probabilistic ontologies. In *PAI*, 41–46. https://api.semanticscholar.org/CorpusID:15406594.

[296] Giuseppe Cota, Fabrizio Riguzzi, Riccardo Zese, Elena Bellodi, and Evelina Lamma. 2018. A modular inference system for probabilistic description logics. In *Scalable Uncertainty Management*. Springer International Publishing, Cham, 78–92. DOI: https://doi.org/10.1007/978-3-030-00461-3_6.

[297] Fabrizio Riguzzi, Elena Bellodi, Evelina Lamma, and Riccardo Zese. 2013. BUNDLE: A reasoner for probabilistic ontologies. In *Web Reasoning and Rule Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, 183–197. DOI: https://doi.org/10.1007/978-3-642-39666-3_14.

[298] Riccardo Zese, Elena Bellodi, Evelina Lamma, Fabrizio Riguzzi, and Fabiano Aguiari. 2014. Semantics and inference for probabilistic description logics. In *Uncertainty Reasoning for the Semantic Web III*. Springer International Publishing, Cham, 79–99. ISBN: 978-3-319-13413-0. DOI: 10.1007/978-3-319-13413-0_5.

[299] Fabrizio Riguzzi, Elena Bellodi, Evelina Lamma, and Riccardo Zese. 2015. Reasoning with probabilistic ontologies. In *IJCAI*, 4310–4316. https://api.semanticscholar.org/CorpusID:17702296.

[300] RICCARDO ZESE, GIUSEPPE COTA, EVELINA LAMMA, ELENA BELLODI, and FABRIZIO RIGUZZI. 2019. Probabilistic DL reasoning with pinpointing formulas: a prolog-based approach. *Theory and Practice of Logic Programming*, 19, 3, 449–476. DOI: 10.1017/S1471068418000480.

[301] Giuseppe Cota, Riccardo Zese, Elena Bellodi, Evelina Lamma, and Fabrizio Riguzzi. 2020. A framework for reasoning on probabilistic description logics. In *Applications and Practices in Ontology Design, Extraction, and Reasoning*. IOS Press, 127–144. https://api.semanticscholar.org/CorpusID:222124856.

[302] Li Weigang, Alessandro F. Leite, Vitor F. Ribeiro, Jose A. Fregnani, and Italo R. de Oliveira. 2017. Towards intelligent system wide information management for air traffic management. In *Security, Privacy, and Anonymity in Computation, Communication, and Storage*. Springer International Publishing, Cham, 584–593. DOI: https://doi.org/10.1007/978-3-319-72389-1_46.

[303] Joerg Schoenfisch. 2014. Querying probabilistic ontologies with SPARQL. In *GI-Jahrestagung*. https://api.semanticscholar.org/CorpusID:4863694.

[304] Andreas Ecke, Rafael Peñaloza, and Anni-Yasmin Turhan. 2014. Completion-based generalization inferences for the description logic ELOR with subjective probabilities. *International Journal of Approximate Reasoning*, 55, 9, 1939–1970. DOI: https://doi.org/10.1016/j.ijar.2014.03.001.

[305] R. Ferrari, H. Dibowski, and S. Baldi. 2017. A message passing algorithm for automatic synthesis of probabilistic fault detectors from building automation ontologies. *IFAC-PapersOnLine*, 50, 1, 4184–4190. DOI: https://doi.org/10.1016/j.ifacol.2017.08.809. https://www.sciencedirect.com/science/article/pii/S2405896317312636.

[306] Oliver Gries and Ralf Möller. 2010. Gibbs sampling in probabilistic description logics with deterministic dependencies. In *UniDL*. https://api.semanticscholar.org/CorpusID:11730989.

[307] Axel Bodart, Keyvin Evrard, James Ortiz, and Pierre-Yves Schobbens. 2014. ArThUR: a tool for markov logic network. In *On the Move to Meaningful Internet Systems: OTM 2014 Workshops*. Springer Berlin Heidelberg, Berlin, Heidelberg, 319–328. DOI: https://doi.org/10.1007/978-3-662-45550-0_32.

[308] Melisachew Wudage Chekol, Jakob Huber, Christian Meilicke, and Heiner Stuckenschmidt. 2016. Markov logic networks with numerical constraints. In *Proceedings of the Twenty-Second European Conference on Artificial Intelligence* (ECAI'16). IOS Press, NLD, 1017–1025. ISBN: 9781614996712. DOI: 10.3233/978-1-61499-672-9-1017.

[309] Denis D. Mauá and Fabio G. Cozman. 2015. DL-Lite bayesian networks: a tractable probabilistic graphical model. In *Scalable Uncertainty Management*. Springer International Publishing, Cham, 50–64. DOI: https://doi.org/10.1007/978-3-319-23540-0_4.

[310] Thomas Lukasiewicz, Maria Vanina Martinez, Giorgio Orsi, and Gerardo I. Simari. 2012. Heuristic ranking in tightly coupled probabilistic description logics. In *Uncertainty in Artificial Intelligence - Proceedings of the 28th Conference, UAI 2012*, 554 –563. DOI: https://doi.org/10.48550/arXiv.1210.4894.

[311] İsmail İlkan Ceylan. 2015. Query answering in bayesian description logics. In *Proceedings of the 28th International Workshop on Description Logics (DL 2015)*. Volume 1350. CEUR Workshop Proceedings. https://api.semanticscholar.org/CorpusID:8071162.

[312] Ismail Ilkan Ceylan and Rafael Penaloza. 2015. Probabilistic query answering in the bayesian description logic BEL. In *Scalable Uncertainty Management*. Springer International Publishing, Cham, 21–35. DOI: https://doi.org/10.1007/978-3-319-23540-0_2.

[313] İsmail İlkan Ceylan and Rafael Peñaloza. 2014. Reasoning in the description logic BEL using bayesian networks. In *Proceedings of the 4th International Workshop on Statistical Relational AI (starAI 2014)* (AAAI Workshops). Volume WS-14-13. AAAI Press. https://api.semanticscholar.org/CorpusID:1523574.

[314] Laécio L dos Santos, Rommel N Carvalho, Marcelo Ladeira, Li Weigang, and Gilson L Mendes. 2015. PR-OWL 2 RL - a language for scalable uncertainty reasoning on the semantic web. In *11th International Workshop on Uncertainty Reasoning for the Semantic Web (URSW 2015)*, 14. https://api.semanticscholar.org/CorpusID:18297442.

[315] Rafael Peñaloza and Nico Potyka. 2016. Probabilistic reasoning in the description logic ALCP with the principle of maximum entropy. In *Scalable Uncertainty Management*. Springer International Publishing, Cham, 246–259. DOI: https://doi.org/10.1007/978-3-319-45856-4_17.

[316] Rafael Penaloza and Nico Potyka. 2017. Towards statistical reasoning in description logics over finite domains. In *Scalable Uncertainty Management*. Springer International Publishing, Cham, 280–294. DOI: https://doi.org/10.1007/978-3-319-67582-4_20.

[317] Yuqicheng Zhu, Nico Potyka, Bo Xiong, Trung-Kien Tran, Mojtaba Nayyeri, Steffen Staab, and Evgeny Kharlamov. 2023. Towards statistical reasoning with ontology embeddings. In *ISWC (Posters/Demos/Industry)*. https://api.semanticscholar.org/CorpusID:267502021.

[318] Pavel Klinov and Bijan Parsia. 2011. Representing sampling distributions in P-SROIQ. In *Proceedings of the 7th International Conference on Uncertainty Reasoning for the Semantic Web - Volume 778* (URSW'11). CEUR-WS.org, Aachen, DEU, 39–50. https://api.semanticscholar.org/CorpusID:16590022.

[319] Pawel Kozak and Karsten Tolle. 2011. Analyzing web profiles using probabilistic ontologies. In *Proceeding of 3rd International Conference on Web Science*. https://api.semanticscholar.org/CorpusID:11344524.

[320] Kate Revoredo, José Eduardo Ochoa Luna, and Fabio Gagliardi Cozman. 2011. Semantic link prediction through probabilistic description logics. In *URSW*, 87–97. https://api.semanticscholar.org/CorpusID:18821736.

[321] Fabiano Corrêa, Fábio Gagliardi Cozman, and Jun Okamoto Jr. 2011. Collective classification in semantic mapping with a probabilistic description logic. In *Description Logics*. https://api.semanticscholar.org/CorpusID:14819591.

[322] Suwan Tongphu and Boontawee Suntisrivaraporn. 2012. Toward composite object classification using a probabilistic inference engine. In *Multi-disciplinary Trends in Artificial Intelligence*. Chattrakul Sombattheera, Nguyen Kim Loi, Rajeev Wankar, and Tho Quan, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 35–46. DOI: https://doi.org/10.1007/978-3-642-35455-7_4.

[323] Rim Helaoui, Daniele Riboni, and Heiner Stuckenschmidt. 2013. A probabilistic ontological framework for the recognition of multilevel human activities. *UbiComp 2013 - Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, (September 2013). DOI: 10.1145/2493432.2493501.

[324] Claudio Bettini, Gabriele Civitarese, Davide Giancane, and Riccardo Presotto. 2020. ProCAVIAR: hybrid data-driven and probabilistic knowledge-based activity recognition. *IEEE Access*, 8, 146876–146886. DOI: 10.1109/ACCESS.2020.3015091.

[325] Elena Bellodi, Evelina Lamma, Fabrizio Riguzzi, Riccardo Zese, and Giuseppe Cota. 2017. A web system for reasoning with probabilistic OWL. *Software: Practice and Experience*, 47, 1, 125–142. DOI: https://doi.org/10.1002/spe.2410. https://onlinelibrary.wiley.com/doi/abs/10.1002/spe.2410.

[326] Erman Acar. 2014. Computing subjective expected utility using probabilistic description logics. In *STAIRS*, 21–30. DOI: 10.3233/978-1-61499-421-3-21.

[327] 2019. *Explanation-friendly query answering under uncertainty. Reasoning Web. Explainable Artificial Intelligence: 15th International Summer School 2019, Bolzano, Italy, September 20–24, 2019, Tutorial Lectures*. Springer International Publishing, Cham, 65–103. DOI: 10.1007/978-3-030-31423-1_2.

[328] Georg Gottlob, Thomas Lukasiewicz, Maria Vanina Martinez, and Gerardo I. Simari. 2013. Query answering under probabilistic uncertainty in datalog +/− ontologies. *Annals of Mathematics and Artificial Intelligence*, 69, 1, 37–72. https://doi.org/10.1007/s10472-013-9342-1.

[329] Christophe Feltus, Djamel Khadraoui, and Jocelyn Aubert. 2010. A security decision-reaction architecture for heterogeneous distributed network. In *2010 International Conference on Availability, Reliability and Security*, 1–8. DOI: 10.1109/ARES.2010.57.

[330] Stephen. C. Dinkel, William Hafner, Paulo C. G. Costa, and Sumitra Mukherjee. 2013. Service-based situational awareness on the semantic web. In *2013 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, 246–253. DOI: 10.1109/CogSIMA.2013.6523854.

[331] Álvaro Carrera and Carlos A. Iglesias. 2012. Improving diagnosis agents with hybrid hypotheses confirmation reasoning techniques. In *Multi-Agent Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, 48–62. DOI: https://doi.org/10.1007/978-3-642-34799-3_4.

[332] Spiros Nikolopoulos, Georgios Th. Papadopoulos, Ioannis Kompatsiaris, and Ioannis Patras. 2011. Evidence-driven image interpretation by combining implicit and explicit knowledge in a bayesian network. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 41, 5, 1366–1381. DOI: 10.1109/TSMCB.2011.2147781.

[333] Spiros Nikolopoulos, Georgios Th. Papadopoulos, Ioannis Kompatsiaris, and Ioannis Patras. 2012. Image interpretation by combining ontologies and bayesian networks. In *Artificial Intelligence: Theories and Applications*. Springer Berlin Heidelberg, Berlin, Heidelberg, 307–314. DOI: https://doi.org/10.1007/978-3-642-30448-4_39.

[334] F. Maurelli, Z. A. Saigol, G. Papadimitriou, T. Larkworthy, V. De Carolis, and D.M. Lane. 2013. Probabilistic approaches in ontologies: joining semantics and uncertainty for AUV persistent autonomy. In *2013 OCEANS - San Diego*, 1–6. DOI: 10.23919/OCEANS.2013.6741183.

[335] Chia-Li Kuo, David Buchman, Arzoo Katiyar, and David Poole. 2013. Probabilistic reasoning with undefined properties in ontologically-based belief networks. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence* (IJCAI '13). AAAI Press, Beijing, China, 2532–2539. ISBN: 9781577356332. DOI: 10.14288/1.0052177.

[336] Chia-Li Kuo and David Poole. 2013. On integrating ontologies with relational probabilistic models. *AAAI Workshop - Technical Report*, (January 2013), 18–24. https://api.semanticscholar.org/CorpusID:1525925.

[337] Abdelweheb Gueddes and Mohamed Ali Mahjoub. 2022. Combining logical and probabilistic reasoning to improve a home care platform. In *2022 8th International Conference on Control, Decision and Information Technologies (CoDIT)*. Volume 1, 373–378. doi: 10.1109/CoDIT55151.2022.9803979.

[338] Abdelweheb Gueddes and Mohamed Ali Mahjoub. 2020. Enhancing ontology-based home care services platform using bayesian networks. In *2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 1310–1316. doi: 10.1109/ICMLA51294.2020.00205.

[339] Zhen Lei, Peizhi Cui, and Yanyan Huang. 2020. Research on tactical intention recognition method based on probabilistic network ontology language. In *2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications(AEECA)*, 957–960. doi: 10.1109/AEECA49918.2020.9213589.

[340] Shuai Wang, Yu Zhang, and Zhiyong Liao. 2019. Building domain-specific knowledge graph for unmanned combat vehicle decision making under uncertainty. In *2019 Chinese Automation Congress (CAC)*, 4718–4721. doi: 10.1109/CAC48633.2019.8996418.

[341] Shrinivasan Patnaikuni and Sachin Gengaje. 2019. Syntactico-semantic reasoning using PCFG, MEBN & PP attachment ambiguity. In *2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*. Volume 1. IEEE, 1–7. doi: https://doi.org/10.48550/arXiv.1809.07607.

[342] Márcio J. da Silva, Carlos E. Pereira, and Marcelo Götz. 2019. On the use of inductive biases for semantic characterization of industrial alarm systems. In *2019 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)*, 1–6. doi: 10.1109/ISGT-LA.2019.8894963.

[343] Márcio J. da Silva, Carlos E. Pereira, and Marcelo Götz. 2018. Context-aware recommendation for industrial alarm system. *IFAC-PapersOnLine*, 51, 11, 229–234. 16th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2018. issn: 2405-8963. doi: https://doi.org/10.1016/j.ifacol.2018.08.266.

[344] Hongfei Yao, Chunsong Han, and Fengxia Xu. 2022. Reasoning methods of unmanned underwater vehicle situation awareness based on ontology and bayesian network. *Complexity*, 2022, (April 2022), 1–10. doi: 10.1155/2022/7143974.

[345] Jose Aguilar, José Torres, and Kristell Aguilar. 2016. Autonomie decision making based on bayesian networks and ontologies. In *2016 International Joint Conference on Neural Networks (IJCNN)*, 3825–3832. doi: 10.1109/IJCNN.2016.7727693.

[346] Suvarna Kadam and Madhuri Potey. 2014. Probabilistic ontology for social background verification: a belief fusion use case. In *International Conference on Information Communication and Embedded Systems (ICICES2014)*, 1–7. doi: 10.1109/ICICES.2014.7033823.

[347] Xin Li, José-Fernán Martínez, Gregorio Rubio, and David Gómez. 2017. Context reasoning in underwater robots using MEBN. *CoRR*, abs/1706.07204. doi: https://doi.org/10.48550/arXiv.1706.07204.

[348] Xin Li, Sonia Bilbao, Tamara Martín-Wanton, Joaquim Bastos, and Jonathan Rodriguez. 2017. SWARMs Ontology: a common information model for the cooperation of underwater robots. *Sensors (Basel)*, 17. doi: 10.3390/s17030569.

[349] Obed Jules, Abdelhakim Hafid, and Mohamed Adel Serhani. 2014. Bayesian network, and probabilistic ontology driven trust model for sla management of cloud services. In *2014 IEEE 3rd International Conference on Cloud Networking (CloudNet)*, 77–83. doi: 10.1109/CloudNet.2014.6968972.

[350] Laécio L Santos, Rommel N Carvalho, Marcelo Ladeira, and Li Weigang. 2014. Probabilistic ontologies incremental modeling using UnBBayes. *Proceedings of encontro nacional de inteligência artificial e computacional, ENIAC*, 266–271.

[351] Rommel N. Carvalho, Shou Matsumoto, Kathryn B. Laskey, Paulo C. G. Costa, Marcelo Ladeira, and Laécio L. Santos. 2013. Probabilistic ontology and knowledge fusion for procurement fraud detection in brazil. In *Uncertainty Reasoning for the Semantic Web II*. Springer Berlin Heidelberg, Berlin, Heidelberg, 19–40. doi: https://doi.org/10.1007/978-3-642-35975-0_2.

[352] Rommel N. Carvalho, Richard Haberlin, Paulo Cesar G. Costa, Kathryn B. Laskey, and K C Chang. 2011. Modeling a probabilistic ontology for maritime domain awareness. In *14th International Conference on Information Fusion*, 1–8.

[353] Henrique Costa Marques, José M. P. de Oliveira, and Paulo Cesar G. da Costa. 2011. COA modeling with probabilistic ontologies. In *Proceedings of the Sixth International Conference on Semantic Technologies for Intelligence, Defense, and Security, Fairfax, VA, USA, November 16-17, 2011* (CEUR Workshop Proceedings). Paulo Cesar G. da Costa and Kathryn B. Laskey, editors. Volume 808. CEUR-WS.org, 60–67. https://ceur-ws.org/Vol-808/STIDS2011_CR_T8_MarquesEtAl.pdf.

[354] Kathryn B. Laskey, Richard Haberlin, Rommel N. Carvalho, and Paulo C. G. Costa. 2011. PR-OWL 2 case study: a maritime domain probabilistic ontology. In *Semantic Technologies for Intelligence, Defense, and Security*. https://api.semanticscholar.org/CorpusID:15853353.

[355] Rommel Carvalho, KB Laskey, Paulo Costa, Marcelo Ladeira, Laécio Santos, and Shou Matsumoto. 2010. UnBBayes: modeling uncertainty for plausible reasoning in the semantic web. *Semantic web*, 953–978. doi: 10.5772/7315.

[356] Rommel N. Carvalho, Paulo Cesar G. Costa, Kathryn B. Laskey, and K C Chang. 2010. PROGNOS: predictive situational awareness with probabilistic ontologies. In *2010 13th International Conference on Information Fusion*, 1–8. doi: 10.1109/ICIF.2010.5711970.

[357] Zhongli Wang, Guohui Tian, and Xuyang Shao. 2020. Home service robot task planning using semantic knowledge and probabilistic inference. *Knowledge-Based Systems*, 204, 106174. issn: 0950-7051. doi: https://doi.org/10.1016/j.knosys.2020.106174. https://www.sciencedirect.com/science/article/pii/S0950705120304093.

[358] Laura Antanas, Plinio Moreno, Marion Neumann, Rui Pimentel de Figueiredo, Kristian Kersting, José Santos-Victor, and Luc De Raedt. 2019. Semantic and geometric reasoning for robotic grasping: a probabilistic logic approach. *Autonomous Robots*, 43, 6, 1393–1418. https://doi.org/10.1007/s10514-018-9784-8.

[359] Hayden Freedman, Neda Abolhassani, Jacob Metzger, and Sanjoy Paul. 2023. Ontology modeling for probabilistic knowledge graphs. In *2023 IEEE 17th International Conference on Semantic Computing (ICSC)*, 252–259. DOI: 10.1109/ICSC56153.2023.00049.

[360] Messaouda Fareh, Ishak Riali, Hafsa Kherbache, and Marwa Guemmouz. 2023. Probabilistic reasoning for diagnosis prediction of coronavirus disease based on probabilistic ontology. *Computer Science and Information Systems*, 00, 35–35. DOI: 10.2298/CSIS220829035F.

[361] Luca Arrotta, Gabriele Civitarese, and Claudio Bettini. 2023. Probabilistic knowledge infusion through symbolic features for context-aware activity recognition. *Pervasive and Mobile Computing*, 91, 101780. ISSN: 1574-1192. DOI: https://doi.org/10.1016/j.pmcj.2023.101780.

[362] Erman Acar, Camilo Thorne, and Heiner Stuckenschmidt. 2015. Towards decision making via expressive probabilistic ontologies. In *Algorithmic Decision Theory*. Springer International Publishing, Cham, 52–68. DOI: https://doi.org/10.1007/978-3-319-23114-3_4.

[363] Stefan Ultes and Wolfgang Minker. 2014. Managing adaptive spoken dialogue for intelligent environments. *J. Ambient Intell. Smart Environ.*, 6, 523–539. DOI: 10.3233/AIS-140275.

[364] Mounir Hemam. 2018. An extension of the ontology web language with multi-viewpoints and probabilistic reasoning. *International Journal of Advanced Intelligence Paradigms*, 10, 3, (January 2018), 247–265. ISSN: 1755-0386. DOI: 10.1504/IJAIP.2018.090789. https://doi.org/10.1504/IJAIP.2018.090789.

[365] Anupama Mallik, Hiranmay Ghosh, Santanu Chaudhury, and Gaurav Harit. 2013. MOWL: an ontology representation language for web-based multimedia applications. *ACM Trans. Multimedia Comput. Commun. Appl.*, 10, 1, (December 2013). ISSN: 1551-6857. DOI: 10.1145/2542205.2542210. https://doi.org/10.1145/2542205.2542210.

[366] Abdul-Wahid Mohammed, Yang Xu, Ming Liu, and Brighter Agyemang. 2016. Uncertainty reasoning for smart homes: an ontological decision network based approach. *Web Intelligence*, 14, 199–210. DOI: 10.3233/WEB-160339.

[367] Fei Long, Yu Feng Zhang, and Feng Hu. 2011. A personalized recommender systems framework based on ontology and bayesian network in E-Commerce. *Advanced Materials Research*, 143, 961–965. DOI: 10.4028/www.scientific.net/AMR.143-144.961.

[368] Soon Chong Johnson Lim, Ying Liu, and Han Tong Loh. 2012. An exploratory study of ontology-based platform analysis under user preference uncertainty. In *Proceedings of the ASME Design Engineering Technical Conference*. Volume 2, 519–528. DOI: 10.1115/DETC2012-70756.

[369] Hu Xiao and Pan Shouhui. 2013. Reasoning model of quality and safety accidents of consumer products based on bayesian network. *Journal of Convergence Information Technology*, 8, 5, 229 –34. ISSN: 1975-9320. DOI: 10.4156/JCIT.VOL8.ISSUE5.27.

[370] Hayden Freedman, Jacob Metzger, Neda Abolhassani, Ana Tudor, Bill Tomlinson, and Sanjoy Paul. 2024. A bayesian approach to constructing probabilistic models from knowledge graphs. *International Journal of Semantic Computing*, 18, 01, 25–49. DOI: 10.1142/S1793351X24410022.

[371] Rim Helaoui, Daniele Riboni, Mathias Niepert, Claudio Bettini, and Heiner Stuckenschmidt. 2012. Towards activity recognition using probabilistic description logics. In *AAAI Conference on Artificial Intelligence*. https://api.semanticscholar.org/CorpusID:9535734.

[372] Elena Bellodi, Evelina Lamma, Fabrizio Riguzzi, and Simone Albani. 2011. A distribution semantics for probabilistic ontologies. In *Proceedings of the 7th International Conference on Uncertainty Reasoning for the Semantic Web* (URSW'11). Volume 778. CEUR-WS.org, Aachen, DEU, 75–86. https://api.semanticscholar.org/CorpusID:12164829.

[373] Fabrizio Riguzzi, Elena Bellodi, Evelina Lamma, and Riccardo Zese. 2012. Epistemic and statistical probabilistic ontologies. In *Proceedings of the 8th International Conference on Uncertainty Reasoning for the Semantic Web* (URSW'12). Volume 900. CEUR-WS.org, Aachen, DEU, 3–14. https://api.semanticscholar.org/CorpusID:1203726.

[374] Fabrizio Riguzzi, Elena Bellodi, and Evelina Lamma. 2012. Probabilistic datalog+/-under the distribution semantics. *CEUR Workshop Proceedings*, 846, (January 2012). https://api.semanticscholar.org/CorpusID:348381.

[375] Carsten Lutz and Lutz Schröder. 2010. Probabilistic description logics for subjective uncertainty. In *Twelfth International Conference on the Principles of Knowledge Representation and Reasoning*. (June 2010). DOI: 10.1613/jair.5222.

[376] Georg Gottlob, Thomas Lukasiewicz, and Gerardo I. Simari. 2011. Answering threshold queries in probabilistic datalog+/– ontologies. In *Scalable Uncertainty Management*. Springer Berlin Heidelberg, Berlin, Heidelberg, 401–414. DOI: https://doi.org/10.1007/978-3-642-23963-2_31.

[377] Octavian Arsene, Ioan Dumitrache, and Ioana Mihu. 2010. Collaborative medical software agents. *IFAC Proceedings Volumes*, 43, 22, 34–39. 3rd IFAC Workshop on Intelligent Control Systems. ISSN: 1474-6670. DOI: https://doi.org/10.3182/20100929-3-RO-4017.00007.

[378] Octavian Arsene, Ioan Dumitrache, and Ioana Mihu. 2011. Medicine expert system dynamic bayesian network and ontology based. *Expert Systems with Applications*, 38, 12, 15253–15261. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2011.05.074.

[379] Stephen. C. Dinkel, William Hafner, Paulo Costa, and Sumitra Mukherjee. 2011. Uncertainty reasoning for service-based situational awareness information on the semantic web. In *2011 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, 102–105. DOI: 10.1109/COGSIMA.2011.5753425.

[380] Alvaro Carrera and Carlos A. Iglesias. 2011. Multi-agent architecture for heterogeneous reasoning under uncertainty combining MSBN and ontologies in distributed network diagnosis. In *2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*. Volume 2. IEEE, 159–162. DOI: 10.1109/WI-IAT.2011.106.

[381] David Ausín, Federico Castanedo, and Diego López-de Ipiña. 2012. TURAMBAR: an approach to deal with uncertainty in semantic environments. In *Ambient Assisted Living and Home Care*. Springer Berlin Heidelberg, Berlin, Heidelberg, 329–337. DOI: https://doi.org/10.1007/978-3-642-35395-6_45.

[382] Leonard Botha, Thomas Meyer, and Rafael Peñaloza. 2020. The probabilistic description logic BALC. *ArXiv*, abs/2009.13407. DOI: https://doi.org/10.48550/arXiv.2009.13407.

[383] Ismail Ilkan Ceylan and Rafael Penaloza. 2014. The bayesian description logic BEL. In *Automated Reasoning*. Springer International Publishing, Cham, 480–494. DOI: https://doi.org/10.1007/978-3-319-08587-6_37.

[384] Ismail Ceylan, Rafael Penaloza, et al. 2014. Bayesian description logics. In *CEUR WORKSHOP PROCEEDINGS*. Volume 1193. CEUR, 447–458. https://api.semanticscholar.org/CorpusID:14266367.

[385] Emna Hlel, Salma Jamoussi, and Abdelmajid Ben Hamadou. 2015. A probabilistic ontology for the prediction of author's interests. In *Computational Collective Intelligence*. Springer International Publishing, Cham, 492–501. DOI: https://doi.org/10.1007/978-3-319-24306-1_48.

[386] Rommel N. Carvalho, Kathryn B. Laskey, and Paulo C. G. Costa. 2013. PR-OWL 2.0 –Bridging the gap to OWL semantics. In *Uncertainty Reasoning for the Semantic Web II*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1–18. DOI: https://doi.org/10.1007/978-3-642-35975-0_1.

[387] Xiangyang Li, Yung-Wen Liu, and Joseph Tan. 2011. MEDIA: a framework to reconfigure today's complex health care. *61st Annual IIE Conference and Expo Proceedings*, (January 2011). https://api.semanticscholar.org/CorpusID:114588988.

[388] Yawei Ge and Xiqian Hou. 2020. The probabilistic reasoning of hierarchical diagnosis decision-making. *Journal of Physics: Conference Series*, 1507, 10, (March 2020), 102016. DOI: 10.1088/1742-6596/1507/10/102016.

[389] Hlel Emna, Jamoussi Salma, Turki Mohamed, and Ben Hamadou Abdelmajid. 2016. Probabilistic ontology definition meta-model. In *Intelligent Decision Technologies 2016*. Springer International Publishing, Cham, 243–254. DOI: https://doi.org/10.1007/978-3-319-39630-9_20.

[390] Hlel Emna, Jamoussi Salma, Turki Mohamed, and Ben Hamadou Abdelmajid. 2017. Building probabilistic ontologies based on meta-model PODM. In *Proceedings of the 16th International Conference on Hybrid Intelligent Systems (HIS 2016)*. Springer International Publishing, Cham, 288–297. DOI: https://doi.org/10.1007/978-3-319-52941-7_29.

[391] Hoifung Poon and Pedro Domingos. 2010. Unsupervised ontology induction from text. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics* (ACL '10). Association for Computational Linguistics, USA, 296–305. https://api.semanticscholar.org/CorpusID:1473515.

[392] Xifeng Wang, Bixin Li, Li Liao, and Chunli Xie. 2011. Ontology-based reliability evaluation for web service. In *2011 IEEE 35th Annual Computer Software and Applications Conference*, 348–349. DOI: 10.1109/COMPSAC.2011.125.

[393] Xiaowei Zhang, Daniel Cao, Philip Moore, Jing Chen, Lin Zhou, Yang Zhou, and Xu Ma. 2012. A bayesian network (BN) based probabilistic solution to enhance emotional ontology. In *Human Centric Technology and Service in Smart Space*. Springer Netherlands, Dordrecht, 181–190. DOI: 10.1007/978-94-007-5086-9_24.

[394] Emil Vassev and Mike Hinchey. 2012. Knowledge representation with knowlang the marXbot case study. In *2012 IEEE 11th International Conference on Cybernetic Intelligent Systems (CIS)*. IEEE, 18–23. DOI: 10.1109/CIS.2013.6782155.

[395] Jian Ming Guo, Zheng Wang, and Yan Ping Wang. 2013. Study on intelligent monitoring and controlling system in oil drilling optimization. In *Information Technology Applications in Industry* (Applied Mechanics and Materials). Volume 263. Trans Tech Publications Ltd, (February 2013), 1319–1322. DOI: 10.4028/www.scientific.net/AMM.263-266.1319.

[396] Rong Hua Cong. 2014. About the studying on the uncertainty of ontology on the bayesian network. *Applied Mechanics and Materials*, 513-517, 1717 –1721. https://api.semanticscholar.org/CorpusID:62164927.

[397] AT Mouenis, K Mohamed, and A Souhaib. 2014. Towards probabilistic ontology based on bayesian networks. *International Journal of Software and Web Sciences*, 1, 10, 102–106. https://api.semanticscholar.org/CorpusID:13197115.

[398] Sonika Malik and Sarika Jain. 2021. A bayesian network approach to handle uncertainty in web ontology language. *International Journal of Reasoning-based Intelligent Systems*, 13, (January 2021), 243. DOI: 10.1504/IJRIS.2021.118661.

[399] Maciej J. Swat, Pierre Grenon, and Sarala Wimalaratne. 2016. ProbOnto: ontology and knowledge base of probability distributions. *Bioinformatics*, 32, 17, (April 2016), 2719–2721. DOI: 10.1093/bioinformatics/btw170.

[400] Daniele Riboni, Timo Sztyler, Gabriele Civitarese, and Heiner Stuckenschmidt. 2016. Unsupervised recognition of interleaved activities of daily living through ontological and probabilistic reasoning. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. Association for Computing Machinery, New York, NY, USA, 1–12. DOI: 10.1145/2971648.2971691. https://doi.org/10.1145/2971648.2971691.

[401] Gabriele Civitarese, Claudio Bettini, Timo Sztyler, Daniele Riboni, and Heiner Stuckenschmidt. 2019. NewNECTAR: collaborative active learning for knowledge-based probabilistic activity recognition. *Pervasive and Mobile Computing*, 56, 88–105. ISSN: 1574-1192. DOI: https://doi.org/10.1016/j.pmcj.2019.04.006.

[402] Kasper Apajalahti, Eero Hyvönen, Juha Niiranen, and Vilho Räisänen. 2017. Combining ontological modelling and probabilistic reasoning for network management. *Journal of Ambient Intelligence and Smart Environments*, 9, (January 2017), 63–76. DOI: 10.3233/AIS-160419.

[403] Shangpu Jiang, Daniel Lowd, and Dejing Dou. 2012. Learning to refine an automatically extracted knowledge base using markov logic. In *2012 IEEE 12th International Conference on Data Mining*, 912–917. DOI: 10.1109/ICDM.2012.156.

[404] Lucas Drumond and Rosario Girardi. 2010. Extracting ontology concept hierarchies from text using markov logic. In *Proceedings of the 2010 ACM Symposium on Applied Computing* (SAC '10). Association for Computing Machinery, New York, NY, USA, 1354–1358. DOI: 10.1145/1774088.1774379.

[405] K.S. Gayathri, K.S. Easwarakumar, and Susan Elias. 2017. Probabilistic ontology based activity recognition in smart homes using markov logic network. *Knowledge-Based Systems*, 121, 173–184. ISSN: 0950-7051. DOI: https://doi.org/10.1016/j.knosys.2017.01.025.

[406] Sina Dami, Ahmad Abdollahzadeh Barforoush, and Hossein Shirazi. 2018. News events prediction using markov logic networks. *Journal of Information Science*, 44, 1, 91–109. DOI: 10.1177/0165551516673285.

[407] Jay Pujara, Hui Miao, Lise Getoor, and William Cohen. 2013. Knowledge graph identification. In *The Semantic Web – ISWC 2013*. Springer Berlin Heidelberg, Berlin, Heidelberg, 542–557. DOI: https://doi.org/10.1007/978-3-642-41335-3_34.

[408] Jay Pujara, Hui Miao, Lise Getoor, and William W Cohen. 2015. Using semantics and statistics to turn data into knowledge. *AI Magazine*, 36, 1, 65–74. DOI: https://doi.org/10.1609/aimag.v36i1.2568.

[409] Michael N Huhns, Marco G Valtorta, and Jingsong Wang. 2010. Design principles for ontological support of bayesian evidence management. In *Ontologies and Semantic Technologies for Intelligence*. IOS Press, 163–178. DOI: 10.3233/978-1-60750-581-5-163.

[410] Rafał Kozik, Michał Choraś, and Witold Hołubowicz. 2010. Fusion of bayesian and ontology approach applied to decision support system for critical infrastructures protection. In *Mobile Lightweight Wireless Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, 451–463. DOI: https://doi.org/10.1007/978-3-642-16644-0_39.

[411] Octavian Arsene, Ioan Dumitrache, and Ioana Mihu. 2015. Expert system for medicine diagnosis using software agents. *Expert Systems with Applications*, 42, 4, 1825–1834. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2014.10.026.

[412] Quratulain Rajput and Sajjad Haider. 2011. BNOSA: A bayesian network and ontology based semantic annotation framework. *Journal of Web Semantics*, 9, 2, 99–112. Provenance in the Semantic Web. ISSN: 1570-8268. DOI: https://doi.org/10.1016/j.websem.2011.04.002.

[413] Montassar Ben Messaoud, Philippe Leray, and Nahla Ben Amor. 2011. SemCaDo: a serendipitous strategy for learning causal bayesian networks using ontologies. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*. Springer Berlin Heidelberg, Berlin, Heidelberg, 182–193. DOI: 10.1016/j.knosys.2014.12.006.

[414] Giacomo Bucci, Valeriano Sandrucci, and Enrico Vicario. 2011. Ontologies and bayesian networks in medical diagnosis. *2011 44th Hawaii International Conference on System Sciences*, 1–8. https://api.semanticscholar.org/CorpusID:13963833.

[415] Álvaro Carrera, Carlos A. Iglesias, Javier García-Algarra, and Dušan Kolařík. 2014. A real-life application of multi-agent systems for fault diagnosis in the provision of an internet business service. *Journal of Network and Computer Applications*, 37, 146–154. ISSN: 1084-8045. DOI: https://doi.org/10.1016/j.jnca.2012.11.004.

[416] Shutaro Hashimoto, Takeo Tadono, Masahiko Onosato, Masahiro Hori, and Takashi Moriyama. 2011. A framework of ontology-based knowledge information processing for change detection in remote sensing data. In *2011 IEEE International Geoscience and Remote Sensing Symposium*, 3927–3930. DOI: 10.1109/IGARSS.2011.6050090.

[417] Tomislav Stipančić, Bojan Jerbić, and Petar Ćurković. 2011. A robot group control based on bayesian reasoning. In *Proceedings of the World Congress on Engineering 2011, WCE 2011 Vol II*. Volume 2, 1056–1060. https://api.semanticscholar.org/CorpusID:13887487.

[418] Tomislav Stipancic, Yoshimasa Ohmoto, Sara Akaoka Badssi, and Toyoaki Nishida. 2017. Computation mechanism for situated sentient robot. In *2017 Computing Conference*, 64–73. DOI: 10.1109/SAI.2017.8252082.

[419] Fabio G. Cozman, Rodrigo B. Polastro, Felipe I. Takiyama, and Kate C. Revoredo. 2013. Computing inferences for relational bayesian networks based on ALC constructs. In *Revised Selected Papers of the ISWC International Workshops on Uncertainty Reasoning for the Semantic Web III - Volume 8816*. Springer-Verlag, Berlin, Heidelberg, 21–40. ISBN: 9783319134123. DOI: https://doi.org/10.1007/978-3-319-13413-0_2.

[420] Kamran Farooq, Amir Hussain, Stephen Leslie, Chris Eckl, Calum MacRae, and Warner Slack. 2012. An ontology driven and bayesian network based cardiovascular decision support framework. In *Advances in Brain Inspired Cognitive Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, 31–41. DOI: https://doi.org/10.1007/978-3-642-31561-9_4.

[421] Hakan Bulu, Adil Alpkocak, and Pinar Balci. 2013. Uncertainty modeling for ontology-based mammography annotation with intelligent BI-RADS scoring. *Computers in Biology and Medicine*, 43, 4, 301–311. ISSN: 0010-4825. DOI: https://doi.org/10.1016/j.compbiomed.2013.01.001.

[422] Ani Grubišić, Slavomir Stankov, and Ivan Peraić. 2013. Ontology based approach to bayesian student model design. *Expert Systems with Applications*, 40, 13, 5363–5371. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2013.03.041.

[423] Moges Tesfaye, Jones Kevontrez, Feng Shaw, Witherell Paul, and Ameta Gaurav. 2022. On characterizing uncertainty sources in laser powder-bed fusion additive manufacturing models. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 8, 1, 011110. DOI: 10.1115/1.4052039. https://doi.org/10.1115/1.4052039.

[424] Qiran Chen, Wei-Chun Chou, and Zhoumeng Lin. 2022. Integration of toxicogenomics and physiologically based pharmacokinetic modeling in human health risk assessment of perfluorooctane sulfonate. *Environ Sci Technol*, 56, 6, (March 2022), 3623–3633. DOI: 10.1021/acs.est.1c06479.

[425] Mélanie Munch, Patrice Buche, Stéphane Dervaux, Juliette Dibie, Liliana Ibanescu, Cristina Manfredotti, Pierre-Henri Wuillemin, and Hélène Angellier-Coussy. 2022. Combining ontology and probabilistic models for the design of bio-based product transformation processes. *Expert Systems with Applications*, 203, 117406. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2022.117406.

[426] Hossain Kordestani, Roghayeh Mojarad, Abdelghani Chibani, Kamel Barkaoui, Yacine Amirat, and Wagdy Zahran. 2021. Extended hapicare: a telecare system with probabilistic diagnosis and self-adaptive treatment. *Expert Systems with Applications*, 186, 115749. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2021.115749.

[427] Hossain Kordestani, Roghayeh Mojarad, Abdelghani Chibani, Aomar Osmani, Yacine Amirat, Kamel Barkaoui, and Wagdy Zahran. 2019. Hapicare: a healthcare monitoring system with self-adaptive coaching using probabilistic reasoning. In *2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA)*, 1–8. DOI: 10.1109/AICCSA47632.2019.9035291.

[428] Willem Coetzer, Deshendran Moodley, and Aurona Gerber. 2017. A knowledge-based system for generating interaction networks from ecological data. *Data & Knowledge Engineering*, 112, 55–78. ISSN: 0169-023X. DOI: https://doi.org/10.1016/j.datak.2017.09.005.

[429] Alan M. Kalet, Jason N. Doctor, John H. Gennari, and Mark H. Phillips. 2017. Developing bayesian networks from a dependency-layered ontology: a proof-of-concept in radiation oncology. *Medical Physics*, 44, 8, 4350–4359. DOI: https://doi.org/10.1002/mp.12340.

[430] Baydaa Taha. 2016. An emergency unit support system to diagnose chronic heart failure embedded with SWRL and bayesian network. *International Journal of Advanced Computer Science and Applications*, 7, (July 2016). DOI: 10.14569/IJACSA.2016.070761.

[431] Lu Zhong-Jun, Li Guan-Yu, and Pan Ying. 2016. A method of meta-context ontology modeling and uncertainty reasoning in SWoT. In *2016 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*, 128–135. DOI: 10.1109/CyberC.2016.34.

[432] Jinhyuk Mok and Hongki Min. 2015. Ontology-based context-aware model by applying bayesian network. In *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*, 2660–2664. DOI: 10.1109/FSKD.2015.7382377.

[433] Zhiqiang Liu, Yila Su, Fei Wang, and Huimin Li. 2014. The research on reasoning of tourism information and knowledge based on semantic web. *Computer Modelling and New Technologies*, 18, 12, 1061 –1067. https://api.semanticscholar.org/CorpusID:61452229.

[434] Edgar Jembere, Sibusiso S. Xulu, and Matthew O. Adigun. 2010. Automatic ontology evolution in open and dynamic computing environments. In *Computational Collective Intelligence. Technologies and Applications*. Springer Berlin Heidelberg, Berlin, Heidelberg, 122–132. DOI: https://doi.org/10.1007/978-3-642-16696-9_14.

[435] Maurice Van Keulen and Mena B. Habib. 2011. Handling uncertainty in information extraction. *CEUR Workshop Proceedings*, 778, 109 –112. ISSN: 16130073. https://api.semanticscholar.org/CorpusID:36926815.

[436] Marco Sotomayor, Kewen Wang, Yidong Shen, and John Thornton. 2012. Probabilistic multi-context systems. In *The Semantic Web*. Springer Berlin Heidelberg, Berlin, Heidelberg, 366–375. DOI: https://doi.org/10.1007/978-3-642-29923-0_26.

[437] Warner Marzocchi and Thomas H Jordan. 2018. Experimental concepts for testing probabilistic earthquake forecasting and seismic hazard models. *Geophysical Journal International*, 215, 2, (July 2018), 780–798. https://doi.org/10.1093/gji/ggy276.

[438] Pouya Foudeh and Naomie Salim. 2011. Probabilistic ontologies and probabilistic ontology learning: significance and challenges. *2011 International Conference on Research and Innovation in Information Systems*, 1–4. https://api.semanticscholar.org/CorpusID:16779740.

[439] Clemens Dubslaff, Patrick Koopmann, and Anni-Yasmin Turhan. 2021. Enhancing probabilistic model checking with ontologies. *Formal Aspects of Computing*, 33, 6, 885–921. https://doi.org/10.1007/s00165-021-00549-0.

[440] Pedro Chahuara, François Portet, and Michel Vacher. 2013. Making context aware decision from uncertain information in a smart home: a markov logic network approach. In *Ambient Intelligence*. Springer International Publishing, Cham, 78–93. DOI: https://doi.org/10.1007/978-3-319-03647-2_6.

[441] Erman Acar, Rafael Penaloza, and Livia Predoiu. 2019. Making decisions over contextual ontologies. In *32nd International Workshop on Description Logics, DL 2019*, 1–12. https://api.semanticscholar.org/CorpusID:186206465.

[442] Laécio L. Santos, Rommel N. Carvalho, Marcelo Bronzo Ladeira, and Weigang Li. 2016. A new algorithm for generating situation-specific bayesian networks using bayes-ball method. In *URSW@ISWC*. https://api.semanticscholar.org/CorpusID:9074007.

[443] Thomas Scharrenbach, Claudia d'Amato, Nicola Fanizzi, Rolf Grütter, Bettina Waldvogel, and Abraham Bernstein. 2010. Unsupervised conflict-free ontology evolution without removing axioms. In *CEUR Workshop Proceedings*. University of Zurich. DOI: 10.5167/uzh-44854.

[444] Thomas Lukasiewicz, Maria Vanina Martinez, and Gerardo I. Simari. 2012. Consistent answers in probabilistic datalog+/- ontologies. In *Web Reasoning and Rule Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, 156–171. DOI: https://doi.org/10.1007/978-3-642-33203-6_12.

[445] James Mcglothlin and Latifur Khan. 2010. Materializing and persisting inferred and uncertain knowledge in RDF datasets. *Proceedings of the National Conference on Artificial Intelligence*, 3, (January 2010). DOI: 10.1609/aaai.v24i1.7522.

[446] Salima Benbernou and Mourad Ouziri. 2018. Query answering on uncertain big RDF data using apache spark framework. In *2018 IEEE International Conference on Big Data (Big Data)*, 4854–4860. DOI: 10.1109/BigData.2018.8621905.

[447] Abdelhamid Malki, Djamal Benslimane, Sidi-Mohamed Benslimane, Mahmoud Barhamgi, Mimoun Malki, Parisa Ghodous, and Khalil Drira. 2016. Data services with uncertain and correlated semantics. *World Wide Web*, 19, 1, 157–175. DOI: 10.1007/s11280-014-0317-x.

[448] Abdelhamid Malki, Mahmoud Barhamgi, Sidi-Mohamed Benslimane, Djamal Benslimane, and Mimoun Malki. 2015. Composing data services with uncertain semantics. *IEEE Transactions on Knowledge and Data Engineering*, 27, 4, 936–949. DOI: 10.1109/TKDE.2014.2359661.

[449] Richard Haberlin, Paulo Costa, and Kathryn Laskey. 2013. A reference architecture for probabilistic ontology development. In *CEUR Workshop Proceedings*. Volume 1097. (November 2013). https://api.semanticscholar.org/CorpusID:566876.

[450] Rommel N Carvalho, Kathryn B Laskey, and Paulo CG Da Costa. 2016. Uncertainty modeling process for semantic technology. *PeerJ Computer Science*, 2, e77. DOI: 10.7287/peerj.preprints.2045v1.

[451] Diego M. Azevedo, Marcelo Bronzo Ladeira, Laécio L. Santos, and Rommel N. Carvalho. 2016. Automatic generation of probabilistic ontologies from UMP-ST model. In *URSW@ISWC*. https://api.semanticscholar.org/CorpusID:6056885.

[452] Cheol Young Park, Kathryn Blackmond Laskey, and Paulo CG Costa. 2016. An extended maritime domain awareness probabilistic ontology derived from human-aided multi-entity bayesian networks learning. In *STIDS*, 28–36. https://api.semanticscholar.org/CorpusID:18470654.

[453] Richard Haberlin, PC Costa, and Kathryn B Laskey. 2014. Probabilistic ontology architecture for a terrorist identification decision support system. In *Proceedings of the nineteenth international command and control research and technology symposium*. https://api.semanticscholar.org/CorpusID:60327672.

[454] Rommel N. Carvalho, Laécio L. dos Santos, Marcelo Ladeira, Henrique A. da Rocha, and Gilson L. Mendes. 2014. UMP-ST Plug-in: documenting, maintaining and evolving probabilistic ontologies using UnBBayes framework. In *Uncertainty Reasoning for the Semantic Web III*. Springer International Publishing, Cham, 1–20. DOI: https://doi.org/10.1007/978-3-319-13413-0_1.

[455] Rommel N. Carvalho, Marcelo Ladeira, Rafael Mezzomo De Souza, Shou Matsumoto, Henrique A. Da Rocha, and Gilson Libório Mendes. 2013. UMP-ST plug-in: a tool for documenting, maintaining, and evolving probabilistic ontologies. In *Proceedings of the 9th International Conference on*

Uncertainty Reasoning for the Semantic Web - Volume 1073 (URSW'13). CEUR-WS.org, Aachen, DEU, 15–26. https://api.semanticscholar.org/CorpusID:1514231.

[456] Foni A. Setiawan, Eko K. Budiardjo, and Wahyu C. Wibowo. 2019. ByNowLife: a novel framework for OWL and bayesian network integration. *Information*, 10, 3. ISSN: 2078-2489. DOI: 10.3390/info10030095. https://www.mdpi.com/2078-2489/10/3/95.

[457] Abhijit Boruah and Shyamanta M. Hazarika. 2014. An MEBN framework as a dynamic firewall's knowledge flow architecture. In *2014 International Conference on Signal Processing and Integrated Networks (SPIN)*, 249–254. DOI: 10.1109/SPIN.2014.6776957.

[458] Pavel Klinov and Bijan Parsia. 2010. Relationships between probabilistic description and first-order logics. *CEUR Workshop Proceedings*, 613, (January 2010). https://api.semanticscholar.org/CorpusID:1755203.

[459] Luigi Bellomarini, Eleonora Laurenza, Emanuel Sallinger, and Evgeny Sherkhonov. 2020. Reasoning under uncertainty in knowledge graphs. In *Rules and Reasoning*. Víctor Gutiérrez-Basulto, Tomáš Kliegr, Ahmet Soylu, Martin Giese, and Dumitru Roman, editors. Springer International Publishing, Cham, 131–139. DOI: https://doi.org/10.1007/978-3-030-57977-7_9.

[460] Luigi Bellomarini, Eleonora Laurenza, Emanuel Sallinger, and Evgeny Sherkhonov. 2023. Swift markov logic for probabilistic reasoning on knowledge graphs. *Theory and Practice of Logic Programming*, 23, 3, 507–534. DOI: 0.48550/arXiv.2210.00283.

[461] Fabio G Cozman and Denis D Mauá. 2015. Specifying probabilistic relational models with description logics. *Proceedings of the XII encontro nacional de inteligência artificial e computacional (ENIAC)*. https://api.semanticscholar.org/CorpusID:102344808.

[462] Shenyong Zhang, Yi Sun, Yun Peng, and Xiaopu Wang. 2010. A practical tool for uncertainty in OWL ontologies. *Proceedings of the 10th IASTED International Conference on Artificial Intelligence and Applications, AIA 2010*, (January 2010). DOI: 10.2316/P.2010.674-007.

[463] Mouna BenIshak, Philippe Leray, and Nahla Ben Amor. 2011. A two-way approach for probabilistic graphical models structure learning and ontology enrichment. *KEOD 2011 - Proceedings of the International Conference on Knowledge Engineering and Ontology Development*, (January 2011), 189–194. DOI: 10.5220/0003634801890194.

[464] Mouna Ben Ishak, Philippe Leray, and Nahla Ben Amor. 2011. Ontology-based generation of object oriented bayesian networks. In *CEUR Workshop Proceedings*. Volume 818. (January 2011), 9–17. https://api.semanticscholar.org/CorpusID:1607289.

[465] Bellandi Andrea and Turini Franco. 2012. Mining bayesian networks out of ontologies. *Journal of Intelligent Information Systems*, 38, 2, 507–532. DOI: 10.1007/s10844-011-0165-4. https://doi.org/10.1007/s10844-011-0165-4.

[466] Stefan Fenz. 2012. An ontology-based approach for constructing bayesian networks. *Data & Knowledge Engineering*, 73, 73–88. ISSN: 0169-023X. DOI: https://doi.org/10.1016/j.datak.2011.12.001.

[467] Giovanni Pilato, Agnese Augello, Michele Missikoff, and Francesco Taglino. 2012. Integration of ontologies and bayesian networks for maritime situation awareness. In *2012 IEEE Sixth International Conference on Semantic Computing*. IEEE, 170–177. DOI: 10.1109/ICSC.2012.25.

[468] Yvonne Fischer and Jürgen Beyerer. 2013. Ontologies for probabilistic situation assessment in the maritime domain. In *2013 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*. IEEE, 102–105. DOI: 10.1109/CogSIMA.2013.6523830.

[469] L. Yan, C.H. Wei, W.H. Tang, and Q.H. Wu. 2013. Development of a novel asset management system for power transformers based on ontology. In *2013 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, 1–6. DOI: 10.1109/APPEEC.2013.6837182.

[470] H. Lin, W. H. Tang, T. Y. Ji, and Q. H. Wu. 2014. A novel approach to power transformer fault diagnosis based on ontology and bayesian network. In *2014 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, 1–6. DOI: 10.1109/APPEEC.2014.7066069.

[471] David Ausın, Diego López-de Ipina, and Federico Castanedo. 2014. A probabilistic OWL reasoner for intelligent environments. In *10th International Workshop on Uncertainty Reasoning for the Semantic Web (URSW 2014)*, 1. https://api.semanticscholar.org/CorpusID:15650257.

[472] Abdelweheb Gueddes and Mohamed Ali Mahjoub. 2022. A Jena API for combining ontologies and bayesian object-oriented networks. In *2022 8th International Conference on Control, Decision and Information Technologies (CoDIT)*. Volume 1, 355–360. DOI: 10.1109/CoDIT55151.2022.9804150.

[473] Messaouda Fareh. 2019. Modeling incomplete knowledge of semantic web using bayesian networks. *Applied Artificial Intelligence*, 33, 11, 1022–1034. DOI: 10.1080/08839514.2019.1661578.

[474] Yao Hongfei, Wang Hongjian, Lv Hongli, and Wang Ying. 2016. Research on situation awareness based on ontology for UUV. In *2016 IEEE International Conference on Mechatronics and Automation*, 2500–2506. DOI: 10.1109/ICMA.2016.7558959.

[475] Malak Khreis, Iyad Zaarour, and Hamid Mcheick. 2016. SMILE: smart monitoring intelligent learning engine: an ontology-based context-aware system for supporting patients subjected to severe emergencies. *International Journal of Healthcare Technology and Management*, 15, (January 2016), 194. DOI: 10.1504/IJHTM.2016.078346.

[476] Hamid Mcheick, Hoda Nasser, Mohamed Dbouk, and Ahmad Nasser. 2016. Stroke prediction context-aware health care system. In *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, 30–35. DOI: 10.1109/CHASE.2016.49.

[477] Hamid Mcheick, Malak Khreiss, Hala Sweidan, and Iyad Zaarour. 2015. PHEN: Parkinson helper emergency notification system using bayesian belief network. In *E-Technologies*. Springer International Publishing, Cham, 212–223. DOI: https://doi.org/10.1007/978-3-319-17957-5_14.

[478] Quan Liu, Ayeley Tchangani, Bernard Kamsu-Foguem, and François Pérèsx. 2015. Modelling a large scale system for risk assessment. In *2015 International Conference on Industrial Engineering and Systems Management (IESM)*, 203–208. DOI: 10.1109/IESM.2015.7380158.

[479] Abdul-Wahid Mohammed, Yang Xu, and Ming Liu. 2015. Ontology-based reasoning with uncertain context in a smart home: a decision network approach. In *2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*. Volume 1, 515–522. DOI: 10.1109/WI-IAT.2015.160.

[480] Thomas Lukasiewicz, Maria Vanina Martinez, and Gerardo I. Simari. 2013. Preference-based query answering in probabilistic datalog+/- ontologies. In *On the Move to Meaningful Internet Systems: OTM 2013 Conferences*. Springer Berlin Heidelberg, Berlin, Heidelberg, 501–518. DOI: https://doi.org/10.1007/978-3-642-41030-7_36.

[481] Thomas Lukasiewicz, Maria Vanina Martinez, Gerardo I. Simari, and Oana Tifrea-Marciuska. 2013. Query answering in datalog+/- ontologies under group preferences and probabilistic uncertainty. In *Current Trends in Web Engineering*. Springer International Publishing, Cham, 192–206. DOI: https://doi.org/10.1007/978-3-319-04244-2_18.

[482] Thomas Lukasiewicz, Maria Vanina Martinez, Gerardo I. Simari, and Oana Tifrea-Marciuska. 2013. Query answering in probabilistic datalog+/- ontologies under group preferences. In *2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*. Volume 1, 171–178. DOI: 10.1109/WI-IAT.2013.26.

[483] Thomas Lukasiewicz, Maria Vanina Martinez, Gerardo I. Simari, and Oana Tifrea-Marciuska. 2014. Ontology-based query answering with group preferences. *ACM Trans. Internet Techn.*, 14, 25:1–25:24. https://api.semanticscholar.org/CorpusID:8309952.

[484] Hlel Emna, Salma Jamoussi, and Abdelmajid Ben Hamadou. 2017. A new method for building probabilistic ontology (Prob-Ont). *International Journal of Information Technology and Web Engineering*, 12, (April 2017), 1–25. DOI: 10.4018/IJITWE.2017040101.

[485] José Eduardo Ochoa-Luna, Kate Revoredo, and Fábio Gagliardi Cozman. 2011. Learning probabilistic description logics: a framework and algorithms. In *Advances in Artificial Intelligence*. Springer Berlin Heidelberg, Berlin, Heidelberg, 28–39. DOI: https://doi.org/10.1007/978-3-642-25324-9_3.

[486] Fabrizio Riguzzi, Elena Bellodi, Evelina Lamma, Riccardo Zese, and Giuseppe Cota. 2014. Learning probabilistic description logics. In *Uncertainty Reasoning for the Semantic Web III*. Springer International Publishing, Cham, 63–78. DOI: https://doi.org/10.1007/978-3-319-13413-0_4.

[487] Shyi-Ming Chen, Yun-Hou Huang, Rung-Ching Chen, Szu-Wei Yang, and Tian-Wei Sheu. 2012. Using fuzzy reasoning techniques and the domain ontology for anti-diabetic drugs recommendation. In *Intelligent Information and Database Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, 125–135. ISBN: 978-3-642-28487-8. DOI: https://doi.org/10.1007/978-3-642-28487-8_13.

[488] Piotr Szwed. 2013. Application of fuzzy ontological reasoning in an implementation of medical guidelines. In *2013 6th International Conference on Human System Interactions (HSI)*. IEEE, 342–349. DOI: 10.1109/HSI.2013.6577845.

[489] Christer Carlsson, Matteo Brunelli, and József Mezei. 2013. A soft computing approach to mastering paper machines. In *2013 46th Hawaii International Conference on System Sciences*. IEEE, 1394–1401. DOI: 10.1109/HICSS.2013.61.

[490] Chkiwa Mounira, Jedidi Anis, and Faiez Gargouri. 2013. Handling uncertainty in semantic information retrieval process. In *Proceedings of the 9th International Conference on Uncertainty Reasoning for the Semantic Web-Volume 1073*, 29–33. DOI: 10.13140/2.1.3562.9129.

[491] Carmen De Maio, Giuseppe Fenza, Domenico Furno, Vincenzo Loia, and Sabrina Senatore. 2012. OWL-FC: an upper ontology for semantic modeling of fuzzy control. *Soft Computing*, 16, 1153–1164. DOI: https://doi.org/10.1007/s00500-011-0790-4.

[492] Fernando Bobillo, Miguel Delgado, and Juan Gómez-Romero. 2012. DeLorean: a reasoner for fuzzy OWL 2. *Expert Systems with Applications*, 39, 1, 258–272. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2011.07.016.

[493] Fernando Bobillo, Miguel Delgado, and Juan Gómez-Romero. 2013. Reasoning in fuzzy OWL 2 with DeLorean. In *Uncertainty Reasoning for the Semantic Web II*. Springer Berlin Heidelberg, Berlin, Heidelberg, 119–138. ISBN: 978-3-642-35975-0. DOI: https://doi.org/10.1007/978-3-642-35975-0_7.

[494] Fernando Bobillo, Miguel Delgado, and José César Sánchez–Sánchez. 2013. Parallel algorithms for fuzzy ontology reasoning. *IEEE Transactions on Fuzzy Systems*, 21, 4, 775–781. DOI: 10.1109/TFUZZ.2012.2230266.

[495] Silvia Calegari, Davide Ciucci, and Matteo Mondini. 2013. Linguistic-variable definition in OWL 2 - A Protégé plugin. In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development (IC3K 2013) - KEOD*. INSTICC. SciTePress, 269–277. DOI: 10.5220/0004548002690277.

[496] Aloui Amira and Grissa Amel. 2015. An extension of Protégé for an automatic fuzzy-ontology building using clustering and FCA. *International Journal of Computer Science and Information Technology*, 7, (April 2015), 13–25. DOI: 10.5121/ijcsit.2015.7202.

[497] Fernando Bobillo and Umberto Straccia. 2016. The fuzzy ontology reasoner fuzzyDL. *Knowledge-Based Systems*, 95, 12–34. ISSN: 0950-7051. DOI: https://doi.org/10.1016/j.knosys.2015.11.017. https://www.sciencedirect.com/science/article/pii/S0950705115004621.

[498] Fernando Bobillo. 2016. The role of crisp elements in fuzzy ontologies: the case of fuzzy OWL 2 EL. *IEEE Transactions on Fuzzy Systems*, 24, 5, 1193–1209. DOI: 10.1109/TFUZZ.2015.2505329.

[499] Fernando Bobillo and Umberto Straccia. 2018. Reasoning within fuzzy OWL 2 EL revisited. *Fuzzy Sets and Systems*, 351, 1–40. DOI: https://doi.org/10.1016/j.fss.2018.03.011.

[500] Stamatia Dasiopoulou, Ioannis Kompatsiaris, and Michael G. Strintzis. 2010. Investigating fuzzy DLs-based reasoning in semantic image analysis. *Multimedia Tools and Applications*, 49, 1, 167–194. DOI: 10.1007/s11042-009-0393-6.

[501] Dorothea Tsatsou, Stamatia Dasiopoulou, Ioannis Kompatsiaris, and Vasileios Mezaris. 2014. LiFR: a lightweight fuzzy DL reasoner. In *The Semantic Web: ESWC 2014 Satellite Events*. Springer International Publishing, Cham, 263–267. DOI: 10.1007/978-3-319-11955-7_32.

[502] Giorgos Stoilos and Giorgos Stamou. 2014. Reasoning with fuzzy extensions of OWL and OWL 2. *Knowledge and Information Systems*, 40, 1, 205–242. DOI: 10.1007/s10115-013-0641-y.

[503] Kamaluddeen Usman Danyaro, Jafreezal Jaafar, and Mohd Shahir Liew. 2012. Completeness knowledge representation in fuzzy description logics. In *Knowledge Technology*. Springer Berlin Heidelberg, Berlin, Heidelberg, 164–173. ISBN: 978-3-642-32826-8. DOI: https://doi.org/10.1007/978-3-642-32826-8_17.

[504] Theofilos Mailis, Rafael Peñaloza, and Anni-Yasmin Turhan. 2014. Conjunctive query answering in finitely-valued fuzzy description logics. In *Web Reasoning and Rule Systems*. Springer International Publishing, Cham, 124–139. DOI: https://doi.org/10.1007/978-3-319-11113-1_9.

[505] Stefan Borgwardt and Rafael Peñaloza. 2014. Consistency reasoning in lattice-based fuzzy description logics. *International Journal of Approximate Reasoning*, 55, 9, 1917–1938. DOI: https://doi.org/10.1016/j.ijar.2013.07.006.

[506] Céline Hudelot, Jamal Atif, and Isabelle Bloch. 2015. ALC(F): a new description logic for spatial reasoning in images. In *Computer Vision - ECCV 2014 Workshops*. Springer International Publishing, Cham, 370–384. DOI: https://doi.org/10.1007/978-3-319-16181-5_26.

[507] Stefan Borgwardt, Theofilos Mailis, Rafael Peñaloza, and Anni-Yasmin Turhan. 2016. Answering fuzzy conjunctive queries over finitely valued fuzzy ontologies. *Journal on Data Semantics*, 5, 2, 55–75. DOI: 10.1007/s13740-015-0055-y.

[508] Dorian Merz, Rafael Peñaloza, and Anni-Yasmin Turhan. 2014. Reasoning in ALC with fuzzy concrete domains. In *KI 2014: Advances in Artificial Intelligence*. Springer International Publishing, Cham, 171–182. ISBN: 978-3-319-11206-0. DOI: https://doi.org/10.1007/978-3-319-11206-0_17.

[509] Agnieszka Kułacka, Dirk Pattinson, and Lutz Schröder. 2013. Syntactic labelled tableaux for łukasiewicz fuzzy alc. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence* (IJCAI '13). AAAI Press, Beijing, China, 962–968. ISBN: 9781577356332.

[510] Theofilos Mailis and Anni-Yasmin Turhan. 2015. Employing DL_Lite_R -Reasoners for fuzzy query answering. In *Semantic Technology*. Springer International Publishing, Cham, 63–78. ISBN: 978-3-319-15615-6. DOI: https://doi.org/10.1007/978-3-319-15615-6_5.

[511] Stefan Borgwardt, José Galano, and Rafael Peñaloza. 2014. The fuzzy description logic G-FL_0 with greatest fixed-point semantics. In *Logics in Artificial Intelligence*. Springer International Publishing, Cham, 62–76. DOI: https://doi.org/10.1007/978-3-319-11558-0_5.

[512] Fernando Bobillo and Umberto Straccia. 2011. Reasoning with the finitely many-valued łukasiewicz fuzzy description logic SROIQ. *Information Sciences*, 181, 4, 758–778. ISSN: 0020-0255. DOI: https://doi.org/10.1016/j.ins.2010.10.020.

[513] Marco Cerami, Francesc Esteva, and Àngel García-Cerdaña. 2012. On finitely valued fuzzy description logics: the łukasiewicz case. In *Advances in Computational Intelligence*. Springer Berlin Heidelberg, Berlin, Heidelberg, 235–244. ISBN: 978-3-642-31715-6. DOI: https://doi.org/10.1007/978-3-642-31715-6_26.

[514] Cristiane A. Yaguinuma, Walter C. P. Magalhães, Marilde T. P. Santos, Heloisa A. Camargo, and Marek Reformat. 2014. Combining fuzzy ontology reasoning and mamdani fuzzy inference system with HyFOM reasoner. In *Enterprise Information Systems*. Springer International Publishing, Cham, 174–189. DOI: https://doi.org/10.1007/978-3-319-09492-2_11.

[515] Gabriella Pasi and Rafael Penaloza. 2023. Answering fuzzy queries over fuzzy DL-Lite ontologies. *Theory and Practice of Logic Programming*, 23, 3, 594–623. DOI: https://doi.org/10.48550/arXiv.2111.11779.

[516] Chunchen Liu, Dayou Liu, and Shengsheng Wang. 2010. Situation modeling and identifying under uncertainty. In *2010 Second Pacific-Asia Conference on Circuits, Communications and System*. Volume 1, 296–299. DOI: 10.1109/PACCS.2010.5626909.

[517] Mye Sohn, Sunghwan Jeong, and Hyun Jung Lee. 2014. Case-based context ontology construction using fuzzy set theory for personalized service in a smart home environment. *Soft Computing*, 18, 9, 1715–1728. DOI: 10.1007/s00500-014-1288-7.

[518] Houda Akremi, Sami Zghal, and Gayo Diallo. 2016. Modeling of uncertainty: fuzzification of medical ontology. In *Proceedings of the 6th International Conference on Web Intelligence, Mining and Semantics*, 1–4. DOI: 10.1145/2912845.2912879.

[519] Ahmad C. Bukhari and Yong-Gi Kim. 2011. Incorporation of fuzzy theory with heavyweight ontology and its application on vague information retrieval for decision making. *The International Journal of Fuzzy Logic and Intelligent Systems*, 11, 3, (September 2011), 171–177. http://www.ijfis.org/journal/view.html?doi=.

[520] Mateus Ferreira-Satler, Francisco P. Romero, Victor H. Menendez-Dominguez, Alfredo Zapata, and Manuel E. Prieto. 2012. Fuzzy ontologies-based user profiles applied to enhance e-learning activities. *Soft Computing*, 16, 7, 1129–1141. DOI: 10.1007/s00500-011-0788-y.

[521] Juan Antonio Morente Molinera, Ignacio Javier Pérez Gálvez, Robin Wikström, Enrique Herrera Viedma, and Christer Carlsson. 2015. Designing a decision support system for recommending smartphones using fuzzy ontologies. In *Intelligent Systems'2014*. Springer International Publishing, Cham, 323–334. DOI: https://doi.org/10.1007/978-3-319-11310-4_28.

[522] Ignacio Pérez, Robin Wikström, József Mezei, Christer Carlsson, Karina Anaya, and Enrique Herrera-Viedma. 2013. Linguistic consensus models based on a fuzzy ontology. *Procedia Computer Science*, 17, 498–505. DOI: https://doi.org/10.1016/j.procs.2013.05.064.

[523] C. Porcel, A. Ching-López, G. Lefranc, V. Loia, and E. Herrera-Viedma. 2018. Sharing notes: an academic social network based on a personalized fuzzy linguistic recommender system. *Engineering Applications of Artificial Intelligence*, 75, 1–10. DOI: https://doi.org/10.1016/j.engappai.2018.07.007.

[524] J.A. Morente-Molinera, R. Wikström, E. Herrera-Viedma, and C. Carlsson. 2016. A linguistic mobile decision support system based on fuzzy ontology to facilitate knowledge mobilization. *Decision Support Systems*, 81, 66–75. DOI: https://doi.org/10.1016/j.dss.2015.09.001.

[525] A.C.M. Fong, Baoyao Zhou, Siu Hui, Jie Tang, and Guan Hong. 2012. Generation of personalized ontology based on consumer emotion and behavior analysis. *IEEE Transactions on Affective Computing*, 3, 2, 152–164. DOI: 10.1109/T-AFFC.2011.22.

[526] Hichem Bannour and Céline Hudelot. 2014. Building and using fuzzy multimedia ontologies for semantic image annotation. *Multimedia Tools and Applications*, 72, 3, 2107–2141. DOI: 10.1007/s11042-013-1491-z.

[527] Natalia Díaz Rodríguez, Manuel P. Cuéllar, Johan Lilius, and Miguel Delgado Calvo-Flores. 2014. A fuzzy ontology for semantic modelling and recognition of human behaviour. *Knowledge-Based Systems*, 66, 46–60. DOI: https://doi.org/10.1016/j.knosys.2014.04.016.

[528] Shaker El-Sappagh, Mohammed Elmogy, and A.M. Riad. 2015. A fuzzy-ontology-oriented case-based reasoning framework for semantic diabetes diagnosis. *Artificial Intelligence in Medicine*, 65, 3, 179–208. DOI: https://doi.org/10.1016/j.artmed.2015.08.003.

[529] Heba Fouda, Mohammed Elmogy, Ahmed Aboelfetoh, and Abdel Raziq Maat. 2015. Constructing fuzzy ontology for cardiac arrhythmias. In *2015 Tenth International Conference on Computer Engineering & Systems (ICCES)*, 402–409. DOI: 10.1109/ICCES.2015.7393084.

[530] Juan Antonio Morente-Molinera, Ignacio J Pérez, M Raquel Ureña, and Enrique Herrera-Viedma. 2016. Creating knowledge databases for storing and sharing people knowledge automatically using group decision making and fuzzy ontologies. *Information Sciences*, 328, 418–434. DOI: https://doi.org/10.1016/j.ins.2015.08.051.

[531] Mohammad Elhefny, Mohammed Elmogy, Ahmed Aboelfetouh, and Farid Badria. 2017. Developing a fuzzy OWL ontology for obesity related cancer domain. *International Journal of Medical Engineering and Informatics*, 9, (January 2017), 162–187. DOI: 10.1504/IJMEI.2017.10002627.

[532] Shaker El-Sappagh and Mohammed Elmogy. 2017. A fuzzy ontology modeling for case base knowledge in diabetes mellitus domain. *Engineering Science and Technology, an International Journal*, 20, 3, 1025–1040. ISSN: 2215-0986. DOI: https://doi.org/10.1016/j.jestch.2017.03.009.

[533] Tommaso Di Noia, Marina Mongiello, Francesco Nocera, and Umberto Straccia. 2019. A fuzzy ontology-based approach for tool-supported decision making in architectural design. *Knowledge and Information Systems*, 58, 1, 83–112. DOI: 10.1007/s10115-018-1182-1.

[534] Allassane Issa Salaou and Abdelghani Ghomari. 2021. Fuzzy ontology-based complex and uncertain video surveillance events recognition. In *2021 International Conference on Information Systems and Advanced Technologies (ICISAT)*, 1–5. DOI: 10.1109/ICISAT54145.2021.9678479.

[535] Nora Shoaip, Amira Rezk, Shaker El-Sappagh, Louai Alarabi, Sherif Barakat, and Mohammed M. Elmogy. 2021. A comprehensive fuzzy ontology-based decision support system for alzheimer's disease diagnosis. *IEEE Access*, 9, 31350–31372. DOI: 10.1109/ACCESS.2020.3048435.

[536] Olaide N. Oyelade, Absalom E. Ezugwu, and Sunday A. Adewuyi. 2022. Enhancing reasoning through reduction of vagueness using fuzzy OWL-2 for representation of breast cancer ontologies. *Neural Computing and Applications*, 34, 4, 3053–3078. DOI: 10.1007/s00521-021-06517-2.

[537] Houda AKREMI, Mouhamed Gaith AYADI, and Sami Zghal. 2022. To medical ontology fuzzification purpose: COVID-19 study case. *Procedia Computer Science*, 207, 1027–1036. DOI: https://doi.org/10.1016/j.procs.2022.09.158.

[538] Ahmad C. Bukhari and Yong-Gi Kim. 2013. A research on an intelligent multipurpose fuzzy semantic enhanced 3D virtual reality simulator for complex maritime missions. *Applied Intelligence*, 38, 2, 193–209. DOI: 10.1007/s10489-012-0365-9.

[539] Tsung-Chih Lin, Zhen-Shu Mi, Ahmad C. Bukhari, and Yong-Gi Kim. 2014. An obstacle recognizing mechanism for autonomous underwater vehicles powered by fuzzy domain ontology and support vector machine. *Mathematical Problems in Engineering*, 676729. DOI: 10.1155/2014/676729.

[540] Farman Ali, Kyung-Sup Kwak, and Yong-Gi Kim. 2016. Opinion mining based on fuzzy domain ontology and support vector machine: a proposal to automate online review classification. *Applied Soft Computing*, 47, 235–250. ISSN: 1568-4946. DOI: https://doi.org/10.1016/j.asoc.2016.06.003.

[541] Ehsan Sharifi, Reza A. Moghadam, Fernando Bobillo, and Mohamad M. Ebadzadeh. 2011. A fuzzy framework for semantic web service description, matchmaking, ranking and selection. In *2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)*. Volume 1, 621–625. DOI: 10.1109/FSKD.2011.6019550.

[542] Houda Akremi and Sami Zghal. 2021. DOF: a generic approach of domain ontology fuzzification. *Frontiers of Computer Science*, 15, 1–12. DOI: https://doi.org/10.1007/s11704-020-9354-z.

[543] Firas Zekri, Rafik Bouaziz, and Emna Turki. 2015. A fuzzy-based ontology for alzheimer's disease decision support. In *2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 1–6. DOI: 10.1109/FUZZ-IEEE.2015.7337922.

[544] Lianlian He, Peng Yue, Liangcun Jiang, and Mingda Zhang. 2015. Fuzzy spatial relation ontology driven detection of complex geospatial features in a web service environment. *Earth Science Informatics*, 8, 1, 63–76. https://doi.org/10.1007/s12145-014-0186-y.

[545] Umberto Straccia. 2010. Softfacts: a top-k retrieval engine for ontology mediated access to relational databases. In *2010 IEEE International Conference on Systems, Man and Cybernetics*, 4115–4122. DOI: 10.1109/ICSMC.2010.5641780.

[546] Nicola Vitucci, Mario Arrigoni Neri, and Giuseppina Gini. 2010. Using f-SHIN to represent objects: an aid to visual grasping. *UniDL*, 10, 80–87. https://api.semanticscholar.org/CorpusID:12796608.

[547] Yuncheng Jiang, Yong Tang, Qimai Chen, and Ju Wang. 2011. Reasoning and change management in modular fuzzy ontologies. *Expert Systems with Applications*, 38, 11, 13975–13986. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2011.04.205.

[548] Chunchen Liu, Dayou Liu, and Shengsheng Wang. 2012. Fuzzy geospatial information modeling in geospatial semantic retrieval. *Adv. Math. Comput. Methods*, 2, 4, 47–53. DOI: 10.4156/jcit.vol6.issue6.10.

[549] Fernando Bobillo, Juan Gómez-Romero, and Pilar León Araúz. 2012. Fuzzy ontologies for specialized knowledge representation in wordnet. In *Advances on Computational Intelligence*. Salvatore Greco, Bernadette Bouchon-Meunier, Giulianella Coletti, Mario Fedrizzi, Benedetto Matarazzo, and Ronald R. Yager, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 430–439. ISBN: 978-3-642-31709-5. DOI: https://doi.org/10.1007/978-3-642-31709-5_44.

[550] Natalia Díaz-Rodríguez, Olmo León Cadahía, Manuel Pegalajar Cuéllar, Johan Lilius, and Miguel Delgado Calvo-Flores. 2014. Handling real-world context awareness, uncertainty and vagueness in real-time human activity tracking and recognition with a fuzzy ontology-based hybrid method. *Sensors*, 14, 10, 18131–18171. DOI: 10.3390/s141018131.

[551] J. Gómez-Romero, F. Bobillo, M. Ros, M. Molina-Solana, M.D. Ruiz, and M.J. Martín-Bautista. 2015. A fuzzy extension of the semantic building information model. *Automation in Construction*, 57, 202–212. DOI: https://doi.org/10.1016/j.autcon.2015.04.007.

[552] Tommaso Di Noia, Marina Mongiello, and Umberto Straccia. 2015. Fuzzy description logics for component selection in software design. In *Software Engineering and Formal Methods*. Springer Berlin Heidelberg, Berlin, Heidelberg, 228–239. DOI: 10.1007/978-3-662-49224-6_19.

[553] Piotr Szwed. 2016. Modeling and recognition of video events with fuzzy semantic petri nets. In *Knowledge, Information and Creativity Support Systems: Recent Trends, Advances and Solutions*. Andrzej M.J. Skulimowski and Janusz Kacprzyk, editors. Springer International Publishing, Cham, 507–518. DOI: https://doi.org/10.1007/978-3-319-19090-7_38.

[554] Hai Bang Truong, Ngoc Thanh Nguyen, and Phi Khu Nguyen. 2011. Fuzzy ontology building and integration for fuzzy inference systems in weather forecast domain. In *Intelligent Information and Database Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, 517–527. DOI: https://doi.org/10.1007/978-3-642-20039-7_52.

[555] Bicheng Zhu and Utpal Roy. 2013. Uncertain information representation and its usage in disassembly modeling. In *33rd Computers and Information in Engineering Conference* (Proceedings of the ASME Design Engineering Technical Conference). American Society of Mechanical Engineers. DOI: 10.1115/DETC2013-12477.

[556] James A. Rodger. 2013. A fuzzy linguistic ontology payoff method for aerospace real options valuation. *Expert Systems with Applications*, 40, 8, 2828–2840. DOI: https://doi.org/10.1016/j.eswa.2012.12.001.

[557] Juan Antonio Morente-Molinera, Ignacio Javier Pérez, Francisco Javier Cabrerizo, Carlos Porcel, and Enrique Herrera-Viedma. 2016. Improving queries and representing heterogeneous information in fuzzy ontologies using multi-granular fuzzy linguistic modelling methods. In *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 000290–000295. DOI: 10.1109/SMC.2016.7844256.

[558] Abolfazl Doostparast Torshizi, Mohammad Hossein Fazel Zarandi, Ghazaleh Doostparast Torshizi, and Kamyar Eghbali. 2014. A hybrid fuzzy-ontology based intelligent system to determine level of severity and treatment recommendation for benign prostatic hyperplasia. *Computer Methods and Programs in Biomedicine*, 113, 1, 301–313. ISSN: 0169-2607. DOI: https://doi.org/10.1016/j.cmpb.2013.09.021.

[559] Mohamed Said Hamani, Ramdane Maamri, Yacine Kissoum, and Maamar Sedrati. 2014. Unexpected rules using a conceptual distance based on fuzzy ontology. *Journal of King Saud University - Computer and Information Sciences*, 26, 1, 99–109. ISSN: 1319-1578. DOI: https://doi.org/10.1016/j.jksuci.2013.06.001.

[560] Vincenzo Loia. 2011. Fuzzy ontologies and fuzzy markup language: a novel vision in web intelligence. In *Advances in Intelligent Web Mastering – 3*. Springer Berlin Heidelberg, Berlin, Heidelberg, 3–10. DOI: https://doi.org/10.1007/978-3-642-18029-3_1.

[561] Sameera Al Shayji, Nahla El Zant El Kadhi, and Zidong Wong. 2011. On fuzzy-logic-based ontology decision support system for government sector. In *Proceedings of the 12th WSEAS International Conference on Neural Networks, Fuzzy Systems, Evolutionary Computing & Automation* (NNECFSIC'12). World Scientific, Engineering Academy, and Society (WSEAS), Stevens Point, Wisconsin, USA, 34–41. ISBN: 9789604742929. https://api.semanticscholar.org/CorpusID:38355149.

[562] 2012. *Semantic information retrieval based on RDF and fuzzy ontology for university scientific research management. Affective Computing and Intelligent Interaction*. Springer Berlin Heidelberg, Berlin, Heidelberg, 661–668. DOI: 10.1007/978-3-642-27866-2_81.

[563] Chang-Shing Lee and Mei-Hui Wang. 2011. A fuzzy expert system for diabetes decision support application. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 41, 1, 139–153. DOI: 10.1109/TSMCB.2010.2048899.

[564] Hanêne Ghorbel, Afef Bahri, and Rafik Bouaziz. 2010. Fuzzy ontologies model for semantic web. In *Proceedings of the International Conference on Information and Knowledge Management, St. Maarten, The Netherlands*.

[565] Ahmad C. Bukhari and Yong-Gi Kim. 2012. Integration of a secure type-2 fuzzy ontology with a multi-agent platform: a proposal to automate the personalized flight ticket booking domain. *Information Sciences*, 198, 24–47. ISSN: 0020-0255. DOI: https://doi.org/10.1016/j.ins.2012.02.036.

[566] Farman Ali, Eun Kyoung Kim, and Yong-Gi Kim. 2015. Type-2 fuzzy ontology-based semantic knowledge for collision avoidance of autonomous underwater vehicles. *Information Sciences*, 295, 441–464. DOI: https://doi.org/10.1016/j.ins.2014.10.013.

[567] Farman Ali, Eun Kyoung Kim, and Yong-Gi Kim. 2015. Type-2 fuzzy ontology-based opinion mining and information extraction: a proposal to automate the hotel reservation system. *Applied Intelligence*, 42, 3, 481–500. DOI: 10.1007/s10489-014-0609-y.

[568] Diksha Hooda and Rinkle Rani. 2022. An interval type-2 fuzzy ontological model: predicting water quality from sensory data. *Concurrency and Computation: Practice and Experience*, 34, 28. DOI: https://doi.org/10.1002/cpe.7377.

[569] Ignacio Huitzil, Lacramioara Dranca, Jorge Bernad, and Fernando Bobillo. 2019. Gait recognition using fuzzy ontologies and kinect sensor data. *International Journal of Approximate Reasoning*, 113, 354–371. DOI: https://doi.org/10.1016/j.ijar.2019.07.012.

[570] Francesca Lisi and Corrado Mencar. 2018. A granular computing method for OWL ontologies*. *Fundamenta Informaticae*, 159, (March 2018), 147–174. DOI: 10.3233/FI-2018-1661.

[571] Fernando Bobillo, Lacramioara Dranca, and Jorge Bernad. 2017. A fuzzy ontology-based system for gait recognition using kinect sensor. In *Scalable Uncertainty Management*. Springer International Publishing, Cham, 397–404. ISBN: 978-3-319-67582-4. DOI: https://doi.org/10.1007/978-3-319-67582-4_29.

[572] Hong Wang, Jian Li, Jing Wang, and Hongjie Fan. 2014. Research on fuzzy extension of domain ontology for emergency management of civil aviation. *The Journal of Information and Computational Science*, 11, 6019–6028. https://api.semanticscholar.org/CorpusID:61963284.

[573] Zhenglian Su, Haisong Chen, Liang Zhu, and Yonghua Zeng. 2012. Framework of semantic web service discovery based on fuzzy logic and multi-phase matching. *Journal of Information and Computational Science*, 9, 1, 203 –214. ISSN: 15487741. https://api.semanticscholar.org/CorpusID:58032365.

[574] Nouredine Tamani, Ludovic Liétard, and Daniel Rocacher. 2013. A fuzzy ontology for database querying with bipolar preferences. *International Journal of Intelligent Systems*, 28, 1, 4–36. DOI: https://doi.org/10.1002/int.21572.

[575] Yannick Naudet, Valentin Groues, and Muriel Foulonneau. 2010. Introduction to fuzzy-ontological context-aware recommendations in mobile environments. *CEUR Workshop Proceedings*, 585, 51 –57. ISSN: 16130073.

[576] Huawei Zhang and Kerui Chen. 2012. Building social relationship ontology model based on fuzzy set. *International Journal of Digital Content Technology and its Applications*, 6, 15, 459 –466. ISSN: 19759339. DOI: 10.4156/jdcta.vol6.issue15.52.

[577] G. Besbes, H. Baazaoui-Zghal, and H. Ben Ghezela. 2014. Fuzzy ontology-based system for personalized information retrieval. In *6th International Conference on Knowledge Discovery and Information Retrieval (KDIR 2014). Proceedings*. Setubal, Portugal, 286 –93. https://api.semanticscholar.org/CorpusID:9045414.

[578] 2010. *Compensatory fuzzy ontology. ICT Innovations 2009*. Springer Berlin Heidelberg, Berlin, Heidelberg, 35–44. DOI: https://doi.org/10.1007/978-3-642-10781-8_5.

[579] Fernando Bobillo and Umberto Straccia. 2010. Representing fuzzy ontologies in OWL 2. In *International Conference on Fuzzy Systems*. IEEE, 1–6. DOI: 10.1109/FUZZY.2010.5584661.

[580] Chunchen Liu, Liu Dayou, and Shengsheng Wang. 2011. A fuzzy geospatial ontology model and its application in geospatial semantic retrieval. *Journal of Convergence Information Technology*, 6, (June 2011), 85–97. DOI: 10.4156/jcit.vol6.issue6.10.

[581] Umberto Straccia. 2012. Description logics with fuzzy concrete domains. *Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence, UAI 2005*, 140, (July 2012). DOI: 10.3233/FI-2015-1259.

[582] Qingchao Dong, Zhixue Wang, Weixing Zhu, and Hongyue He. 2012. Capability requirements modeling and verification based on fuzzy ontology. *Journal of Systems Engineering and Electronics*, 23, 1, 78–87. DOI: 10.1109/JSEE.2012.00011.

[583] Piotr Szwed and Mateusz Komorkiewicz. 2013. Object tracking and video event recognition with fuzzy semantic petri nets. In *2013 Federated Conference on Computer Science and Information Systems*. IEEE, 167–174. DOI: 10.1007/978-3-319-02309-0_47.

[584] Piotr Szwed. 2014. Video event recognition with fuzzy semantic petri nets. In *Man-Machine Interactions 3*. Dr. Aleksandra Gruca, Tadeusz Czachórski, and Stanisław Kozielski, editors. Springer International Publishing, Cham, 431–439. DOI: https://doi.org/10.1007/978-3-319-02309-0_47.

[585] Zhang Tingting, Liu Xiaoming, Wang Zhixue, and Dong Qingchao. 2015. Capability-oriented architectural analysis method based on fuzzy description logic. *Computer Science and Information Systems*, 13, (January 2015), 46–46. DOI: 10.2298/CSIS150222046T.

[586] Fernando Bobillo, Ehsan Sharifi, Mohamad M. Ebadzadeh, and Reza A. Moghadam. 2010. On fuzzy semantic web services. In *International Conference on Fuzzy Systems*, 1–6. DOI: 10.1109/FUZZY.2010.5584162.

[587] ASM Kayes, Wenny Rahayu, Tharam Dillon, Elizabeth Chang, and Jun Han. 2017. Context-aware access control with imprecise context characterization through a combined fuzzy logic and ontology-based approach. In *On the Move to Meaningful Internet Systems. OTM 2017 Conferences: Confederated International Conferences: CoopIS, C&TC, and ODBASE 2017, Rhodes, Greece, October 23-27, 2017, Proceedings, Part I*. Springer, 132–153. DOI: https://doi.org/10.1007/978-3-319-69462-7_10.

[588] Cristiane A. Yaguinuma, Marilde T.P. Santos, Heloisa A. Camargo, and Tatiane M. Nogueira. 2010. A meta-ontology approach for representing vague linguistic terms and fuzzy rules for classification in ontologies. In *2010 14th IEEE International Enterprise Distributed Object Computing Conference Workshops*, 263–271. DOI: 10.1109/EDOCW.2010.41.

[589] Mateus Ferreira-Satler, Francisco P. Romero, Victor H. Menendez, Alfredo Zapata, and Manuel E. Prieto. 2010. A fuzzy ontology approach to represent user profiles in e-learning environments. In *International Conference on Fuzzy Systems*, 1–8. DOI: 10.1109/FUZZY.2010.5584789.

[590] Christer Carlsson, Matteo Brunelli, and Jozsef Mezei. 2010. Fuzzy ontology and information granulation: an approach to knowledge mobilisation. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Applications*. Springer Berlin Heidelberg, Berlin, Heidelberg, 420–429. DOI: https://doi.org/10.1007/978-3-642-14058-7_44.

[591] Christer Carlsson, Matteo Brunelli, and Jozsef Mezei. 2010. Fuzzy ontologies and knowledge mobilisation: turning amateurs into wine connoisseurs. In *International Conference on Fuzzy Systems*. IEEE, 1–7. DOI: 10.1109/FUZZY.2010.5583998.

[592] Carlos Porcel, Carmen Martinez-Cruz, Juan Bernabé-Moreno, Álvaro Tejeda-Lorente, and Enrique Herrera-Viedma. 2015. Integrating ontologies and fuzzy logic to represent user-trustworthiness in recommender systems. *Procedia Computer Science*, 55, 603–612. ISSN: 1877-0509. DOI: https://doi.org/10.1016/j.procs.2015.07.050.

[593] J.A. Morente-Molinera, I.J. Pérez, M.R. Ureña, and E. Herrera-Viedma. 2015. Building and managing fuzzy ontologies with heterogeneous linguistic information. *Knowledge-Based Systems*, 88, 154–164. DOI: https://doi.org/10.1016/j.knosys.2015.07.035.

[594] Francesca Alessandra Lisi and Corrado Mencar. 2015. Towards fuzzy granulation in OWL ontologies. In *Italian Conference on Computational Logic*. https://api.semanticscholar.org/CorpusID:1191966.

[595] Sameera Alshaiji, Nahla ElKadhi, Zidong Wang, and Fawaz S. Al-Anzi. 2011. Fuzzy-based ontology intelligent dss to strengthen government bilateral economic relations. In *Proceedings of the Second Kuwait Conference on E-Services and e-Systems* (KCESS '11). Association for Computing Machinery, New York, NY, USA. DOI: 10.1145/2107556.2107576.

[596] Cristiane A. Yaguinuma, Walter C. P. Magalhães, Marilde Terezinha Prado Santos, Heloisa A. Camargo, and Marek Z. Reformat. 2013. Hyfom reasoner: hybrid integration of fuzzy ontology and mamdani reasoning. In *International Conference on Enterprise Information Systems*. https://api.semanticscholar.org/CorpusID:20091124.

[597] Chang-Shing Lee, Mei-Hui Wang, Giovanni Acampora, Chin-Yuan Hsu, and Hani Hagras. 2010. Diet assessment based on type-2 fuzzy ontology and fuzzy markup language. *International Journal of Intelligent Systems*, 25, 12, 1187–1216. DOI: https://doi.org/10.1002/int.20449.

[598] Chang-Shing Lee, Mei-Hui Wang, Zhi-Wei Chen, Chin-Yuan Hsu, Su-E Kuo, Hui-Ching Kuo, Hui-Hua Cheng, and Akio Naito. 2011. Genetic fuzzy markup language for diet application. In *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)*, 1791–1798. DOI: 10.1109/FUZZY.2011.6007634.

[599] Chang-Shing Lee, Mei-Hui Wang, Hani Hagras, Zhi-Wei Chen, Shun-Teng Lan, Chin-Yuan Hsu, Su-E Kuo, Hui-Ching Kuo, and Hui-Hua Cheng. 2012. A novel genetic fuzzy markup language and its application to healthy diet assessment. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 20, (October 2012), 247–278. DOI: 10.1142/S0218488512400235.

[600] Chang-Shing Lee, Mei-Hui Wang, Min-Hsiang Wu, Chin-Yuan Hsu, Yi-Chen Lin, and Shi-Jim Yen. 2010. A type-2 fuzzy personal ontology for meeting scheduling system. In *International Conference on Fuzzy Systems*, 1–8. DOI: 10.1109/FUZZY.2010.5584517.

[601] Hsien-De Huang, Giovanni Acampora, Vincenzo Loia, Chang-Shing Lee, and Hung-Yu Kao. 2011. Applying FML and fuzzy ontologies to malware behavioural analysis. In *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)*, 2018–2025. DOI: 10.1109/FUZZY.2011.6007716.

[602] Zhu HongKang, Yu XueLi, Zhuo GuangPing, and Hu Kun. 2010. Research on services matching and ranking based on fuzzy QoS ontology. In *2010 International Conference on Computational Aspects of Social Networks*, 579–582. DOI: 10.1109/CASoN.2010.134.

[603] Tomasz Wiktor Wlodarczyk, Chunming Rong, Marting O'Connor, and Mark Musen. 2011. SWRL-F: a fuzzy logic extension of the semantic web rule language. In *Proceedings of the International Conference on Web Intelligence, Mining and Semantics* (WIMS '11). Association for Computing Machinery, New York, NY, USA. ISBN: 9781450301480. DOI: 10.1145/1988688.1988735.

[604] Jie Liu and Lin Cui. 2014. Extending description logic SHOIN based on cloud model. *Applied Mechanics and Materials*, 543-547, (March 2014), 3586–3589. DOI: 10.4028/www.scientific.net/AMM.543-547.3586.

[605] Alessandro Ciaramella, Mario G. C. A. Cimino, Francesco Marcelloni, and Umberto Straccia. 2010. Combining fuzzy logic and semantic web to enable situation-awareness in service recommendation. In *Database and Expert Systems Applications*. Pablo García Bringas, Abdelkader Hameurlain, and Gerald Quirchmayr, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 31–45. DOI: https://doi.org/10.1007/978-3-642-15364-8_3.

[606] Uraiwan Inyaem, Phayung Meesad, Choochart Haruechaiyasak, and Dat Tran. 2010. Construction of fuzzy ontology-based terrorism event extraction. In *2010 third international conference on knowledge discovery and data mining*. IEEE, 391–394. DOI: 10.1109/WKDD.2010.113.

[607] Mateus Ferreira-Satler, Francisco P. Romero, Jose A. Olivas, and Jesus Serrano-Guerrero. 2014. Fuzzy ontology-based approach for automatic construction of user profiles. In *Rough Sets and Current Trends in Computing*. Springer International Publishing, Cham, 339–346. ISBN: 978-3-319-08644-6. DOI: https://doi.org/10.1007/978-3-319-08644-6_35.

[608] R Wikström. 2013. Ontology of imprecision and fuzzy ontology applications. In *WEBIST 2013 - Proceedings of the 9th International Conference on Web Information Systems and Technologies*, 284–287. DOI: 10.5220/0004361002840287.

[609] Lien-Fu. Lai, Liang-Tsung Huang, Chao-Chin Wu, and Shi-Shan Chen. 2010. Fuzzy knowledge management through knowledge engineering and fuzzy logic. *Journal of Convergence Information Technology*, 5, 3, 7 –15. ISSN: 19759320. DOI: 10.4156/jcit.vol5.issue3.1.

[610] Sun Yi, Zhang Dezheng, and Cheng Li. 2010. Fuzzy ontology constructing and its application in traditional chinese medicine. In *2010 IEEE International Conference on Intelligent Computing and Intelligent Systems*. Volume 2, 592–595. DOI: 10.1109/ICICISYS.2010.5658410.

[611] Singh Aarti, Juneja Dimple, and Sharma Ashok. 2011. A fuzzy integrated ontology model to manage uncertainty in semantic web: the FIOM. *International Journal on Computer Science and Engineering*, 3, (February 2011). https://api.semanticscholar.org/CorpusID:7823152.

[612] Lizhen Liu, Xinhui Nie, and Hanshi Wang. 2012. Toward a fuzzy domain sentiment ontology tree for sentiment analysis. In *2012 5th International Congress on Image and Signal Processing*, 1620–1624. DOI: 10.1109/CISP.2012.6469930.

[613] Salisu Sani and Teh NM Aris. 2014. Proposal for ontology based approach to fuzzy student model design. In *2014 5th International Conference on Intelligent Systems, Modelling and Simulation*. IEEE, 35–37. DOI: 10.1109/ISMS.2014.14.

[614] Shital V. Chavan, S. S. Sambare, and Aniruddha Joshi. 2016. Diet recommendation based on prakriti and season using fuzzy ontology and type-2 fuzzy logic. In *2016 International Conference on Computing Communication Control and automation (ICCUBEA)*, 1–6. DOI: 10.1109/ICCUBEA.2016.7860026.

[615] Madiha Liaqat, Sharifullah Khan, and Muhammad Majid. 2016. Fuzzy ontology based model for image retrieval. In *Mobile Web and Intelligent Information Systems*. Springer International Publishing, Cham, 108–120. DOI: https://doi.org/10.1007/978-3-319-44215-0_9.

[616] Nizar Elleuch, Mohamed Zarka, Anis Ben Ammar, and Adel M. Alimi. 2011. A fuzzy ontology: based framework for reasoning in visual video content analysis and indexing. In *MDMKDD '11*. https://api.semanticscholar.org/CorpusID:11059404.

[617] Stasinos Konstantopoulos and Angelos Charalambidis. 2010. Formulating description logic learning as an inductive logic programming task. In *International Conference on Fuzzy Systems*, 1–7. DOI: 10.1109/FUZZY.2010.5584417.

[618] Josúe Iglesias and Jens Lehmann. 2011. Towards integrating fuzzy logic capabilities into an ontology-based inductive logic programming framework. In *2011 11th International Conference on Intelligent Systems Design and Applications*, 1323–1328. DOI: 10.1109/ISDA.2011.6121843.

[619] Z.M. Ma, Fu Zhang, Li Yan, and Jingwei Cheng. 2011. Representing and reasoning on fuzzy UML models: a description logic approach. *Expert Systems with Applications*, 38, 3, 2536–2549. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2010.08.042.

[620] Carmen Martinez-Cruz, Ignacio J. Blanco, and M. Amparo Vila. 2012. An ontology as a tool for representing fuzzy data in relational databases. *International Journal of Computational Intelligence Systems*, 5, 6, 1089–1108. DOI: 10.1080/18756891.2012.747663.

[621] Carmen Martinez-Cruz, Ignacio Blanco, and M. Amparo Vila. 2011. An ontology to represent queries in fuzzy relational databases. In *2011 11th International Conference on Intelligent Systems Design and Applications*, 1317–1322. DOI: 10.1109/ISDA.2011.6121842.

[622] Carmen Martínez-Cruz, Jose María Serrano, and Amparo Vila. 2017. A tool to perform semantic and imprecise queries on non-scalar data. In *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 1–6. DOI: 10.1109/FUZZ-IEEE.2017.8015693.

[623] Fernando Bobillo and Umberto Straccia. 2011. Aggregation operators and fuzzy OWL 2. In *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)*, 1727–1734. DOI: 10.1109/FUZZY.2011.6007406.

[624] Fernando Bobillo and Umberto Straccia. 2011. Fuzzy ontologies and fuzzy integrals. In *2011 11th International Conference on Intelligent Systems Design and Applications*. IEEE, 1311–1316. DOI: 10.1109/ISDA.2011.6121841.

[625] Fernando Bobillo and Umberto Straccia. 2013. Aggregation operators for fuzzy ontologies. *Applied Soft Computing*, 13, 9, 3816–3830. ISSN: 1568-4946. DOI: https://doi.org/10.1016/j.asoc.2013.05.008.

[626] Francesca Alessandra Lisi and Umberto Straccia. 2013. A system for learning GCI axioms in fuzzy description logics. In *Description Logics*. https://api.semanticscholar.org/CorpusID:11963015.

[627] Francesca A. Lisi and Umberto Straccia. 2014. A FOIL-Like method for learning under incompleteness and vagueness. In *Inductive Logic Programming*. Gerson Zaverucha, Vítor Santos Costa, and Aline Paes, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 123–139. DOI: https://doi.org/10.1007/978-3-662-44923-3_9.

[628] Francesca A. Lisi, Umberto Straccia, Domenico Cantone, Marianna Nicolosi Asmundo, and Alberto Pettorossi. 2015. Learning in description logics with fuzzy concrete domains. *Fundam. Inf.*, 140, 3–4, (September 2015), 373–391. DOI: 10.3233/FI-2015-1259.

[629] Oliver Fernández Gil, Franz Baader, and Gerhard Brewka. 2015. Adding threshold concepts to the description logic EL. In *Description Logics*. https://api.semanticscholar.org/CorpusID:10873198.

[630] Ioana Şora and Doru Todinca. 2015. Dealing with fuzzy QoS properties in service composition. In *2015 IEEE 10th Jubilee International Symposium on Applied Computational Intelligence and Informatics*, 197–202. DOI: 10.1109/SACI.2015.7208198.

[631] Emilio Fdez-Viñas, Mateus Ferreira-Satler, Francisco P. Romero, Jesus Serrano-Guerrero, Jose A. Olivas, and Natalia Saavedra. 2012. Fuzzy ontologies to represent background knowledge: some experimental results on modelling climate change knowledge. In *Eurofuse 2011*. Springer Berlin Heidelberg, Berlin, Heidelberg, 179–191. DOI: https://doi.org/10.1007/978-3-642-24001-0_18.

[632] Norbert Sram. 2012. Minnesota code: evaluation of fuzzy ontology based approach. In *2012 IEEE 16th International Conference on Intelligent Engineering Systems (INES)*. IEEE, 477–481. DOI: 10.1109/INES.2012.6249881.

[633] Václav Slavíček. 2013. An ontology-driven fuzzy workflow system. In *SOFSEM 2013: Theory and Practice of Computer Science*. Peter van Emde Boas, Frans C. A. Groen, Giuseppe F. Italiano, Jerzy Nawrocki, and Harald Sack, editors. Springer Berlin Heidelberg, Berlin, Heidelberg, 515–527. DOI: https://doi.org/10.1007/978-3-642-35843-2_44.

[634] Aniello Minutolo, M Esposito, and G De Pietro. 2014. An ontology-based fuzzy approach for encoding cognitive processes in medical decision making. In *2014 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, 397–402. DOI: 10.1109/CogInfoCom.2014.7020487.

[635] Yakup Yildirim, Adnan Yazici, and Turgay Yilmaz. 2013. Automatic semantic content extraction in videos using a fuzzy ontology and rule-based model. *IEEE Transactions on Knowledge and Data Engineering*, 25, 1, 47–61. DOI: 10.1109/TKDE.2011.189.

[636] Panos Alexopoulos, Manolis Wallace, Konstantinos Kafentzis, and Dimitris Askounis. 2012. Ikarus-onto: a methodology to develop fuzzy ontologies from crisp ones. *Knowledge and Information Systems*, 32, 3, 667–695. DOI: 10.1007/s10115-011-0457-6.

[637] Xin Li, José-Fernán Martínez, and Gregorio Rubio. 2016. A new fuzzy ontology development methodology (FODM) proposal. *IEEE Access*, 4, 7111–7124. DOI: 10.1109/ACCESS.2016.2621756.

[638] Hai Bang Truong and Ngoc Thanh Nguyen. 2012. A multi-attribute and multi-valued model for fuzzy ontology integrationon instance level. In *Intelligent Information and Database Systems*. Springer Berlin Heidelberg, Berlin, Heidelberg, 187–197. ISBN: 978-3-642-28487-8. DOI: https://doi.org/10.1007/978-3-642-28487-8_19.

[639] Amir Vahid Dastjerdi and Rajkumar Buyya. 2014. Compatibility-aware cloud service composition under fuzzy preferences of users. *IEEE Transactions on Cloud Computing*, 2, 1, 1–13. DOI: 10.1109/TCC.2014.2300855.

[640] P. A. Khodke, M. G. Tingane, A. P. Bhagat, S. P. Chaudhari, and M. S. Ali. 2016. Neuro fuzzy intelligent e-learning systems. In *2016 Online International Conference on Green Engineering and Technologies (IC-GET)*, 1–7. DOI: 10.1109/GET.2016.7916766.

[641] Zainab Harbi, Yulia Hicks, and Rossi Setchi. 2017. Clock drawing test interpretation system. *Procedia Computer Science*, 112, (December 2017), 1641–1650. DOI: 10.1016/j.procs.2017.08.259.

[642] Dhanalakshmi Teekaraman, S. Sendhilkumar, and G. S. Mahalakshmi. 2020. Semantic provenance based trustworthy users classification on book-based social network using fuzzy decision tree. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 28, 01, 47–77. DOI: 10.1142/S0218488520500038.

[643] Nikos Manolis and Yannis Tzitzikas. 2011. Interactive exploration of fuzzy RDF knowledge bases. In *The Semantic Web: Research and Applications*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1–16. DOI: https://doi.org/10.1007/978-3-642-21034-1_1.

[644] Valerie Cross and Meenakshi Kandasamy. 2011. Fuzzy concept lattice construction: a basis for building fuzzy ontologies. In *2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)*, 1743–1750. DOI: 10.1109/FUZZY.2011.6007623.

[645] 2013. *Fuzzy markup language for malware behavioral analysis. On the Power of Fuzzy Markup Language*. Springer Berlin Heidelberg, Berlin, Heidelberg, 113–132. ISBN: 978-3-642-35488-5. DOI: 10.1007/978-3-642-35488-5_7.

[646] Che-Hung Liu, Chang-Shing Lee, Mei-Hui Wang, Yu-Yang Tseng, Yi-Ling Kuo, and Yi-Chen Lin. 2013. Apply fuzzy ontology and FML to knowledge extraction for university governance and management. *Journal of Ambient Intelligence and Humanized Computing*, 4, 4, 493–513. DOI: https://doi.org/10.1007/s12652-012-0139-6.

[647] Mustapha Bourahla. 2018. Description and reasoning for vague ontologies using logic programming. *IET Software*, 12, 5, 397–409. DOI: https://doi.org/10.1049/iet-sen.2017.0226.

[648] Yuncheng Jiang, Hai Liu, Yong Tang, and Qimai Chen. 2011. Semantic decision making using ontology-based soft sets. *Mathematical and Computer Modelling*, 53, 5, 1140–1149. ISSN: 0895-7177. DOI: https://doi.org/10.1016/j.mcm.2010.11.080.

[649] Mustapha Bourahla. 2015. Reasoning over vague concepts. In *Artificial Intelligence and Soft Computing*. Springer International Publishing, Cham, 591–602. DOI: https://doi.org/10.1007/978-3-319-19369-4_52.

[650] Paolo Pareti and Ewan Klein. 2014. Learning vague concepts for the semantic web. *ArXiv*, abs/1410.5078. DOI: https://doi.org/10.48550/arXiv.1410.5078.

[651] Paulo C. G. Costa, Kathryn B. Laskey, Erik Blasch, and Anne-Laure Jousselme. 2012. Towards unbiased evaluation of uncertainty reasoning: the URREF ontology. In *2012 15th International Conference on Information Fusion*. IEEE, 2301–2308. https://api.semanticscholar.org/CorpusID:6332767.

[652] Xiangjun Li, Taorong Qiu, Qing Liu, and Xiaoming Bai. 2010. Ontology building from incomplete information system based on granular computing. In *2010 IEEE International Conference on Granular Computing*, 292–296. DOI: 10.1109/GrC.2010.100.

Chapter 3

Possibilistic extension of Domain Information System (DIS) Framework

This chapter presents a novel extension of the DIS framework to support uncertainty modelling and reasoning under incomplete information grounded in possibility theory, which effectively captures partial knowledge through necessity-weighted formulas. It addresses the fourth objective of this thesis: to develop a structured approach for representing uncertainty of imperfections in DIS framework. The extended DIS framework enables the representation of uncertain concepts, relationships, and assertions, and supports necessity-based reasoning tasks such as subsumption and concept satisfiability. The chapter introduces the model's formal semantics, system architecture, and reasoning procedure.

Possibilistic extension of Domain Information System (DIS) Framework

Deemah Alomair[1,2][a], and Ridha Khedri[1][b]

[1]*Department of Computing and Software, McMaster University, Hamilton, Canada*

[2]*Department of Computer Information Systems, Imam Abdulrahman Bin Faisal University, Kingdom of Saudi Arabia*

{*dnalomair, khedri*}*@mcmaster.ca*

Keywords: Uncertainty, Incomplete Information, Knowledge Representation and Reasoning, Ontology Modelling, Ontologies, Possibilistic Logic, Ontology Reasoning.

Abstract: Uncertainty poses a significant challenge in ontology-based systems, manifesting in forms such as incomplete information, imprecision, vagueness, ambiguity, or inconsistency. This paper addresses this challenge by introducing a quantitative possibilistic approach to manage and model incomplete information systematically. Ontologies are modelled using the Domain Information System (DIS) framework, which is designed to handle Cartesian data structured as sets of tuples or lists, enabling the construction of ontologies grounded in the dataset under consideration. Possibility theory is employed to extend the DIS framework, enhancing its ability to represent and reason with incomplete information. The proposed extension captures uncertainty associated with instances, attributes, relationships, and concepts. Furthermore, we propose a reasoning mechanism within DIS that leverages necessity-based possibilistic logic to draw inferences under uncertainty. The proposed approach is characterized by its simplicity. It improves the expressiveness of DIS-based systems, introducing a foundation for flexible and robust decision-making in the presence of incomplete information.

1 INTRODUCTION

One of the primary challenges in knowledge-based systems, particularly those that rely on ontologies for domain reasoning, is managing uncertainty stemming from incomplete information. In dataset-driven ontologies, data is contextualized to define concepts, relationships, and instances. However, real-world applications frequently suffer from missing or partial information, leading to epistemic uncertainty (Sentz and Ferson, 2002). This type of uncertainty affects instance classification, attribute reliability, relationship strength, and concept validity. When unaddressed, such uncertainty can render ontologies either overly rigid, failing to accommodate partial knowledge, or misleading, by permitting unjustified inferences. Effectively managing uncertainty is therefore essential to ensure the expressiveness, reliability, and adaptability of ontology-based systems, especially in the context of decision support or automated reasoning systems. To illustrate, consider a customer service ontology; the concept `PositiveFeedback` may depend on attributes like `Satisfaction`, `Quality`, and `ResponseTime`. If one of these values is missing or partially available, classical inference systems may fail to classify an instance as `PositiveFeedback` or do so incorrectly. This highlights the need for a framework that can represent and reason under partial knowledge.

This paper introduces a quantitative possibilistic extension to the Domain Information System (DIS) framework (Marinache et al., 2021) to represent and reason under partial knowledge. DIS is a bottom-up, data-centric formalism that constructs ontologies from datasets, structurally separating the domain ontology from the data view and linking them via a mapping operator. Unlike Description Logic (DL)-based ontologies, which separate the A-Box and T-Box logically, DIS achieves this separation structurally and grounds the ontology in data, reducing data-ontology mismatches. DIS is useful for aligning ontologies with real-world datasets, which makes it particularly effective for domains where ontologies must be generated or adapted from existing data sources, improving modularity, transparency, and maintainability in ontology design. In contrast to traditional ontology languages like Web Ontology Language (OWL), which struggle to directly represent mereological relationships in Cartesian datasets (i.e., the structured data itself) without complex extensions, DIS lever-

[a] https://orcid.org/0000-0001-9397-9999

[b] https://orcid.org/0000-0003-2499-1040

ages cylindric algebra and Boolean algebra to model both data structures and conceptual part-whole relations. This enables more natural and robust handling of mereological reasoning within structured data. However, the original DIS model does not capture domain uncertainty and information uncertainty. The proposed approach overcomes this by associating each ontological component with quantified certainty.

Unlike vagueness or imprecision, the focus here is on uncertainty due to incompleteness, typically addressed via probability theory (e.g., (Laha and Rohatgi, 2020)), possibility theory (e.g., (Dubois and Prade, 2015)), or Dempster–Shafer theory (e.g., (Sentz and Ferson, 2002)), as discussed in (Alomair et al., 2025). In this study, possibility theory is adopted and rationale behind this selection is explained in the section 5.

The proposed approach models uncertainty across all key ontological elements: attributes, concepts, relationships, and instances. The key contributions of this paper are as follows:

1. Modelling Uncertainty of Attributes: Introduces a necessity-based mapping from the dataset's attributes to ontology concepts.

2. Modelling Uncertainty of Instances: Proposes an instance distribution relation ($SV^{\mathcal{D}}$), allowing a datum (instance) to be assigned to multiple sorts (attributes) with varying degrees of certainty.

3. Modelling Uncertainty of Relationships: Introduces necessity-based relationship, which allows relationships to hold with varying levels of certainty.

4. Modelling Uncertainty of Concepts: Refines the construction of datascape concepts (which depend on available data values) by incorporating uncertainty modelling into their data-specializing predicate.

5. Possibilistic Reasoning for Uncertainty-Aware Inference: Develops a reasoning mechanism within the DIS framework, leveraging necessity-based possibilistic logic to support inference under incomplete information.

The paper is structured as follows: Section 2 introduces foundational theories. Section 3 presents the integration of possibilistic components into the DIS framework, followed by uncertainty-aware reasoning in Section 4. Section 5 reviews related work and offers a discussion. Section 6 concludes the paper and outlines future directions.

2 PRELIMINARIES

This section reviews uncertainty in ontology, introduces possibility theory and possibilistic logic, and presents the theoretical background of the DIS framework.

2.1 Uncertainty and Ontology

Information imperfection includes incompleteness, imprecision, vagueness, ambiguity, and inconsistency (Ma et al., 2013; Bosc and Prade, 1997). The paper adopts a broad interpretation, considering uncertainty as arising from any of these deficiencies, as adopted in (Anand and Kumar, 2022; Ceravolo et al., 2008). *Incompleteness* arises when information is partial. This creates uncertainty about which interpretation of a statement to rely on, often addressed by calculating an estimation degree for possible worlds (Straccia, 2013). *Imprecision* refers to the lack of exactness, occurring when data is expressed in approximate or qualitative terms instead of precise values (Ma et al., 2013). *Vagueness* emerges when terms or concepts lack clear boundaries (Straccia and Bobillo, 2017). *Ambiguity* arises from multiple interpretations (Ma et al., 2013), and *inconsistency* involves contradictions, such as conflicting statements (Bosc and Prade, 1997).

An extensive review of uncertainty modelling in domain ontologies is presented in (Alomair et al., 2025). The survey examines over 550 studies published between 2010 and 2024 on this topic. A guiding taxonomy is proposed, classifying ontological uncertainty into *concept uncertainty* and *information uncertainty*. This classification supports the systematic identification of uncertainty types across ontological frameworks and the selection of appropriate formalisms to manage them. *Concept uncertainty* involves uncertainty of relationships, uncertainty of attributes defining a concept, and uncertainty due to semantic ambiguity, where context influences the interpretation of a concept. *Information uncertainty* concerns associating instances with concepts or relations. The identified uncertainties are attributed to incomplete, imprecise, vague, or inconsistent information. Then, various formalisms are presented to manage these uncertainties. This taxonomy offers a structured approach to understanding and addressing uncertainty in ontology-driven systems. A visual representation of the taxonomy is shown in Figure 1.

2.2 Possibility Theory and Possibilistic Logic

Possibility theory models incomplete and inconsistent knowledge using qualitative (ordinal) or quantitative (numerical) approaches (Dubois and Prade, 2015). The qualitative approach ranks events with-

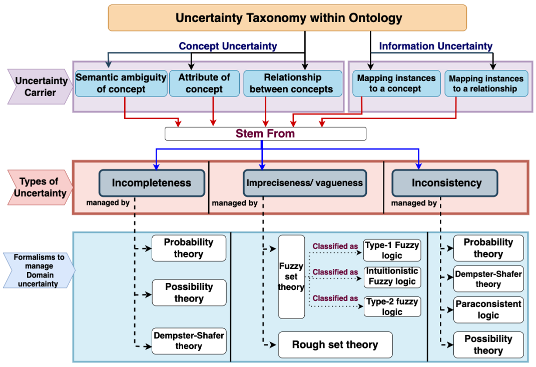


Figure 1: Ontological Uncertainty Taxonomy.

out numerical degree (e.g., ”highly possible”, ”possible”, or ”less possible”), while the quantitative approach assigns a numerical degree to represent degrees of possibility. Possibility distribution represents an agent's knowledge about the world by assigning plausibility degrees to states in a set S, which may be finite or infinite. Formally, it is a function $\pi : S \to L$, where L is a totally ordered scale (often $[0,1]$). The value $\pi(x)$ expresses how plausible the state $x \in S$. A value of $\pi(x) = 0$ means state x is impossible, while $\pi(x) = 1$ means it is fully plausible. If S is exhaustive, at least one state must have plausibility 1. The possibilistic framework captures both complete and incomplete knowledge. Complete knowledge is represented by assigning possibility 1 to a single state and 0 to all others. Complete ignorance is modeled by assigning possibility 1 to all states, indicating that any state could be true. The possibility distribution forms the basis for defining possibility and necessity measures over any subset $X \subseteq S$:

$$\Pi(X) = \sup_{x \in X} \pi(x) \text{ and } N(X) = \inf_{x \notin X}(1 - \pi(x)), \quad (1)$$

where $\Pi(X)$ indicates feasibility, and $N(X)$ expresses certainty (Alola et al., 2013). The measures are dual via: $N(X) = 1 - \Pi(X')$, where X' is the complement of X. Possibility measures follow the *maxitivity axiom*: $\Pi(A \cup B) = \mathtt{max}(\Pi(A), \Pi(B))$, while necessity measure satisfies the dual *minitivity axiom*: $N(A \cap B) = \mathtt{min}(N(A), N(B))$. The necessity degree for the union of two sets satisfies the following property, expressed as (Dubois and Prade, 2014):

$$N(A \cup B) \geq \mathtt{max}(N(A), N(B)) \quad (2)$$

Uncertainty has often been treated as if it were only randomness, managed through probability theory, which is statistical in nature and grounded in measure theory (Zadeh, 1977). Yet uncertainty also has other facets (e.g., vagueness, incompleteness), requiring a distinct treatment provided by possibility theory, which is non-statistical and rooted in fuzzy set theory. Probability is associated with likelihood, belief, frequency, or proportion, while possibility expresses feasibility or ease of attainment; what is possible may not be probable, and what is improbable need not be impossible. In probability theory, measures are additive, requiring the sum of probabilities for mutually exclusive events in a universe of discourse to equal one. In contrast, possibility measures follow max–min rules, as indicated earlier. Random variables correspond to probability distributions, whereas fuzzy variables correspond to possibility distributions, defined by membership functions rather than frequencies. The imprecision intrinsic to human cognition and natural language is better modeled by possibility theory rather than probability theory. Moreover, probability and possibility are connected by a consistency principle: decreasing possibility reduces probability, but not vice versa (Zadeh, 1977). Furthermore, a distinction between possibility and probability theories has been made through an example of *”Hans is eating eggs for breakfast”*. In his example, the possibility distribution of ($\pi_X(3) = 1$) suggests it is entirely possible for Hans to eat three eggs, but the probability ($P_X(3) = 0.1$) indicates this outcome is statistically rare. This demonstrates that high possibility does not imply high probability, though an impossible event ($\pi_X(u) = 0$) has zero probability ($P_X(u) = 0$).

Possibility theory underpins possibilistic logic, which we limit here to necessity-based possibilistic logic (Dubois et al., 1994; Dubois and Prade, 2014; Nieves et al., 2007). In this logic, a formula is a pair (θ, α), where θ is a classical first-order logic formula, and $\alpha \in [0,1]$ is a certainty or priority degree. This pair indicates that θ is certain at least to level α, (i.e., $N(\theta) \geq \alpha$). The interval $[0,1]$ can be replaced by any linearly ordered scale. Standard limit conditions hold: $\Pi(\perp) = N(\perp) = 0$, $\Pi(\top) = N(\top) = 1$, where $\perp$ and $\top$ denote contradiction and tautology, respectively. In the formal system of this logic, the following properties hold: $N(\theta \wedge \gamma) = \mathtt{min}(\{N(\theta), N(\gamma)\})$ and $N(\theta \vee \gamma) \geq \mathtt{max}(\{N(\theta), N(\gamma)\})$, where θ and γ are formulae. One of its main rules is the weakest link resolution rule:

$$(\neg\theta \vee \gamma, \alpha), (\theta \vee \delta, \beta) \vdash (\gamma \vee \delta, \mathtt{min}(\alpha, \beta)), \quad (3)$$

Here, the conclusion's certainty is the smallest among the premises, reflecting that an inference chain is limited by its weakest premise.

The weighted minimum and maximum operations, introduced in (Grabisch, 1998; Dubois and Prade, 1986) within the framework of possibility theory, generalize the standard min and max functions to account for elements from different contexts, each associated with a distinct weight of importance. These operations refine aggregation by modulating the influence of each element based on its assigned weight.

Let $X = \{x_1, \cdots, x_n\}$ be a set of criteria. Let a_i and w_i be, respectively, the score and the weight of importance attributed to criterion x_i such that $\Sigma_{i=1}^{n} w_i = 1$. Then we have:

$$\texttt{Weighted\_Min}(a_1, \cdots, a_n) = \texttt{min}(i \mid 1 \leq i \leq n : \texttt{max}((1 - w_i), a_i)).$$

This formulation ensures that elements with lower weights contribute less to the overall minimum computation. Similarly, the operation `Weighted_Max` is given by:

$$\texttt{Weighted\_Max}(a_1, \cdots, a_n) = \texttt{max}(i \mid 1 \leq i \leq n : \texttt{min}(w_i, a_i)).$$

2.3 Domain Information System

DIS is an ontology framework consisting of three primary components (Marinache et al., 2021; Marinache, 2025): Domain Ontology View (DOnt) $\mathcal{O}$, Domain Data View (DDV) $\mathcal{A}$, and mapping function τ linking $\mathcal{A}$ to $\mathcal{O}$, forming the structure $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$.

The DOnt, $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$, is composed of three elements. The concept structure $\mathcal{C} = (C, \oplus, e_c)$, which is a commutative idempotent monoid where the carrier set C includes an empty concept (e_c), a set of atomic concepts ($\mathcal{T}$) derived directly from dataset attributes, and composite concepts formed using the $\oplus$ operator. The Boolean lattice $\mathcal{L} = (L, \sqsubseteq_c)$ organizes concepts hierarchically based on a natural order $\sqsubseteq_c$, defined as $c_1 \sqsubseteq_c c_2 \iff c_1 \oplus c_2 = c_2$. Lastly, the set of rooted graphs $\mathcal{G}$ provides additional expressiveness by capturing concepts and relations beyond those defined by the lattice structure. Each rooted graph $G_{t_i} = (C_i, R_i, t_i)$ consists of a set of vertices $C_i \subseteq C$, a set of edges R_i, and a root vertex $t_i \in L$.

The DDV, $\mathcal{A} = (A, +, \star, -, 0_A, 1_A, \{c_k\}_{k \in \mathcal{U}})$, is formalized as a diagonal-free cylindric algebra, where $\mathcal{U}$ is a finite set of sorts (the universe). The main notion of this view is sort, which corresponds to an attribute in the dataset. The ordered pair of a sort and its value is known as `Sorted Value` (`SV`). A set of `SV` with a maximum of one `SV` for each sort forms `Sorted Datum` (`S_Datum`). The carrier set A consists of `Sorted Data` (`S_Data`), structured as a set of `S_Datum`. The cylindrification operators c_k are indexed by the sorts used in the data, corresponding to the elements of L, the carrier set of the Boolean lattice $\mathcal{L}$. For a deeper understanding of cylindric algebra, readers are referred to (Imieliński and Lipski, 1984).

The final component of DIS is the mapping function $\tau : A \to L$, which links the elements of $\mathcal{A}$ in DDV to their corresponding concepts in the Boolean lattice $\mathcal{L}$ within DOnt. To define τ, several helper operators introduced, one of which is the helper mapping operator $\eta : \mathcal{U} \to L$. This ensures a one-to-one correspondence between the sorts in DDV and the atomic concepts in the Boolean lattice of DOnt. Ensuring a seamless mapping from data attributes to ontology concepts: $\eta(S_{attr}) = attr$, where S_{attr} and $attr$ are a sort and an atomic concept, respectively.

In DIS, concepts are categorized based on their dependence on objective reality or data elements, leading to the distinction between *objective concepts* and *datascape concepts*, denoted by C_d. Objective concepts exist independently of any dataset. For instance, consider the objective statement $\exists(x \mid x \in \text{Animal} : \text{Pet}(x))$. The concept *Pet* remains valid regardless of whether supporting data is available. In contrast, *datascape concepts* rely on data for their definition and existence. For instance, consider the modified example $\exists(x \mid x \in \text{Animal} : \text{Active\_Pet}(x))$. The concept *Active_Pet*, defined as a pet that exercises for at least one hour daily, depends on a specific data source such as daily activity logs. If such data is unavailable or does not meet the required conditions, the concept cannot be realized. Formally, a datascape concept in a DIS is defined as follows:

Definition 1 (From (Alomair and Khedri, 2025), Datascape Concept)**.** *Let $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$ be a given DIS. For a carrier set A in $\mathcal{A}$ and a lattice $\mathcal{L}$ in $\mathcal{O}$, a datascape concept C_d is defined as follows: $C_d \stackrel{def}{=} \{a \mid \tau(a) \in L \wedge \Phi(a)\}$, where $a \in A$ and Φ is a data-specializing predicate expressed in Disjunctive Normal Form (DNF). This predicate Φ is given by: $\Phi(a) = \vee\ (i \mid 1 \leq i \leq N : \Psi_i(a))$, with N is a natural number, and each conjunctive clause $\Psi_i(a)$ is defined as: $\Psi_i(a) = \wedge\ (j \mid 1 \leq j \leq M : \Omega_{(i,j)}(a))$, where M is a natural number and $\Omega_{(i,j)}(a) = (f_{(i,j)}(a \cdot \texttt{sort\_name}_{(i,j)}), c_{(i,j)}) \in R_{(i,j)}$, where $f_{(i,j)} \in \mathcal{F}$, and $\mathcal{F} = \{\oplus, e_c, \top_{\mathcal{L}}, +, \star, -, 0, 1, \tau, cyl\}$ is the set of function symbols, $c_{(i,j)}$ is a ground term in the DIS language, and $R_{(i,j)}$ is a relator.* □

Based on the above definition, we define the operation $\oplus$ as an operation on concepts.

Definition 2. *Let $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$ be a given DIS. Let $C_{d1} = \{a \mid \tau(a) \in L \wedge \Phi_1(a)\}$, and $C_{d2} = \{a \mid \tau(a) \in L \wedge \Phi_2(a)\}$ be two datascape concepts defined on $\mathcal{D}$. We have $C_{d1} \oplus C_{d2} =$*

$$C_{d1} \cup C_{d2} = \{a \mid \tau(a) \in L \wedge (\Phi_1(a) \vee \Phi_2(a))\}.$$

The structure of the $C_{d1} \oplus C_{d2}$ is that of a datascape as $(\Phi_1(a) \vee \Phi_2(a))$ is in DNF and the other conditions stipulated by Definition 1 are satisfied. Moreover, the empty concept e can be perceived as a datascape concept defined as $e = \{a \mid \tau(a) = e_c \wedge \mathsf{false}\} = \emptyset$. Hence, if we take, for a given DIS, C_d is the set of datascape concepts, then $(C_d, \oplus, e_c)$ is a commutative monoid due to the properties of set union.

Illustrative Example of DIS Construction. We consider a `CustomerService` dataset with the attributes: `Satisfaction`, `Quality`, and `ResponseTime`. The corresponding DIS structure is built as follows:

1. Lattice construction: Each dataset attribute is mapped to an atomic concept: $\tau = \{(\texttt{Quality}, \texttt{Status}), (\texttt{ResponseTime}, \texttt{Duration}), (\texttt{Satisfaction}, \texttt{Comfort})\}$. Then the rest of the Boolean lattice is generated, where each node represents a possible composition of atomic concepts (e.g., `status Tenure` $=$ `Status` $\oplus$ `Duration`).
2. Objective rooted graph concept: Rooted graphs enrich the ontology beyond lattice nodes. One such objective concept is `Feedback`, rooted at `CustomerService`, and defined abstractly as follows: `Feedback` $\stackrel{\text{def}}{=} \{a \mid \tau(a) \in$ `CustomerService`$\}$.
3. Datascape rooted graph concept: A rooted graph concept might be a datascape concept, in which its definition depends on data. For example, the concept `PositiveFeedback` can be defined as: `PositiveFeedback`$= \{a \mid \tau(a) \in$ `CustomerService` $\wedge$ a.`Satisfaction` $\geq 0.6\}$. The predicate here indicates that an instance a of the `Satisfaction` attributes should have a value greater than or equal to 0.6.
4. Construction of the domain data view: An example of `SV` is (`Quality`, `Good`). An example of `S_Datum` is $dt\_1 = \{$(`Quality`, `Good`), (`ResponseTime`, `Fast`), (`Satisfaction`, `Yes`)$\}$. An example of `S_Data` is $a = \{dt\_1, dt\_n\}$.
5. Building the whole DIS system: The DIS is then formed by $(\mathcal{O}, \mathcal{A}, \tau)$. A full illustration of the DIS structure is shown in Figure 2.

3 UNCERTAINTY MODELLING IN DIS FRAMEWORK

In this section, we extend the DIS framework to handle uncertainty by addressing two key questions: What type of uncertainty can be modelled, and where in the DIS framework it can be introduced. As noted in section 1, we focus on incomplete information and adopt possibility theory as the formalism.

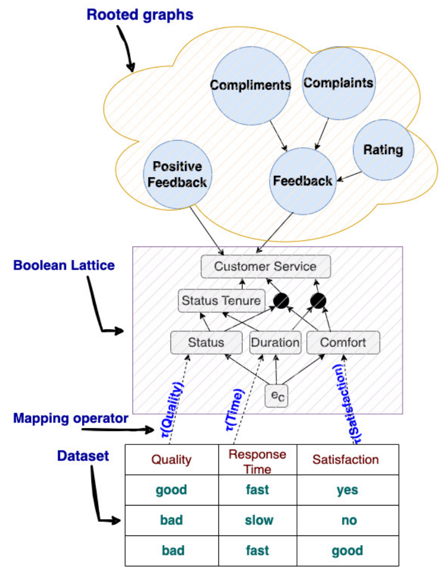


Figure 2: Customer Service DIS Framework.

To illustrate where uncertainty can arise, Figure 3 shows an example using a customer service dataset. Database attributes may assign values, introducing uncertainty of instances. These attributes are mapped via the operator τ (shown by arrows) to atomic concepts introducing uncertainty of attributes. The lattice is further expanded with multiple rooted graphs, such as `PositiveFeedback` and `Feedback`, introducing uncertainty of concepts. The `Feedback` graph includes specialized concepts like `Rating`, with arrows indicating semantic paths among these concepts, capturing the uncertainty of relationships.

Since the focus is on uncertainty due to incomplete information, it is crucial to distinguish between data and information. In our formalism, a datum is strictly a raw value without any assigned context (e.g., the number 3.7 isolated from metadata, units, or semantics). At this stage, it has no uncertainty; Uncertainty arises only when contextual interpretation is applied (e.g., labelling 3.7 as “sensor voltage reading with ±0.2 error”). We acknowledge that the broader literature often treats data as implicitly contextualized (and thus uncertain), but our formalism explicitly separates raw values from their contextual layers. It is also important to emphasize that the assigned degree is explicitly interpreted as a measure of certainty, not as a degree of truth or graded quality. For this reason, we adopt necessity-based possibilistic logic, where necessity degrees directly correspond to the degree of certainty. This interpretation aligns naturally with our setting, in which the degree reflects the certainty in the existence of concepts, in instance-to-concept and attribute-to-concept associations, and in the presence of relationships. In our approach, we examine four types of uncertainty:

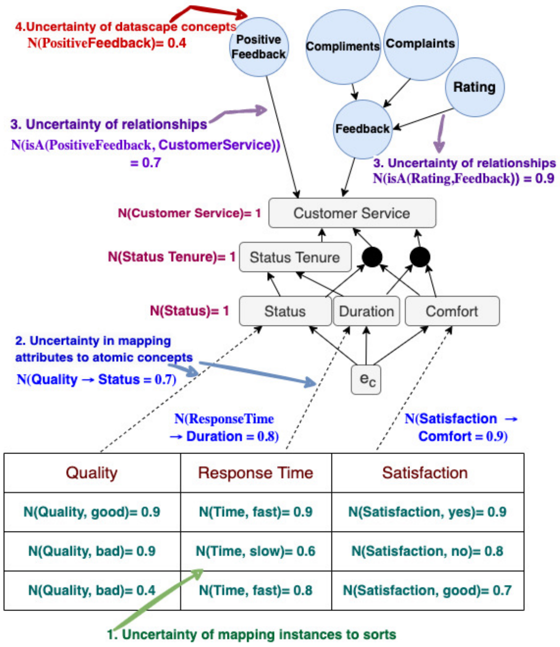


Figure 3: Necessity Degrees Assigned to DIS

1. *Uncertainty of mapping instances to sorts*: When mapping a value to a sort(attribute), for example, as indicated in Figure 3, the $\texttt{Quality}$ attribute being assigned values like $\texttt{Good}$ or $\texttt{Bad}$, with a necessity degree reflecting the degree of certainty with which the value belongs to a given sort. For instance, assigning $(\texttt{Good}, 0.9)$ to $\texttt{Quality}$ indicates that for this particular instance, it is 0.9 certain that the $\texttt{Quality}$ is $\texttt{Good}$.

2. *Uncertainty in mapping attributes to atomic concepts*: When mapping a sort to a lattice concept, such as associating $\texttt{Quality}$ with the $\texttt{Status}$ concept as $N(\texttt{Quality} \rightarrow \texttt{Status}) = 0.7$.

3. *Uncertainty of relationships*: When defining relationships among rooted graph concepts, like the relationship between $\texttt{Rating}$ and $\texttt{Feedback}$ is associated with $N(\texttt{isA}(\texttt{Rating}, \texttt{Feedback})) = 0.9$.

4. *Uncertainty of datascape concepts*: This uncertainty arises when a concept is defined in terms of data conditions that may themselves be uncertain. For example, consider the datascape concept $\texttt{PositiveFeedback}$, defined as $\texttt{PositiveFeedback} = \{a \mid \tau(a) \in \texttt{CustomerService} \wedge a.\texttt{Satisfaction} \geq 0.6\}$. Here, the condition $(\Phi(a) = a.\texttt{Satisfaction} \geq 0.6)$ is the data-specializing predicate that characterizes the concept. In our framework, the necessity degree of the datascape concept itself, that is, the degree to which the concept $\texttt{PositiveFeedback}$ holds in the presence of incomplete information, is derived directly from the necessity with which its data-specializing predicate is satisfied.

The first three types of uncertainty that are listed above are given by the domain expert, while the last one is calculated.

3.1 Uncertainty of Mapping Instances to Sorts

In the traditional DIS framework, data records (instances) are typically assigned to sorts (attributes) through a certain mapping function. This assignment is $\texttt{SV} : V \rightarrow \mathcal{U}$, where V is a finite set of values assigned to the sort, and $\mathcal{U}$ is a finite set of sorts (the universe). For example, consider a customer service database presented in Figure 3, where the attribute $\texttt{Quality}$ can take values such as $\texttt{Good}$ and $\texttt{Bad}$. The traditional mapping function would assign these values to the $\texttt{Quality}$ sort, as $\texttt{SV}(\texttt{Good}) = \texttt{Quality}$ and $\texttt{SV}(\texttt{Bad}) = \texttt{Quality}$. However, uncertainty brings nondeterminism in this mapping, as a value might be assigned to several sorts with some degree of certainty. Hence, to account for the uncertainty in these assignments, we introduce a new relation called the *instance distribution relation*, denoted $SV^{\mathcal{D}}$, and defined as $SV^{\mathcal{D}} \subseteq V \times \mathcal{U} \times [0,1]$. The relation $SV^{\mathcal{D}}$ relates a data value to sorts and necessity degree that represent the degree of certainty in the assignment. A data value might be assigned to several sorts with varying degrees of certainty. In the example, the instance distribution relation could return values like: $SV^{\mathcal{D}} = \{(\texttt{Good}, (\texttt{Quality}, 0.9)), (\texttt{Bad}, (\texttt{Quality}, 0.9)), (\texttt{Bad}, (\texttt{Quality}, 0.4)), (\texttt{Good}, (\texttt{Satisfaction}, 0.7))\}$. Here, the data value $\texttt{Good}$ is assigned to the $\texttt{Quality}$ sort with a certainty of 0.9, while the value $\texttt{Bad}$ is assigned to the same sort with two different certainty degrees: 0.9 and 0.4. These reflect varying contexts, such as different data records, where assignment certainty differs. Although the notation does not explicitly represent context, it is implicitly captured through association with different instances. Additionally, $\texttt{Good}$ is assigned to the $\texttt{Satisfaction}$ sort with a certainty of 0.7. This extension enables the framework to better reflect uncertainty by accommodating varying degrees of certainty in data-to-sort assignments.

3.2 Uncertainty in Mapping Attributes to Atomic Concepts

As previously discussed in subsection 2.3, the DIS framework defines the helper mapping operator $\eta : \mathcal{U} \rightarrow L$, which assigns each sort (attributes) in $\mathcal{U}$ to its corresponding atomic concept in the Boolean lattice. Similar to the *uncertainty of instances*, uncertainty in attribute mapping introduces non-determinism when associating sorts with atomic lattice concepts. To account for this, we define the *mapping distribution relation* $\eta^{\mathcal{D}}$, which captures the uncertainty in this mapping. The relation $\eta^{\mathcal{D}} \subseteq \mathcal{U} \times L \times [0,1]$ is provided by a domain expert, and assigns a necessity degree to each potential mapping.

Consider the customer service database presented in Figure 3, where the mapping operators η are defined as follows:

$$\eta(\texttt{Quality}) = \texttt{Status}, \eta(\texttt{Satisfaction}) = \texttt{Comfort}, \eta(\texttt{ResponseTime}) = \texttt{Duration}.$$

In this mapping, the sorts `Quality`, `Satisfaction`, and `ResponseTime` correspond to the atomic concepts `Status`, `Comfort`, and `Duration`, respectively. To capture the uncertainty in these mappings, the mapping distribution relation $\eta^{\mathcal{D}}$ assigns a necessity degree to each association:

$$\eta^{\mathcal{D}}(\texttt{Quality}) = (\texttt{Status},\ 0.7),$$
$$\eta^{\mathcal{D}}(\texttt{Satisfaction}) = (\texttt{Comfort},\ 0.9),$$
$$\eta^{\mathcal{D}}(\texttt{ResponseTime}) = (\texttt{Duration},\ 0.8).$$

These degrees indicate the degree of certainty in each mapping, allowing the DIS framework to handle the uncertainty in the alignment between data attributes and ontology concepts.

3.3 Uncertainty of Relationships

Within the DIS framework, there are relationships between the concepts of rooted graphs and a parthood relationship between the concepts of the Boolean lattice. The parthood relationship $\sqsubseteq_c$ forms the relationship between objective concepts given in the lattice. The existence of this relationship among lattice concepts is certain, as they are constructed by a Cartesian construction from the atomic concepts. In other terms, a concept k_1 is considered a `partOf` another concept k, if k_1 is a Cartesian projection of k or if its atomic structure is a subset of that of k. However, the relations among the concepts of the rooted graph might be uncertain. Given a rooted graph $G_{t_i} = (C_i, R_i, t_i)$, $C_i \subseteq C$, $R_i \subseteq C_i \times C_i$, $t_i \in L$, its relation is transformed to give each edge a necessity degree. We extend R_i to a necessity-based $R_i^{\mathcal{D}}$. Hence, $R_i^{\mathcal{D}} \subseteq R_i \times [0,1]$, which incorporates necessity degrees to quantify the degree of certainty associated with each relationship.

In the customer service database illustrated in Figure 3, the relation of the rooted graph, denoted by R_i, is the following: $R_i =$

$$\{(\texttt{isA(PositiveFeedback,CustomerService)}),$$
$$(\texttt{isA(Complaints,Feedback)}),$$
$$(\texttt{isA(Rating,Feedback)}),$$
$$(\texttt{isA(Feedback,CustomerService)}),$$
$$(\texttt{isA(Compliments,Feedback)})\}$$

Hence, the relations $R_i^{\mathcal{D}}$ is given as follows: $R_i^{\mathcal{D}} = \{$

$$(\texttt{isA(PositiveFeedback,CustomerService)}, 0.4),$$
$$(\texttt{isA(Complaints,Feedback)}, 1),$$
$$(\texttt{isA(Rating,Feedback)}, 0.9)$$
$$(\texttt{isA(Feedback,CustomerService)}, 0.5),$$
$$(\texttt{isA(Compliments,Feedback)}, 0.7)\}$$

These necessity degrees quantify the degree of certainty in each relationship, enabling the framework DIS to systematically capture and reason about uncertainty in relational structures.

3.4 Uncertainty of Datascape Concepts

If we examine the elementary predicate $\Omega_{(i,j)}(a)$, which is used in building the data-specializing predicate $\Phi(a)$ of a datascape concept and which is equal to $(f_{(i,j)}(a \cdot \texttt{sort\_name}_{(i,j)}), c_{(i,j)}) \in R_{(i,j)}$, we find that there are two sources of uncertainty. The first comes from mapping a datum a to a sort due to the usage of the term $a \cdot \texttt{sort\_name}_{(i,j)}$, and the second comes from the relator $R_{(i,j)}$ used in $\Omega_{(i,j)}(a)$. Hence, by capturing these two sources of uncertainty, we capture the uncertainty of the datascape concept. For that, we adopt the weighted minimum function, previously defined in subsection 2.2. The weights of instance mapping $w_{\texttt{inst}}$, and the weight of the relationship $w_{\texttt{rel}}$ assign relative importance to the necessity measures $SV^{\mathcal{D}}(a)$ and $R_i^{\mathcal{D}}(a)$, with $w_{\texttt{inst}} + w_{\texttt{rel}} = 1$. Then, we have the following inductive procedure for calculating the necessity degree $N_{\Phi(a)}$ of a datascape concept having Φ as its data-specializing predicate.

Procedure 3.1 (Necessity Degree of a Datascape Predicate)**.** *Let $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$ be a given DIS. Let $C_d = \{a \mid \tau(a) \in L \wedge \Phi(a)\}$ be a datascape concept that is defined within $\mathcal{D}$, and has Φ as its specializing predicate. For a given element $a \in A$, let $\delta = (w_{\texttt{inst}} = w_{\texttt{Wrel}}) \vee \Big((SV^{\mathcal{D}}(a) \leq (1 - w_{\texttt{inst}})) \wedge (R^{\mathcal{D}}_{(i,j)}(a) \leq (1 - w_{\texttt{Wrel}}))\Big)$. The necessity degree $N(\Phi(a))$ is computed inductively as follows:*

- *Base cases:*

1. $N(\textsf{true}) = 1$;
2. $N(\textsf{false}) = 0$.
3. $N(\Omega_{(i,j)}(a)) =$

$$\begin{cases} \mathtt{min}\Big(SV^{\mathcal{D}}(a), R^{\mathcal{D}}_{(i,j)}(a)\Big), & \text{if } \delta = \textsf{true}, \\ \mathtt{min}\Bigg(\mathtt{max}\Big((1 - w_{inst}), SV^{\mathcal{D}}(a)\Big), \\ \quad \mathtt{max}\Big((1 - w_{rel}), R^{\mathcal{D}}_{(i,j)}(a)\Big)\Bigg), & \text{otherwise}. \end{cases}$$

- *Inductive cases:*
  1. *Conjunction of atomic predicates:*

$$N(\Psi_i(a)) = \mathtt{min}\Big(j \mid 1 \leq j \leq M : N(\Omega_{(i,j)}(a))\Big)$$

  2. *Disjunction of conjunctive clauses:*

$$N(\Phi(a)) = \mathtt{max}\Big(i \mid 1 \leq i \leq N : N(\Psi_i(a))\Big)$$

□

In the base case, the necessity degree of each atomic predicate is considered. The necessity degree of the ground terms true and false are, respectively, 1 and 0. For the elementary term $\Omega_{(i,j)}(a)$ forming $\Phi(a)$, we have several cases:

- When we have equal weights of all the criteria, then the weights are omitted in determining $N(\Omega(a))$.
- When both the certainty of a mapped to its sort is below 1 minus the weight assigned to the mapping, and the certainty of the relator $R^{\mathcal{D}}_{(i,j)}(a)$ used in Ω is also below 1 minus the weight assigned to the relationship, then the weights are also omitted. Hence, when $(SV^{\mathcal{D}}(a) \leq (1 - w_{\mathtt{inst}})$ and $R^{\mathcal{D}}_{(i,j)}(a) \leq (1 - w_{\mathtt{Wrel}})$ means that the importance or influence of the mapping of instances to sorts and the relator in the overall-uncertainty determination outweighs the level of uncertainty associated with it. That is why we ignore the weights in this case.

We can extend the necessity degree function to the datascape concepts as follows: $N(C_d) \overset{\text{def}}{=} N(\Phi) = \mathtt{max}\Big(a \mid a \in A : N(\Phi(a))\Big)$, where $C_d = \{a \mid \tau(a) \in L \wedge \Phi(a)\}$ is a datascape concept that is defined within a DIS $\mathcal{D}$. We take the max of the individual necessity degrees due to the union property described earlier in Equation 2.

For objective concepts in the lattice, the composition operator $\oplus$ enables the formation of new concepts by composing existing ones i.e., creating composite concepts from the set of atomic concepts T. Writing $k = k_1 \oplus k_2$ means that concept k is constructed by the Cartesian product of concepts k_1 and k_2. These concepts are certain and carry no uncertainty. An alternative way to consider uncertainty in an objective concept is considering its specializing predicate that is always true, hence its certainty degree is 1.

4 REASONING ON POSSIBILISTIC DIS FRAMEWORK

We discuss several reasoning tasks and their governing inference rules for deriving conclusions in different reasoning scenarios. These tasks are concept satisfiability and concept subsumption. Each of which is explained in detail. We use N to denote the necessity degree function.

4.1 Concept Satisfiability

In this subsection, we examine concept satisfiability in necessity-based reasoning within the DIS framework, distinguishing between the objective and the datascape concept satisfiability.

In the classical DIS framework, a datascape concept is considered satisfiable if its corresponding data values exist within the carrier set of the DDV. However, in the necessity-based extension of DIS, we introduce the necessity degree to account for incomplete information of the data specializing predicate ($\Phi(a)$), which defines the datascape concept. In this extended framework, a datascape concept C_d is deemed satisfiable if there exists at least one instance $a \in C_d$ such that the necessity degree $N(a, C_d)$ of this instance is strictly greater than zero. Therefore, a datascape concept is satisfiable if and only if its necessity degree is strictly greater than zero, indicating that there is sufficient data support for the concept's existence.

Definition 3. *Let $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$ be a given DIS. Let $C_d = \{a \mid \tau(a) \in L \wedge \Phi(a)\}$ be a datascape concept that is defined within $\mathcal{D}$, with $\Phi(a)$ as its data specializing predicate. The datascape concept C_d is satisfiable, denoted by $\mathtt{stsfd}(C_d)$, if and only if* $\exists(a \mid a \in A : N(\Phi(a)) > 0)$.

For objective concepts within the Boolean lattice, their certainty is inherently guaranteed, as they are directly linked to the DDV of the DIS under consideration. Thus, their satisfiability is inherently ensured, meaning they are both valid and certain to exist. The satisfiability of a composite concept is also guaranteed, as its atomic components have a degree of necessity of one. In this case, the combination of their necessity degree results in the composite concept also having a necessity degree of one, ensuring its satisfiability. If, from another perspective, one sees objective concepts as concepts that are independent of datasets, which translates into a data specializing predicate equivalent to true, then using Definition 3 and Procedure 3.1(item 1), one infers that its necessity degree is also equal to one.

Claim 4.1. *Let C_{d1} and C_{d2} be datascape concepts defined in a given DIS. Let $\Phi_1(a)$ and $\Phi_2(a)$ be their data specializing predicates, respectively. We have*
$\mathtt{stsfd}(C_{d1} \oplus C_{d2}) \equiv \mathtt{stsfd}(C_{d1}) \vee \mathtt{stsfd}(C_{d2})$.

Proof. The concepts C_{d1} and C_{d2} are two datascape concepts. Hence, by Definition 1 and for $\mathcal{D} = (O, \mathcal{A}, \tau)$ is a given DIS, we can write C_{d1} and C_{d2} as follows: $C_{d1} \stackrel{\text{def}}{=} \{a \mid \tau(a) \in L \wedge \Phi_1(a)\}$, and $C_{d2} \stackrel{\text{def}}{=} \{a \mid \tau(a) \in L \wedge \Phi_2(a)\}$.

Then, we have $\mathtt{stsfd}(C_{d1} \oplus C_{d2})$

$\equiv$ ⟨ Definition 2 ⟩

$\mathtt{stsfd}\{a \mid \tau(a) \in L \wedge (\Phi_1(a) \vee \Phi_2(a))\}$

$\equiv$ ⟨ Definition 3: Satisfiability of datascape concept ⟩

$\exists(a \mid a \in A : N(\Phi_1(a)) > 0 \vee N(\Phi_2(a)) > 0)$

$\equiv$ ⟨ Axiom Distributivity for $\vee$ ⟩

$\exists(a \mid a \in A : N(\Phi_1(a)) > 0)$
$\vee \exists(a \mid a \in A : N(\Phi_2(a)) > 0)$

$\equiv$ ⟨ Definition 3 ⟩

$\mathtt{stsfd}(C_{d1}) \vee \mathtt{stsfd}(C_{d2})$

□

Example 4.1 (Satisfiability of a Datascape Concept). *Consider the datascape concept $\mathtt{PositiveFeedback} = \{a \mid \tau(a) \in \mathtt{CustomerService} \wedge a.\mathtt{Satisfaction} \geq 0.6\}$. Thus, this concept consists of a single atomic predicate: $\Omega(a) = (a.\mathtt{Satisfaction} \geq 0.6)$. Assume the following information is provided by a domain expert:*

- *Instance distribution relation:*
  $SV^{\mathcal{D}}(a_1) = (\mathtt{Good}, (\mathtt{Satisfaction}, 0.7))$
- *Relator necessity degree:* $R^{\mathcal{D}}(a_1) = (\mathtt{Good} \geq 0.6, 0.8)$
- *Wights of importance:* $w_{\mathtt{inst}} = 0.4, w_{\mathtt{Wrel}} = 0.6$

Then, using Procedure 3.1, we compute the necessity degree of the atomic predicate:

$$\mathtt{min}\Big(\mathtt{max}(1 - w_{\mathtt{inst}}, SV^{\mathcal{D}}(a_1)),\ \mathtt{max}(1 - w_{\mathtt{Wrel}}, R^{\mathcal{D}}(a_1))\Big)$$

$$= \mathtt{min}(\mathtt{max}(0.6, 0.7),\ \mathtt{max}(0.4, 0.8)) = 0.7$$

Since $\Phi(a)$ consists of just this atomic predicate, we have:

$$N(C_d) = N(\Phi(a)) = N(\Omega(a_1)) = 0.7$$

By Definition 3, the concept is satisfiable because $N(\Phi(a)) > 0$. Hence: $\mathtt{stsfd(PositiveFeedback)}$ holds. □

4.2 Necessity-Based Subsumption

In general, we say that a concept C_1 subsumes a concept C_2 if every instance of C_2 is in C_1. In classical DIS, we have an additional kind of subsumption relationship. It is the `partOf` relationship, denoted by $\sqsubseteq_c$, that exists among the members of the Boolean lattice. When we write $C_1 \sqsubseteq_c C_2$, it indicates that the instances of C_1 are obtained through the projection of corresponding instances of C_2 on the attributes defining C_1. In this case, we say that C_2 subsumes C_1. The formal definition of DIS based subsumption is given below:

Definition 4. *Given a DIS $\mathcal{D} = (O, \mathcal{A}, \tau)$. Let C_1 and C_2 be two concepts in the set of concepts of $\mathcal{D}$. We say that $C_2 \sqsubseteq C_1$ iff one of the conditions holds:*

1. $C_1 \in L \wedge C_2 \in L \wedge C_2 \sqsubseteq_c C_1$.

2. C_1 and C_2 are two datascape concepts with data specializing predicates Φ_1 and Φ_2, respectively and $\forall(a \mid a \in A : \Phi_2(a) \implies \Phi_1(a))$.

3. $C_1 \in L$ and C_2 is a datascape concept. We have $(C_1, C_2) \in R^*$, where R^* is the reflexive transitive closure of a relation R of the graph rooted at C_1.

Definition 4 formalizes concept subsumption in DIS, covering both objective and datascape concepts. First, if C_1 and C_2 are objective concepts in the ontology lattice, subsumption holds if $C_2 \sqsubseteq_c C_1$, meaning C_2 is structurally more specific than C_1 per the lattice order, reflecting the traditional subclass relation of the lattice hierarchical structure. Second, if both are datascape concepts defined by data specializing predicates Φ_1 and Φ_2, then $C_2 \sqsubseteq C_1$ holds if $\forall a \in A$, the implication $\Phi_2 \implies \Phi_1$ is satisfied. This ensures all instances satisfying C_2 also satisfy C_1. Third, if C_1 is an objective concept and C_2 is a datascape, subsumption holds if a path exists from C_1 to C_2 in the reflexive-transitive closure R^* of relation R. Notably, all concepts in the graph rooted at C_1 are considered a specialization of C_1. Subsumption is a partial order (reflexive, antisymmetric, transitive) over concepts, as stated in the following claim.

Claim 4.2. *Given a DIS $\mathcal{D} = (O, \mathcal{A}, \tau)$, the subsumption relation on the set of concepts in $\mathcal{D}$ is a partial order.*

Proof. We provide a proof for each case of the subsumption relation as given in Definition 4.

1. In the first case, the subsumption relation is identical to the `partOf` (i.e., $\sqsubseteq_c$) relation. The latter satisfies the properties required for subsumption (reflexivity, transitivity, and anti-symmetry) because it is defined on a Boolean lattice, which is itself a partially ordered set (poset) (Marinache, 2025).

2. In the second case, the subsumption relation $\sqsubseteq$ defines a partial order over the datascape concepts. This is due to the properties of the logical $\implies$ operation that satisfies the properties of partial order (Gries and Schneider, 1993, Pages 57-59).

3. In the third case, subsumption is interpreted as membership to the reflexive transitive closure R^*. It is established that a reflexive transitive closure on an acyclic graph is a partial order. Indeed, the rooted graphs are acyclic, as furthermore each is a Directed Acyclic Graph (DAG).

□

In the context of necessity-based subsumption, the subsumption necessity degree is determined by a domain expert. In addition, as the transitivity of subsumption applies, the degree of transitive subsumption is governed by the weakest link resolution rule, presented previously in subsection 2.2. This approach to possibilistic transitivity reasoning has been adopted in prior research (e.g., (Mohamed et al., 2018; Benferhat and Bouraoui, 2015)) and has shown effectiveness in handling uncertainty within possibilistic ontologies.

Claim 4.3. *Let $\mathcal{D} = (O, \mathcal{A}, \tau)$ be a DIS. Let C_1, C_2, and C_3 be concepts defined in $\mathcal{D}$. Let*

$$R_{\sqsubseteq} = \{(C_1, C_2) \mid C_1 \in C \land C_2 \in C \land C_1 \sqsubseteq C_2\}$$

and $R^{\mathcal{D}}_{\sqsubseteq}$ is its corresponding necessity relation. We have:

$$((C_1, C_2), \alpha) \in R^{\mathcal{D}}_{\sqsubseteq} \land ((C_2, C_3), \beta) \in R^{\mathcal{D}}_{\sqsubseteq} \implies ((C_1, C_3), min(\alpha, \beta)) \in R^{\mathcal{D}}_{\sqsubseteq}$$

Proof. $((C_1, C_2), \alpha) \in R^{\mathcal{D}}_{\sqsubseteq} \land ((C_2, C_3), \beta) \in R^{\mathcal{D}}_{\sqsubseteq}$

$\equiv$ $\langle$ Definition of $R^{\mathcal{D}}_{\sqsubseteq}$ $\rangle$

$(C_1, C_2) \in R^{\mathcal{D}}_{\sqsubseteq} \land (C_2, C_3) \in R^{\mathcal{D}}_{\sqsubseteq} \land N((C_1, C_2)) = \alpha \land N((C_2, C_3)) = \beta$

$\implies$ $\langle$ Transitivity of $R_{\sqsubseteq}$ and the weakest link resolution rule (Equation 3) $\rangle$

$(C_1, C_3) \in R^{\mathcal{D}}_{\sqsubseteq} \land N((C_1, C_3)) = \mathtt{min}(\alpha, \beta)$

$\equiv$ $\langle$ Definition of $R^{\mathcal{D}}_{\sqsubseteq}$ $\rangle$

$((C_1, C_3), \mathtt{min}(\alpha, \beta)) \in R^{\mathcal{D}}_{\sqsubseteq}$

□

The necessity-based subsumption between objective concepts (`partOf` relation) invariably assumes a necessity degree of 1, since the parthood relation among lattice concepts is considered fully certain i.e., $N(\mathtt{partOf}) = 1$, as indicated previously in subsection 3.3. This certainty extends naturally to the transitivity of `partOf` relation, whereby the minimum necessity degree computed over a chain of parthood relations, among lattice concepts, gives a degree of 1.

Example 4.2 (Transitivity of Concept Subsumption in DIS)**.** *Let C_1 = `Complaints`, C_2 = `Feedback`, C_3 = `CustomerService` be three concepts defined in given DIS. Suppose the following necessity-based subsumption relationships are provided by the domain expert:*

$$\{((\mathtt{Complaints}, \mathtt{Feedback}), 0.6), ((\mathtt{Feedback}, \mathtt{CustomerService}), 0.8)\} \subseteq R^{\mathcal{D}}_{\sqsubseteq}$$

By Claim 4.3, the transitive subsumption relation holds with:

$$N(\mathtt{Complaints}, \mathtt{CustomerService}) = 0.6$$

5 RELATED WORK AND DISCUSSION

This section reviews ontology modelling approaches that handle uncertainty using possibility theory, and explains our choice of this formalism.

In possibilistic DL-based approaches, uncertainty is modelled by assigning necessity or possibility degrees to ontology axioms at different levels. For instance, Pb-π-DL-Lite (Boutouhami et al., 2017) assigns necessity degrees only to ABox assertions, allowing uncertain instance membership such as $(\mathtt{Status}, \mathtt{Good}) = 0.9$ in the `CustomerService` domain, without modelling uncertainty at concept or relationship levels. The work of (Sun, 2013) focuses on uncertainty at the TBox level, assigning necessity degrees to TBox axioms such as $N(\mathtt{isA}(\mathtt{Rating}, \mathtt{Feedback})) = 0.8$, yet does not support uncertain instance classification or data-driven concept definitions. Similarly, studies such as (Benferhat and Bouraoui, 2015; Benferhat et al., 2014; Qi et al., 2011) extend possibilistic logic to both TBox and ABox axioms, enriching the expressiveness by allowing weighted axioms at multiple levels; however, their frameworks still assume that concepts like `PositiveFeedback` are defined and do not enable concept definitions directly derived from data (i.e., datascape concepts). The study of (Mohamed et al., 2018) further incorporates possibility distributions over interpretations, adding expressiveness to represent uncertainty about models themselves, but does not provide mechanisms to ground concepts in uncertain data attributes or integrate graded uncertainty at the attribute-concept mapping level.

At the language level, (Safia and Aicha, 2014) propose extending Web Ontology Language 2 (OWL2) with possibilistic annotations, enabling uncertainty representation in both concepts and instances. Using the `CustomerService` example, one could annotate a concept like `PositiveFeedback`, or an instances like `Fast` with possibility degrees, yet the approach still requires concepts to be pre-defined and does not support automatic or context-aware construction of concepts from data conditions. Finally, the work of (Ben Salem et al., 2018) assigns possibility degrees directly to concepts outside DL semantics, like assigning possibility degree to the concept `PositiveFeedback`, but does not address uncertainty propagation from attribute data to instances or model uncertainty in relationships or attribute mappings.

In contrast, our DIS-based framework uniquely integrates uncertainty at all levels: attributes, instances, relationships, and data-driven concepts. For example, we model uncertain attribute-concept mappings such as $N(\mathtt{Quality} \rightarrow \mathtt{Status}) = 0.7$, uncertain instance-concept classification like $(\mathtt{Good}, 0.9)$, and relationships such as $N(\mathtt{isA}(\mathtt{Rating}, \mathtt{Feedback})) = 0.9$. Our datascape concept `PositiveFeedback` is defined by a data-specializing predicate reflecting the actual satisfaction values, allowing uncertainty in satisfaction data to propagate naturally to the membership degree in `PositiveFeedback` concept. This unified, data-grounded modelling allows more expressive, context-aware reasoning about uncertain information compared to existing possibilistic ontology methods.

From a methodological standpoint, choosing an uncertainty formalism requires alignment with the modelling goals and constraints of the framework. While several candidates exist, including probabilistic approaches and Dempster-Shafer Theory (DST), we adopt possibility theory for its suitability to our framework. Probability theory enforces the additivity axiom, requiring the sum of probabilities for mutually exclusive events in a universe of discourse to equal one even under insufficient data (Kovalerchuk, 2017), leading to challenges in accurately representing uncertainty. In contrast, possibility theory relaxes this constraint, making it more suitable for the proposed approach. This rationale is supported by several possibilistic ontology frameworks (e.g., (Bal-Bourai and Mokhtari, 2016; Boutouhami et al., 2017)). Regarding DST, it is primarily designed for belief fusion from multiple sources (McClean, 2003), whereas our approach derives certainty from a single source. Moreover, DST typically assigns belief to sets of hypotheses rather than individual ones, making it more suitable for representing group-level uncertainty. In contrast, the DIS framework demands fine-grained certainty assignments to individual attributes, values, and relationships (Sentz and Ferson, 2002; Gordon and Shortliffe, 1984). Possibility theory directly supports this by enabling necessity degrees to annotate specific elements, making it a natural fit for our ontology-based model.

6 CONCLUSION AND FUTURE WORK

This paper presents a principled extension of the DIS framework to support reasoning under incomplete information using necessity-based possibilistic logic. Unlike most ontology-based systems that assume complete information, our approach models uncertainty across instances, attributes, relationships, and concept definitions, enabling fine-grained, graded reasoning. A key advantage is replacing binary inferences with necessity-valued conclusions, allowing cautious reasoning with partial information. Overall, this approach provides a structured foundation for possibilistic reasoning in ontology-based systems advancing more expressive and uncertainty-aware knowledge representations essential for robust decision-making in complex, data-limited contexts.

We are currently automating necessity-based reasoning tasks using the Domain Information System Extended Language (DISEL) tool (Wang et al., 2022). Future work will focus on automating necessity degree assignment via machine learning, integrating fuzzy logic to handle imprecision, developing a scalable reasoning engine, and applying the framework to real-world domains. Once the automation is in place, we plan to use DISEL to reason over data collected from network security prevention mechanisms. This data is often uncertain and originates from diverse sources with varying levels of reliability. Furthermore, this data originates from log files, whether structured or semi-structured, making it well-suited for DIS modelling. The goal is to pre-process and clean the data (Khedri et al., 2013), then apply the proposed reasoning framework to facilitate reliable and context-aware security decision-making in highly dynamic and complex uncertain landscapes.

REFERENCES

Alola, A., Tunay, M., and Alola, U. (2013). Analysis of possibility theory for reasoning under uncertainty. *International Journal of Statistics and Probability*, 2:12.

Alomair, D. and Khedri, R. (2025). Towards a Cartesian theory of ontology domain adequacy. *Synthese*, pages 1–12. Submitted for publication on July. 25, 2025.

Alomair, D., Khedri, R., and MacCaull, W. (2025). A comprehensive review of uncertainty modelling in domain ontologies. *ACM Computing Surveys*, pages 1–35. Accepted with minor revision.

Anand, S. K. and Kumar, S. (2022). Uncertainty analysis in ontology-based knowledge representation. *New Generation Computing*, 40(1):339–376.

Bal-Bourai, S. and Mokhtari, A. (2016). Poss-SROIQ(D): Possibilistic description logic for uncertain geographic information. In *Trends in Applied Knowledge-Based Systems and Data Science*, pages 818–829, Cham. Springer International Publishing.

Ben Salem, Y., Idoudi, R., Ettabaa, K. S., Hamrouni, K., and Soleiman, B. (2018). Mammographie image based possibilistic ontological representation. In *(ATSIP) 4th International Conference*, pages 1–6.

Benferhat, S. and Bouraoui, Z. (2015). Min-based possibilistic DL-Lite. *Journal of Logic and Computation*, 27(1):261–297.

Benferhat, S., Bouraoui, Z., and Tabia, K. (2014). On the revision of prioritized DL-Lite knowledge bases. In *Scalable Uncertainty Management*, pages 22–36, Cham. Springer International Publishing.

Bosc, P. and Prade, H. (1997). *An Introduction to the Fuzzy Set and Possibility Theory-Based Treatment of Flexible Queries and Uncertain or Imprecise Databases*, pages 285–324. Springer US, Boston, MA.

Boutouhami, K., Benferhat, S., Khellaf, F., and Nouioua, F. (2017). Uncertain lightweight ontologies in a product-based possibility theory framework. *International Journal of Approximate Reasoning*, 88:237–258.

Ceravolo, P., Damiani, E., and Leida, M. (2008). Which role for an ontology of uncertainty? In *International Workshop on Uncertainty Reasoning for the Semantic Web*, volume 423.

Dubois, D., Lang, J., and Prade, H. (1994). Possibilistic logic. *Handbook of logic in artificial intelligence and logic programming*, pages 439–513.

Dubois, D. and Prade, H. (1986). Weighted minimum and maximum operations in fuzzy set theory. *Information Sciences*, 39(2):205–210.

Dubois, D. and Prade, H. (2014). Possibilistic logic—an overview. *Handbook of the History of Logic*, 9:283–342.

Dubois, D. and Prade, H. (2015). Possibility theory and its applications: Where do we stand? *Mathware and Soft Computing Magazine*, 18.

Gordon, J. and Shortliffe, E. H. (1984). The dempster-shafer theory of evidence. *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, 3(832-838):3–4.

Grabisch, M. (1998). *Fuzzy Integral as a Flexible and Interpretable Tool of Aggregation*, pages 51–72. Physica-Verlag HD, Heidelberg.

Gries, D. and Schneider, F. B. (1993). *A logical approach to discrete math*. Springer-Verlag, Berlin, Heidelberg.

Imieliński, T. and Lipski, W. (1984). The relational model of data and cylindric algebras. *Journal of Computer and System Sciences*, 28(1):80–102.

Khedri, R., Chiang, F., and Sabri, K. E. (2013). An algebraic approach for data cleansing. In *the 4th EUSPN*, volume 21 of *Procedia Computer Science*, pages 50 – 59. Procedia Computer Science.

Kovalerchuk, B. (2017). *Relationships Between Probability and Possibility Theories*, pages 97–122. Springer International Publishing, Cham.

Laha, R. G. and Rohatgi, V. K. (2020). *Probability theory*. Courier Dover Publications.

Ma, Z. M., Zhang, F., Wang, H., and Yan, L. (2013). An overview of fuzzy description logics for the semantic web. *The Knowledge Engineering Review*, 28(1):1–34.

Marinache, A. (2025). *Syntax and Semantics of Domain Information System and its usage in conjecture verification*. PhD thesis, School of Graduate Studies, McMaster University, Hamilton, Ontario, Canada.

Marinache, A., Khedri, R., LeClair, A., and MacCaull, W. (2021). DIS: A data-centred knowledge representation formalism. In *(RDAAPS) 2021: A Big Data Challenge*, pages 1–8.

McClean, S. I. (2003). Data mining and knowledge discovery. In *Encyclopedia of Physical Science and Technology (Third Edition)*, pages 229–246. Academic Press, New York, third edition edition.

Mohamed, R., Loukil, Z., and Bouraoui, Z. (2018). Qualitative-based possibilistic el ontology. In *PRIMA 2018: Principles and Practice of Multi-Agent Systems*, pages 552–559, Cham. Springer International Publishing.

Nieves, J. C., Osorio, M., and Cortés, U. (2007). Semantics for possibilistic disjunctive programs. In *LPNMR: 9th International Conference, Tempe, AZ, USA, May 15-17, 2007. Proceedings 9*, pages 315–320. Springer.

Qi, G., Ji, Q., Pan, J. Z., and Du, J. (2011). Extending description logics with uncertainty reasoning in possibilistic logic. *International Journal of Intelligent Systems*, 26(4):353–381.

Safia, B.-B. and Aicha, M. (2014). Poss-OWL 2: Possibilistic extension of OWL 2 for an uncertain geographic ontology. *Procedia Computer Science*, 35:407–416.

Sentz, K. and Ferson, S. (2002). *Combination of evidence in Dempster-Shafer theory*, volume 4015. Sandia National Laboratories Albuquerque.

Straccia, U. (2013). *Foundations of Fuzzy Logic and Semantic Web Languages*. Chapman & Hall/CRC.

Straccia, U. and Bobillo, F. (2017). *From Fuzzy to Annotated Semantic Web Languages*, pages 203–240. Springer International Publishing, Cham.

Sun, S. (2013). A novel semantic quantitative description method based on possibilistic logic. *Journal of Intelligent & Fuzzy Systems*, 25:931–940.

Wang, Y., Chen, Y., Alomair, D., and Khedri, R. (2022). DISEL: A language for specifying DIS-based ontologies. In *(KSEM) 2022*, Lecture Notes in Artificial Intelligence, pages 1–16, Singapore. Springer.

Zadeh, L. A. (1977). Possibility theory vs. probability theory in decision analysis. In *1977 IEEE Conference on Decision and Control including the 16th Symposium on Adaptive Processes and A Special Symposium on Fuzzy Set Theory and Applications*, pages 1267–1269.

Chapter 4

Towards a Cartesian Theory of Ontology Domain Adequacy

This chapter presents the formal theory of domain adequacy, addressing the fifth objective of this thesis: (i) Ensuring the domain adequacy of an ontology through ontological commitment, and (ii) Aligning ontological commitment with data commitment.

This work develops the notion of domain adequacy within the Cartesian framework, with a focus on the DIS model. The domain adequacy is characterized by explicit ontological and data commitments. This chapter formally defines both types of commitments and establishes their associated properties. It explores the interplay between ontological and data commitments and introduces a principled approach for identifying minimal adequate sub-ontologies that are sufficient to ensure domain adequacy. Altogether, this chapter lays the conceptual and theoretical foundation for the plugin system developed in the following chapter.

Towards a Cartesian Theory of Ontology Domain Adequacy

Deemah Alomair[1,2*] and Ridha Khedri[1†]

[1*]Computing and Software, McMaster University, 1280 Main Street West, Hamilton, L8S 4L8, Ontario, Canada.
[2]Computer Information System, Imam Abdulrahman bin Faisal University, King Faisal bin Abdulaziz Street, Dammam, 34212, Eastern Province, Kingdom of Saudi Arabia.

*Corresponding author(s). E-mail(s): alomaird@mcmaster.ca;
Contributing authors: khedri@mcmaster.ca;
†These authors contributed equally to this work.

Abstract

Ontological commitment refers to the entities a theory assumes to exist, whether explicitly stated (e.g., objects, relations) or implied through contextual dependencies. The literature diverges on how such commitments are defined. Authors restrict them to variables bound by extensional quantifiers in bivalent first-order logic, excluding individual constants. By contrast, others hold that ontological commitment arises only through direct reference: when an object is known by acquaintance and explicitly referred to. Despite their differences, both perspectives presuppose a single, objective ontological reality.

This paper extends the concept of ontological commitment to the domain of data analytics by distinguishing between two levels of reality: objective reality, where concepts exist independently of data, and datascape reality, where existence is contingent on specific datasets. It introduces two complementary forms of commitment: ontological commitment, aligning the ontology with conceptual understanding, and data commitment, ensuring compatibility with the data under consideration. Within a formal Cartesian framework grounded in the Domain Information Systems (DIS) model, the paper defines domain adequacy as the satisfaction of both forms of commitment. The proposed approach supports ontology modularization through commitment functions and evaluates how well an ontology reflects both semantic intent and data evidence. This work lays the theoretical foundation for designing ontologies that are both logically coherent and empirically grounded for intelligent systems.

Keywords: Ontological commitment, Data commitment, Domain adequacy, Knowledge representation and reasoning, Modularization.

1 Introduction

1.1 Prelude

Ontologies are representations of entities that are in the world under consideration and that constitute its reality. We call this world domain. Therefore, ontologies represent a cognition of objective reality at its most abstract level. Usually, they are represented as a collection of concepts and the relationships among them. Hence, the fundamental question that should be answered by an ontology is which entities are in the domain under consideration and how they relate to each other [1].

When we contemplate a broad and comprehensive domain, such as that encompassed by *SNOMEDCT* ontology [2] for the medical domain, we find that it involves 367,827 concepts. Working with an ontology of this size is computationally very demanding. Even if we consider only one view of this large ontology, such as *SNOMED_TEST*, which is about medical testing and it involves 4332 concepts [3], the size of the ontology remains significant, and one wonders whether involving all the concepts of an ontology or one of its views is necessary or not. The size alone leads to difficulties in processing or maintaining the ontology due to the complexity arising from the number of concepts and their relations [4]. These large ontologies are commonly referred to as monolithic [5]. Ideally, when we need to perform an ontology-based reasoning on a statement or a query, we should be able to systematically and automatically extract the part of the monolithic ontology that is relevant to our specific needs. The issue is about what ought to guide us in this selection of the relevant part of the domain and how to delimit and extract this relevant part. We need to have the part of our domain/world/ontology that is essential for the proper evaluation of a statement (or query) and its related statements in the domain under consideration. Intuitively, if we take as an example the statement $\exists(x \mid x \in \mathrm{H}uman : \mathrm{D}aughter(x))$[1], it explicitly requires the existence in the domain of a concept *Human* and a concept *Daughter*. Moreover, it implicitly requires the existence of concepts *Parent_1* (e.g., *Mother*) and *Parent_2* (e.g., *Father*). The concepts *Parent_1* and *Parent_2* do not explicitly appear in the statement, but they are related to it due to relationships in the domain and the explicitly given concepts. For instance, the concept *Daughter* is supposedly related to the concept *Mother* by the relationship *IsBirthMotherOf* that is in the ontology of the domain relating *Daughter* and *Mother*. These concepts that are either explicitly or implicitly related to the statement and the relationships among them are commonly referred to as the ontological commitment of the statement under the domain of consideration. One important question is as follows: Are these concepts and their relationships the only ones that are relevant to the considered statement, or are there

[1]Throughout this paper, we adopt the uniform linear notation provided by Gries and Schneider in [6]. The general form of the notation is $\star(x \mid R : P)$ where $\star$ is the quantifier, x is the dummy or quantified variable, R is a predicate representing the range, and P is an expression representing the body of the quantification.

others? Then, we are concerned with how we can get, from a statement and a given ontology, the smallest part of the ontology that includes all the explicit and implicit concepts and their relationships. We refer to this part of the ontology that we get from a statement as the objective reality related to the statement.

From another perspective, if we take the following statement $\exists(x \mid x \in \mathrm{H}uman : \mathrm{S}tudious\_Daughter(x))$ where the concept *Studious_Daughter* is defined as someone who is a daughter and who is scoring 100% in all the courses. Certainly, this statement requires, as indicated above, the existence of concepts *daughter* and all the related concepts and relationships. Hence, it has an objective reality associated with it. Moreover, it has an association with a dataset as it is defined based on some needed data (a list of courses and the grades obtained for these courses). Hence, the existence of this concept *Studious_Daughter* depends on the data. We might have a dataset that indicates that there is no daughter who is scoring 100% in all the courses. Hence, this concept does not exist in reality as given by the data. In this sense, we say that the concept *Studious_Daughter* has to have an objective existence and a data-related existence. The existence of a concept has to be in the objective reality and the datascape reality. Hence, we are referencing the data that underpin the ontology to verify the appropriateness of a particular statement for the domain. This referencing essentially constitutes data commitment, which in turn guarantees the domain-adequacy of the ontology.

From the above, we notice that we have two realities: the objective reality and the datascape reality. When we say that an ontology is "the study of what might exist" [7], we understand from the above that we have an objective existence and a data existence. Therefore, we have two types of concepts: objective concepts and the datascape concept. An objective concept is the concept that, in its definition, does not refer to data, indicating that any dataset does not deny its existence and its existence is ensured independently of the data that is considered. The concept *daughter* presented above does not refer to data and therefore it has by default an existence independently of the dataset under consideration (i.e., objective reality). A datascape concept is the concept that, in its definition, signifies that its existence is contingent upon the data it refers to. In the case of the mentioned concept, *Studious_Daughter*, it is explicitly tied to data, demonstrating its reliance on the dataset being considered for its existence (i.e., datascape reality). Therefore, when we consider a statement, we find that it ought to have a commitment to an objective existence and another to a data existence. Then, when delving into domain adequacy, it is imperative to differentiate between objective commitment and the data commitment as perceived within the dataset landscape.

1.2 Motivations

The motivation for this work arises from the longstanding tension between ontological realism and epistemic access to entities. Ontological commitment, as debated by philosophers, concerns what a theory assumes to exist, whether via quantification over variables or direct reference. However, in practical systems that rely on ontologies, especially in data-driven domains, there is an increasing need to reconcile these abstract commitments with the empirical structure provided by data. The classical view presumes a single layer of reality that ontologies aim to describe. Yet, many real-world systems operate with layered realities: one defined by normative or theoretical constructs, and another shaped by what data captures or omits. For example, while being a "student" may be institutionally defined, determining who qualifies often depends on observable data such as enrollment records. This reveals a deeper philosophical issue: the interplay between what exists in abstract (objective reality) and what is instantiated through empirical evidence (data reality). Such cases motivate a formal framework that respects both dimensions, acknowledging that ontological discourse must sometimes be grounded not only in theory but also in contingent, data-driven representations of the world.

1.3 Contributions

In this paper, we consider a Cartesian world where concepts are formed by a product construction of a given set of atomic concepts. We focus on structured data, represented either as a set of tuples of the same size, such as records in a dataset, or of different sizes, such as entries in log files. In Domain Information System (DIS), each tuple is formed by combining values from defined attributes (sorts), and the space of all possible tuples corresponds to the Cartesian product of these attribute domains. Therefore, we consider the elements of data to lie within a Cartesian space, reflecting that each record is a point determined by its attribute values. Within this world, we present a theory of ontological commitment and data commitment. This Cartesian world allows us to consider datasets that are either lists or tuples, which covers the bulk of datasets available. These datasets are either databases or log files of structured but not necessarily relational data. We use DIS framework [8], presented in Subsection 2.1, which provides a system that is composed of a domain ontology view, a domain data view, and a mapping between them. Our theory gives a mathematical understanding of ontological commitment and data commitment, and it uses them to reduce the size of the ontology. Hence, the obtained ontology could be more efficient when used in tasks as hypothesis verification or conjecture validation. We show that there is a link between ontology modularization [4] and ontological commitment. We advocate using modularization to drive ontological commitment, which is a way to effectively utilize a monolithic ontology by breaking it into smaller components (called modules) through a process of modularization [9, 10]. It brings more efficient query answering and other reasoning tasks [11], and eases the maintenance of ontologies [12]. We find in [13] that modularization helps in the reconciliation of multiple ontologies into one. As a final contribution, we define optimal domain adequacy and elucidate its relationship with ontological and data commitments.

1.4 Structure of the paper

The paper is structured as follows. Section 2 provides an overview of the foundational theories and concepts relevant to our study. Section 3 introduces the notation of the datascape concept. Section 4 presents a comprehensive example of the DIS framework. Section 5 discusses the data and ontological commitments underlying the DIS framework, along with key theorems and properties. Section 6 reviews related work and discusses our findings. Finally, Section 7 offers concluding remarks and directions for future research.

2 Preliminaries

In this section, we introduce the DIS framework, which provides the structural basis for our ontology modelling, and the concept of modularization.

2.1 Domain Information System

DIS [8, 14, 15] is an ontology framework that is formed by three main components: domain ontology $\mathcal{O}$, domain data view $\mathcal{A}$, and a function τ that maps the latter to the former to obtain the final structure $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$.

The first component $\mathcal{O}$ is composed of three elements $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$, where $\mathcal{C}$ is the concept structure whose carrier set C is formed by the empty concept e_c, a collection of atomic concepts, and other composite concepts. Atomic concepts are directly derived from the attributes of the dataset or from the domain knowledge. Composite concepts are unordered tuples of atomic concepts that are obtained by combining atomic concepts using the $\oplus$ operator. The structure $\mathcal{C} = (C, \oplus, e_c)$ is a commutative idempotent monoid.

The second component of the domain ontology is the Boolean lattice $\mathcal{L} = (L, \sqsubseteq_c)$, where $\sqsubseteq_c$ is the natural order defined on the idempotent monoid (i.e., $c_1 \sqsubseteq_c c_2 \iff c_1 \oplus c_2 = c_2$). Hence, $\mathcal{L}$ is a free Boolean lattice that is generated from a given set, T, of atomic concepts from $\mathcal{C}$, and L is the carrier set of $\mathcal{L}$.

The third element of the domain ontology is the set of rooted graphs, where an element $G_{t_i} = (C_i, R_i, t_i)$ is a rooted graph on the set of vertices $C_i \subseteq C$, having the set of edges given by R_i, and is rooted at the vertex $t_i \in L$. The vertices, other than the root of a rooted graph from G_{t_i} are concepts that are not directly generated by the Cartesian product construction of the elements of T (i.e., the set of atomic concepts from which the lattice is formed). In this way, the Boolean lattice generated from the atomic concepts and the rooted graphs presents a more enriched view of the ontological perception that we have from the dataset. The formal definition of domain ontology structure is presented in Definition 1.

Definition 1 (From [15], Domain Ontology Component) Let $\mathcal{C} = (C, \oplus, e_c)$ be a commutative idempotent monoid. Let $\sqsubseteq_c$ be the natural order on $\mathcal{C}$, defined as $\forall(c, d \mid c, d \in C : \text{c} \sqsubseteq_c \text{d} \iff c \oplus d = d)$. Let $\mathcal{L} = (L, \sqsubseteq_c)$ be a free Boolean lattice, generated from a set of atoms in C. Let I be a finite set of indices, and $\mathcal{G} = \{G_{t_i}\}_{t_i \in L, i \in I}$ a set of graphs rooted in L. A domain ontology is the structure $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$. □

The second element of DIS is the domain data view, which is abstracted as diagonal-free cylindric algebra $\mathcal{A} = (A, +, \star, -, 0_A, 1_A, \{c_k\}_{k \in U})$, where U represents the finite set of attributes of the considered dataset. The cylindrification operators c_k are indexed by the sorts used in the data and which correspond to the elements of L, the carrier set of $\mathcal{L}$. The main notion of this view is the notion of `sort`, which corresponds to an attribute in the dataset. The ordered pair of the sort and its value is called *sorted value*, for short, `s_value`. Then, we construct a *sorted datum*, for short, `s_datum`, by a collection of `s_value` with a maximum of one `s_value` for each sort.

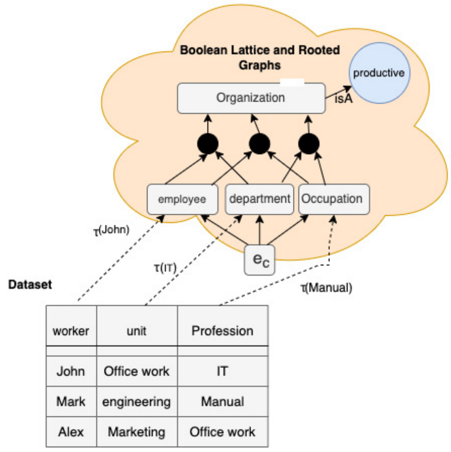


Fig. 1 Abstract View of Organization DIS Framework.

Finally, *sorted data*, for short, `s_data` is a collection of `s_datum`. The cylindric algebra's carrier set, A, is regarded as a collection of `s_data`. The formal definition of data view structure is provided in Definition 2. For more information about cylindric algebra, we refer the reader to [16, 17].

Definition 2 (From [15], Domain Data View Component) Let $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$ be a domain ontology of a DIS $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$. A domain data view associated with $\mathcal{D}$ is a diagonal-free cylindric algebra $\mathcal{A} = (A, +, \star, -, 0_A, 1_A, \{c_k\}_{k \in U})$, where U is the universal set of sorts under consideration. □

The last element of DIS is the mapping operator $\tau : A \to L$, that relates the elements of A to the elements of the Boolean lattice in $\mathcal{O}$. For simplicity and without loss of rigour, we consider that τ maps the sorts of the data to their corresponding objective concepts in $\mathcal{L}$. A partial view of DIS framework corresponding to the organization dataset is presented in Figure 1.

2.2 The language of DIS mathematical structure

In this sub-section, we present the language of DIS, which specifies the syntactical elements, operations, and relations that characterize its structures. It allows us to use the above structures to analyze and reason about them. As introduced above, DIS is a system that involves two mathematical structures: the domain ontology view and the domain data view. The language of the domain ontology is, therefore, the union of the concept language, lattice language, and the relational language coming from the graph structure of the rooted graphs attached to it. The signature of the domain ontology view is $\mathcal{M}_{\mathcal{O}} = (\mathcal{M}_{\mathcal{C}}, \mathcal{M}_{\mathcal{L}}, \mathcal{M}_{\mathcal{G}})$, where $\mathcal{M}_{\mathcal{C}} = (C_{At_{\mathcal{C}}}, \oplus, e_c)$ is the signature of the concept part and $C_{At_{\mathcal{C}}}$ is the set of concepts freely generated by the two operators (composition $\oplus$, and empty concept e_c). The signature of the lattice part $\mathcal{M}_L = (L_{At_\mathcal{L}}, \oplus, e_c, \top_\mathcal{L})$, where $L_{At_\mathcal{L}} \subseteq C_{At_\mathcal{C}}$ that includes the atomic concepts. Also, we have the signature of the relational part of the ontology as given by the rooted graphs. This signature is $\mathcal{M}_\mathcal{G} = (C, R_{t_1}, \cdots, R_{t_n})$ for R_{t_i}, $1 \leq i \leq n$, is a relation of the rooted graph G_{t_i}. The second structure brings an algebraic language coming from cylindric algebra that allows articulating terms representing elements in the data view. The signature of the data view is $\mathcal{M}_A = (A_{(D,U)}, +, \star, -, 0, 1, \{c_k\}_{k\in U})$. The carrier set of the cylindric algebra $A_{D,U}$ is freely generated by all the operators over a set of data elements, called the generating set, which is derived from the universe U. The set of s-datums D may be used to obtain the universe U or vice-versa, as detailed in [15]. In summary, the DIS structure is a many-sorted structure having a many-sorted signature $S = \{C, L, A\}$, where sort C is the sort attributed to the concepts in the domain ontology component. L is the sort of concepts embedded in the Boolean lattice, and A is the sort of domain data view structure. Then, the signature of DIS Σ_{Dis} is defined as follows:

Definition 3 (Borrowed from [15], DIS Signature) Let $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$ be a Domain Information System. The signature Σ_{Dis} is given by the tuple $\Sigma_{Dis} = (\mathcal{F}, \mathcal{R}, r_f, r_r)$, where $\mathcal{F} = \{\oplus, e_c, \top_\mathcal{L}, +, \star, -, 0, 1, \tau, cyl\}$ is the set of function symbols; $\mathcal{R} = \{\sqsubseteq_c, \leq\} \cup \{R_i\}_{i\in I}$ is the set of relation symbols (where I is a set of indeces); and r_f, r_r are the arity mappings, defined as follows:

$$
\begin{gathered}
r_f(\oplus) = (C.C, C) \quad r_r(\sqsubseteq_c) = (L, L) \quad r_f(\tau) = (A, L) \quad r_r(R_i) = (C, C) \\
r_f(-) = (A, A) \quad r_r(\leq) = (A, A) \quad r_f(e_c) = r_f(\top_\mathcal{L}) = (e, L) \\
r_f(cyl) = (A.L, A) \quad r_f(0) = r_f(1) = (e, A) \quad r_f(+) = r_f(\star) = (A.A, A) \quad \square
\end{gathered}
$$

Based on DIS sorts, there are three finitely countable sets of variables: The set $\mathcal{X}^C = \{K, K_1, ..., K_n\}$ denotes the variables of sort C. The set of sort L variables represented by the notation $\mathcal{X}^L = \{k, k_1, .., k_n\}$, and $\mathcal{X}^A$ is the collection of variables of sort A that we represent using the notation $\{a, b, a_1, ..a_n\}$. Then, Let $\mathcal{X} = \{\mathcal{X}^s \mid s \in S\}$ be the S-indexed set. The elements of the set Υ of constants are nullary functions (functions with arity 0), hence Υ is a subset of functions $\mathcal{F}$. Overall, the variables in $\mathcal{X}$, relations in $\mathcal{R}$, and functions in $\mathcal{F}$ provide the non-logical symbols of the DIS language. Moreover, the language provides a set of logical symbols such as $\{\wedge, \vee, \neg, \Rightarrow, \forall, \exists, \mathsf{true}, \mathsf{false}\}$, brackets and parenthesis symbols, and relational symbols. Over this language, the DIS-based expressions (terms and formulae) are constructed inductively. A formal definition of DIS-based terms and formulas is provided in Definition 4. Moreover, if we add to DIS language other operations coming from statistics, we, therefore, augment the set of operations by a set $\mathcal{F}_{stat}$. Hence, terms might involve statistical functions on data values or sets of data values. The statistical operations would allow us to use statistical functions in DIS terms.

Definition 4 (Borrowed from [15], DIS-based Terms and Formulas) Let $\Sigma = (\mathcal{F}, \mathcal{R}, r_f, r_r)$ be a DIS-based signature. Let $\mathcal{X}$ be a set of variables, and a set of constants $\Upsilon \subseteq \mathcal{F}$. We call $T_X(\Sigma)$ the set of X-terms of type Σ defined inductively as follows:

1. Basic terms: $\mathcal{X} \cup \Upsilon \subseteq T_X(\Sigma)$
2. Composite terms: $\forall (t_1, t_2, .., t_{n_f}, f \mid t_1, t_2, .., t_{n_f} \in T_X(\Sigma) \wedge f \in \mathcal{F} : f(t_1, t_2, ..., t_{n_f}) \in T_X(\Sigma))$
3. There are no other terms in $T_X(\Sigma)$.

We call $\Theta(\Sigma)$ the set of formulas of type Σ and define it inductively as follows:

1. Atomic formulas: $\forall (t_1, t_2, .., t_{n_R}, R \mid t_1, t_2, .., t_{n_R} \in T_X(\Sigma) \wedge R \in \mathcal{R} : R(t_1, t_2, ..., t_{n_R}) \in \Theta(\Sigma))$
2. Composite formulas: Let $\diamond \in \{\wedge, \vee, \Rightarrow, \iff\}$ be a logical operator, $\circ \in \{\forall, \exists\}$ a quantifier, $x \in \mathcal{X}$, and $\phi, \psi \in \Theta(\Sigma)$. Then $\neg\phi \in \Theta(\Sigma)$, $\phi \diamond \psi \in \Theta(\Sigma)$, and $\circ(x \mid x \in \mathcal{X} : \phi(x)) \in \Theta(\Sigma)$. □

In DIS, terms and formulas are used to define a concept. The concepts we are distinguishing in the remainder of the paper are defined in the language given by the above signature.

2.3 Ontology Modularization

As we will discuss in Subsection 5.2, ontological commitment leads to a module of the ontology being considered. In this sub-section, we discuss the background related to ontology modularization. A module is a sub-ontology that comprises a subset of the original ontology's concepts and relations. Modularization is the process of retrieving this module from an ontology according to predetermined specifications [4]. In other terms, modularization can be interpreted as a function $m(\mathcal{O})$ that extracts a sub-ontology (i.e., module) $\mathcal{O}_M$ from a given ontology $\mathcal{O}$. The techniques employed for ontology modularization are classified as module extraction or partitioning methods. This classification is based on the type of output produced [18]. Module extraction is the process of building one or more modules that address a desired set of concepts and axioms of an input query. Partitioning, as an alternative, is the process of splitting an ontology into a collection of modules so that the unification of all the modules encompasses the original ontology. Normally, this collection would consist of disjoint modules [19]. However, in certain articles (e.g., [20]), the definition of *partitioning* has been left, in which the disjointness is not a requirement. In some literature, the term *decomposition* is frequently used as another name for the partitioning process [20, 21]. Moreover, approaches to modularization can also be categorized as logical, hybrid, or graphical. For a comprehensive analysis of ontology modularization, we refer the reader to the survey paper [4].

2.4 Ontology Modularization in DIS

In the following, we provide the formal definition of a module. Then, we discuss the view traversal modularization approach of DIS framework. It is important to note that various modularization techniques are explored in the context of the DIS framework [22]. However, our focus in this study is solely on the view traversal modularization technique. Therefore, whenever we refer to the modularization technique in the following sections, we specifically mean view traversal.

Definition 5 (Borrowed from [22], Ontological Module) A module $\mathcal{M}$ of a domain ontology $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$ is the ontology $\mathcal{M} = (\mathcal{C}_M, \mathcal{L}_M, \mathcal{G}_M)$ that satisfies the following:

1. $C_M \subseteq C$
2. $\mathcal{L}_M = (L_M, \sqsubseteq_c)$ such that $L_M \subseteq C_M$, $\mathcal{L}_M$ is a Boolean sub lattice of $\mathcal{L}$, and $e_c \in L_M$
3. $\mathcal{G}_M = \{G_n \mid G_n = (C_n, R_n, t_n) \wedge G_n \in \mathcal{G} \wedge t_n \in L_M\}$,

where C_M and C are the carrier sets of $\mathcal{C}_M$ and $\mathcal{C}$, respectively. □

It is worth noting that rooted graphs of the module include all the rooted graphs in the initial ontology that have roots at one of the nodes of the sublattice $\mathcal{L}_M$ of the module.

We can define simple operations on modules that allow us to combine modules. Let $\mathcal{M}_1 = (\mathcal{C}_{M_1}, \mathcal{L}_{M_1}, \mathcal{G}_{M_1})$ and $\mathcal{M}_2 = (\mathcal{C}_{M_2}, \mathcal{L}_{M_2}, \mathcal{G}_{M_2})$ be two modules of an ontology $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$, where $\mathcal{L}_{M_1}$ (resp., $\mathcal{L}_{M_2}$) is freely generated from the set of atomic concepts T_1 (resp. T_2). We define $\mathcal{M}_1 + \mathcal{M}_2$ as a module $\mathcal{M}$ satisfying the following:

1. $C_M = C_{M_1} \cup C_{M_2}$
2. $\mathcal{L}_M$ is the module that is freely generated from $T_1 \cup T_2$.
3. $\mathcal{G}_M = \mathcal{G}_{M_1} \cup \mathcal{G}_{M_2}$

This operation on modules of an ontology $\mathcal{O}$ is closed on the set of modules of $\mathcal{O}$. Moreover, it is commutative, associative, and has the empty module as its zero.

View traversal is the process of extracting a self-contained sub-ontology from the primary ontology that serves a specific topic [11]. In the view traversal approach, a focal point for creating the module is the identification of the starting concept or group of concepts denoted as C_{st}. Starting from these specified concepts, the module is achieved by forming the principal ideal generated from C_{st}. In the context of lattice theory, the principal ideal is defined as follows:

Let $\mathcal{B}$ be a Boolean algebra with carrier set B, and $b \in B$. The principal ideal generated by b is defined as:

$$L_{\downarrow_b} = \{a \in B \mid a \leq b\}$$

The Principal ideal is a Boolean sub-lattice generated by a set of concepts that encompasses all concepts that are subordinated or part of the starting concept C_{st}. The starting concept might be a concept in the Boolean lattice, a rooted graph concept, or more than one concept. Having established a comprehension of the starting concepts and the principal ideal, we can formulate the extraction of a view traversal module through the modularization function, which is formally defined in Definition 6.

Definition 6 (Borrowed from [22], Modularization Function) Let $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$ be a domain ontology of a DIS $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$. Let the set $\mathcal{M}_D$ include all view traversal modules that are feasible to retrieve from $\mathcal{O}$. The modularization function $\mathsf{m}\Big(\mathcal{O}, c\Big)$ is defined as follows.

$$\mathsf{m} : \mathcal{M}_D \times C \to \mathcal{M}_D,$$

Where $\mathsf{m}(\mathcal{O}, c) = (\mathcal{C}_v, \mathcal{L}_v, \mathcal{G}_v)$, where:

1. $\mathcal{L}_v = (L_{\downarrow_c}, \sqsubseteq_c)$ is a Boolean lattice
2. $\mathcal{G}_v = \{G_i \mid G_i \in \mathcal{G} \wedge t_i \in \mathcal{L}_v\}$
3. $\mathcal{C}_v = \{d \mid d \in L_{\downarrow_c} \vee \exists (G_i \mid \; G_i \in \mathcal{G}_v : d \in C_i\,)\}$. $\square$

As special cases, we have the ontology itself is considered as a view traversal module of $(\mathcal{O}, \top)$, which is exemplified by setting c to $\top$. Conversely, the empty ontology, denoted by $\mathcal{O}_0$, comprises solely the empty concept e_c in its carrier set, is a view traversal module of $(\mathcal{O}, \perp)$. The carrier set of the Boolean lattice of the atomic ontology consists of two elements: an atom, denoted by a, and the empty concept e_c, and involves the rooted graphs rooted at a. Such instances, along with other straightforward scenarios for view traversal, are elucidated in the subsequent Lemma [4].

Lemma 1 Given a domain ontology $\mathcal{O}$, then the following is true:

1. $\mathcal{O} = \mathsf{m}(\mathcal{O}, \top)$
2. $\mathcal{O}_0 = \mathsf{m}(\mathcal{O}, \perp)$
3. $\mathcal{O}_a = \mathsf{m}(\mathcal{O}, a)$ for any atom $a \in L$. $\square$

We can easily prove that for concepts c_1 and c_2 from C. We have

$$\mathsf{m}(\mathcal{O}, c_1 \oplus c_2) = \mathcal{M}_{c_1} + \mathcal{M}_{c_2}, \tag{1}$$

where $\mathcal{M}_{c_1} = \mathsf{m}(\mathcal{O}, c_1)$ and $\mathcal{M}_{c_2} = \mathsf{m}(\mathcal{O}, c_2)$.

3 Datascape Concepts

In Section 1, we discussed that we have two kinds of concepts: the objective ones and the ones that are defined using data elements. We refer to the latter concepts as *datascape concepts*. A datascape is commonly used to indicate a conceptual landscape composed of data that is typically visualized or analyzed to reveal new knowledge. Hence, a datascape concept is a concept from the data landscape, and its definition involves data elements. Formally, we can define a datascape concept as follows:

Definition 7 (Datascape Concept) Let $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$ be a given DIS. Let A be the carrier set of the data view $\mathcal{A}$ and L be the set of the lattice in $\mathcal{O}$. A datascape concept is a concept d that is defined as follows:

$$d \stackrel{\text{def}}{=} \{a \;\mid\; \tau(a) \in L \wedge \Phi(a)\}, \text{ where}$$

$a \in A$, and Φ is a formula given in Disjunctive normal form (DNF) as $\Phi(a) = \vee\,(i \mid \; 1 \leq i \leq N : \Psi_i(a)\,)$, with N is a natural number, and $\Psi_i(a) = \wedge\,(j \mid \; 1 \leq j \leq M : \Omega_{(i,j)}(a)\,)$, where M is a natural number and $\Omega_{(i,j)}(a) = (f_{(i,j)}(a \cdot \mathtt{sort\_name}_{(i,j)}), c_{(i,j)}) \in R_{(i,j)}$, where $f_{(i,j)} \in \mathcal{F}$, $c_{(i,j)}$ is a ground term in DIS language, and $R_{(i,j)}$ is a relator. We refer to Φ as a data specializing predicate, expressed within the formal language of $\mathcal{D}$. $\square$

Φ is also commonly referred to as a conjunctive query (e.g., [23, 24]). The condition $\tau(a) \in L$ ensures that the data that we are considering is relevant to the ontological view. The sorts used in defining a are linked to the lattice of objective concepts $\mathcal{L}$. In the remainder of the paper, we assume that all data specializing predicates are within the language of the DIS under consideration, augmented with statistical functions. Hence, Definition 7 underpins the syntactical structure of a datascape concept.

Most of the time, datascape concepts are used to specialize objective concepts. For instance, the datascape concept of *Studious_Daughter*, which is presented in Section 1, is a specialization of the objective concept *Daughter*. The specialization is by characterizing, using a predicate Φ, a subgroup of objective concepts using data elements.

Based on the above definition, we define the operation $\oplus$ introduced in Subsection 2.1 as an operation on concepts.

Definition 8 Let $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$ be a given DIS. Let $d_1 = \{a \mid \tau(a) \in L \wedge \Phi_{d_1}(a)\}$, and $d_2 = \{a \mid \tau(a) \in L \wedge \Phi_{d_2}(a)\}$ be two datascape concepts defined on $\mathcal{D}$. We have

$$d_1 \oplus d_2 = d_1 \cup d_2 = \{a \mid \tau(a) \in L \wedge (\Phi_{d_1}(a) \vee \Phi_{d_2}(a))\}.$$

The structure of the $d_1 \oplus d_2$ is that of a datascape as $(\Phi_{d_1}(a) \vee \Phi_{d_2}(a))$ is in DNF and the other conditions stipulated by Definition 7 are satisfied. Moreover, the empty concept e can be perceived as a datascape concept defined as $e = \{a \mid \tau(a) = e_c \wedge \mathsf{false}\} = \emptyset$. Hence, if we take, for a given DIS, C_d is the set of datascape concepts, then $(C_d, \oplus, e_c)$ is a commutative monoid due to the properties of set union.

4 Illustrative Example of DIS

We consider an `organization` dataset schema that consists of the five attributes: worker (`wkr`), profession (`prof`), financial (`fincl`), `unit`, and `task`. This means that the concept `organization` is defined as the join of all those attributes: `organization` = `wkr` $\oplus$ `prof` $\oplus$ `fincl` $\oplus$ `unit` $\oplus$ `task`. The construction of the DIS structure $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$ from this dataset is achieved according to the steps below:

Building the lattice of core concepts. We take the dataset attributes and map them to a set of atomic concepts. The attribute `wkr` is mapped to the atomic concept `emp`, `prof` is mapped to the atomic concept `occ`, `fincl` is mapped to `finc`, `unit` is mapped to `dep`, and `task` is mapped to `proj`. Hence, $\tau = \{$(`wkr`, `emp`), (`prof`, `occ`), (`fincl`, `finc`), (`unit`, `dep`), (`task`, `proj`)$\}$. The mapping τ can be, in some cases, the identity function. The set of obtained atomic concepts is used to generate a complete Boolean lattice of concepts as illustrated by Figure 2a. The concepts of the lattice are a subset of the set of concepts $\mathcal{C}$ that is augmented by the concepts of the next step. The mapping τ allows us to translate an attribute into a concept.

Building the family of rooted graphs of the ontology. The lattice alone, while it brings all the concepts obtained from the attributes of the dataset, is not enough to capture all concepts of the domain that are related to the dataset. The domain knowledge allows us to enrich the set of concepts with other concepts that are related to the

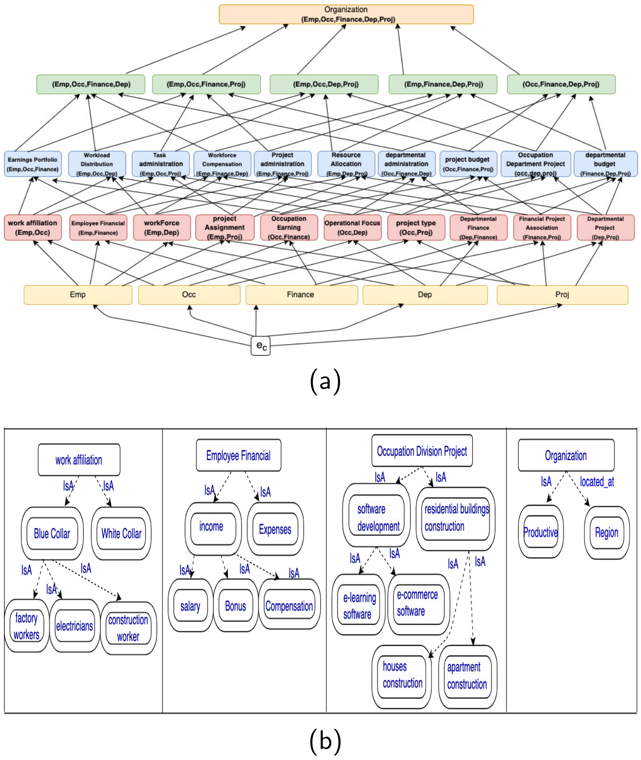


Fig. 2 (a) Constructed Lattice for Organization Ontology. (b) Rooted Graphs Attached to Several Lattices' Concepts.

concepts of the lattice and therefore to the dataset. These concepts and their relationships form the rooted graphs. In our case, rooted graph concepts can be constructed using three distinct approaches.

The first approach is to define rooted graph concepts abstractly based on objective reality, without using the underlying data in their definition, or, in other terms, without providing a mathematical construction for them.

The second approach, referred to as defining datascape concepts, encompasses data constraints within the definition of a rooted graph concept. For instance, a rooted graph concept representing "blue collar" could include specific occupational types such as manual labor. Formally, it can be expressed as follows: $\texttt{Blue\_collar} \stackrel{\text{def}}{=} \{a \mid \tau(a) = \texttt{work\_affiliation} \wedge a.\texttt{prof} = \text{"manual"}\}$. This concept is a specialization of the lattice concept `work_affiliation`, which includes only the workers who hold manual occupations. Another illustration of a datascape concept is the `High_Income_Wrkplce`. This datascape concept is applicable only when the average financial status of the organization exceeds a specified threshold. Formally, it is expressed as follows:

$$\texttt{High\_Income\_Wrkplce} = \{a \mid \tau(a) = \texttt{organization} \wedge \text{average}(\text{a.}\texttt{fincl}) > \text{threshold}\}$$

The third approach involves constructing rooted graph concepts based on other rooted graph concepts, which also constitute a specialization of the latter concepts. For example, starting with a general rooted graph including the `Blue_collar` concept, one can derive further specialized concepts such as `Factory_worker`, `Electrician`, or `Construction_worker`.

In the above examples, some rooted graph concepts are constructed using statistical or aggregation operators. The utilization of these operators introduces complexities during later stages of data validation, necessitating additional analysis to ensure the data relevance of these concepts. Some rooted graphs attached to several lattice concepts are presented in Figure 2b.

Construction of the domain data view. In this example, we have five sorts `wkr`, `prof`, `fincl`, `unit`, and `task`. An example of `s_value` is (`unit`, "IT"). An example of `s_datum` is $dt_1 = \{(\texttt{wkr}, \text{“John”}), (\texttt{prof}, \text{“office\_job”}), (\texttt{unit}, \text{“IT”}) (\texttt{task}, \text{“development”})\}$. An example of `s_data` is $a = \{dt_1, dt_n\}$. We denote the set of s-data as A on which we can have a set of variables $\mathcal{X}^A$. The structure $\mathcal{A} = (A, +, \star, -, 0_A, 1_A, \{c_k\}_{k \in U})$ forms a diagonal-free cylindric algebra and gives us the data view $\mathcal{A}$.

Building the whole DIS system. In our case, the DIS is then formed by $(\mathcal{O}, \mathcal{A}, \tau)$.

The DIS constructed in the above steps has a language that allows us to define terms and formulas. For instance, the constant `wkr` is a term. Forming composite terms involves combining one or more basic terms with an associated function $f \in \mathcal{F}$. For example, we can define a term that uses the language of the ontology and the language of the data view, such as having

$$\tau(a) = \texttt{work\_affiliation} \wedge a.\texttt{prof} = \text{“desk-based”}$$

One can use this term to define a concept

$$\texttt{White\_collar} \stackrel{\text{def}}{=} \{a \mid \tau(a) = \texttt{work\_affiliation} \wedge a.\texttt{prof} = \text{“desk-based”}\}$$

In DIS we can use the relators of the language to combine terms into a formula. For example, we can use the relator $\sqsubseteq_c$ to relate the terms `Proj_Budget` and `budget` to get the basic formula $\texttt{Proj\_Budget} \sqsubseteq_c \texttt{budget}$. An example of a composite formula is obtained by relating composite terms with a relator from the language of DIS. For instance, the concept of `High_Income_Wrkplce`, defined earlier as a datascape concept, is regarded as being defined involving a composite formula.

5 Data commitment and ontological commitment

Given a statement or a formula that we need to perform a reasoning task on, one obvious question that arises is what would be the necessary and sufficient ontological world needed to reason on it. This ontology would include only the necessary concepts and their relationships. This minimality of the ontology brings computational benefits compared to using the whole available ontology. The objective of attaining a small adequate ontology for a formula is, as we show below, achieved using data commitment and ontological commitment of the formula. Verifying the data commitment of a formula leads to the elimination or preservation of irrelevant datascape concepts and the concepts that they specialize. We eliminate a concept when the underlying data does not support its existence. Then, through the ontological commitment, we obtain the minimal sub-ontology that suffices for the reasoning on the formula.

5.1 Data Commitment and Its Usage

Demonstrating data commitment of a formula within the DIS framework entails substantiating the presence of its components within the domain. Specifically, when evaluating the data commitment of a given formula, we are verifying the legitimacy of the incorporated data within that formula according to a specific domain. The data commitment is a function that we denote by P_A, and that evaluates the formula concerning the domain data-view component (i.e., dataset). The evaluation function yields a Boolean value. The returned value indicates the confirmation that the data under consideration supports the existence of the characterized concept, or not. Datascape concepts that their specializing predicates evaluate to false should be omitted from the ontology. Conversely, datascape concepts whose specializing predicates evaluate to true should be included in the subsequent ontology. The formal definition of the data commitment function for a formula within the DIS framework is provided by Definition 9.

Definition 9 (Data Commitment Function) Let $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$ be a given DIS. Let A be the carrier set of component $\mathcal{A}$. Let $a \in A$, Φ be a data specializing predicate as defined in Definition 7. We define the formula data commitment relative to A, that we denote by $P_A(\Phi(a))$, as follows:

- Base cases:
  1. $P_A(\Omega_{(i,j)}(a)) = \exists\Big(b \mid b \in A : \tau(b) \sqsubseteq \tau(a) \wedge (f_{(i,j)}(b \cdot \texttt{sort\_name}_{(i,j)}), c_{(i,j)}) \in R_{(i,j)}\Big)$
  2. $P_A(\lambda(a)) = \texttt{true}$, for any $\lambda(a)$ that is not of the structure of $\Omega_{(i,j)}(a)$;
- Inductive cases:
  1. $P_A(\Psi_i(a)) = \wedge\Big(j \mid 1 \leq j \leq M : P_A(\Omega_{(i,j)}(a))\Big)$
  2. $P_A(\Phi(a)) = \vee\Big(i \mid 1 \leq i \leq N : P_A(\Psi_i(a))\Big)$ □

Definition 9 delimits the data commitment of a specific formula Φ. In the initial base scenario of the data commitment function, the formula $\Omega_{(i,j)}(a)$ assigns a value $c_{(i,j)}$ to $\texttt{sort\_name}_{(i,j)}$ utilizing one of the relators $r \in R_{(i,j)}$. For example, $\Omega_{(i,j)}(a) = average(a.\texttt{Salary} \geq 16000)$. It can be reformulated as: $\Omega_{(i,j)}(a) = (average(a.salary), 16000) \in \geq$. In part two of the base cases, the formula is either $\texttt{true}$, like in the objective concept or $\texttt{false}$, such as for e. As a result, we regard the data commitment in these cases as true. The inductive part puts the results together from all the elementary formulas that form Ψ_i and Φ.

In the following, we consider several valid claims corresponding to the data commitment regarding a given DIS.

Claim 1 Let Φ_1 and Φ_2 be two data specializing predicates defining two datascape concepts d_1 and d_2. Then $P_A(\Phi_1(a) \vee \Phi_2(a)) = P_A(\Phi_1(a)) \vee P_A(\Phi_2(a))$.

Proof $\Phi_1(a)$ and $\Phi_2(a)$ are both in DNF. Since Φ_1 and Φ_2 are two data specializing predicates defining two datascape concepts d_1 and d_2, then we can define them as follows:

- $\Phi_1 \overset{\text{def}}{=} \vee\,(i \mid\ 1 \le i \le N : \Psi_i\,)$, and
- $\Phi_2 \overset{\text{def}}{=} \vee\,(i \mid\ (N+1) \le i \le P : \Psi_i\,)$ for some N and P in $\mathbb{N}$.

$$
\begin{array}{ll}
& P_A(\Phi_1 \vee \Phi_2) \\
= & \qquad\qquad \langle \text{ Definitions of } \Phi_1 \text{ and } \Phi_2 \ \rangle \\
& \quad P_A\Big(\ \vee(i \mid\ 1 \le i \le N : \Psi_i\,) \\
& \qquad\qquad \vee \ \ (\,\vee\,(i \mid\ (N+1) \le i \le P : \Psi_i\,))\Big) \\
= & \qquad\qquad \langle \text{ Quantifier range split } \rangle \\
& \quad P_A\Big(\ \vee\,(i \mid\ 1 \le i \le P : \Psi_i\,)\Big) \\
= & \qquad\qquad \langle \text{ Definition 9 } \rangle \\
& \quad \wedge \Big(i \mid 1 \le i \le P : P_A(\Psi_i)\Big) \\
= & \qquad\qquad \langle \text{ Quantifier range split } \rangle \\
& \quad \wedge \Big(i \mid 1 \le i \le N : P_A(\Psi_i)\Big) \\
& \qquad \vee \wedge \Big(i \mid (N+1) \le i \le P : P_A(\Psi_i)\Big) \\
= & \qquad\qquad \langle \text{ Definition 9; Definition of } \Phi_1 \text{ and } \Phi_2 \ \rangle \\
& \quad P_A(\Phi_1) \vee P_A(\Phi_2)
\end{array}
$$

□

Definition 10 (Data Commitment of a Datascape Concept) Let $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$ be a given DIS. The data commitment of the datascape concept $d = \{a \mid \tau(a) \in L \wedge \Phi(a)\}$, denoted by $\mathcal{C}_A(d)$, is $P_A(\Phi(a))$. □

Claim 2 The data commitment of an Objective_concept is true.

Proof

$$
\begin{array}{ll}
& \quad \mathcal{C}_A(\text{Objective\_concept}) \\
= & \qquad\qquad \langle \text{ Definition 10 } \rangle \\
& \quad P_A(\Phi) \\
= & \qquad\qquad \langle \text{ In an objective concept, we have } \Phi(a) \equiv \mathsf{true}, \text{ for every } a \in A \ \rangle \\
& \quad P_A(\mathsf{true}) \\
= & \qquad\qquad \langle \text{ By Definition 9-Case 2 } \rangle \\
& \mathsf{true}
\end{array}
$$

□

In the claim above, we introduce the notion of data commitment to an objective concept. In the case of an objective concept, we consider $\Phi = \texttt{true}$, which is one of the base cases of the data commitment function to a formula as given previously in Definition 9.

Claim 3 Let d_1 and d_2 be datascape concepts defined in a given DIS. We have $\mathcal{C}_A(d_1 \oplus d_2) = \mathcal{C}_A(d_1) \vee \mathcal{C}_A(d_2)$.

Proof The concepts d_1 and d_2 are two datascape concepts. Hence, by Definition 7 and for $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$ is a given DIS, we can write d_1 and $_2$ as follows:

- $d_1 \stackrel{\text{def}}{=} \{a \mid \tau(a) \in L \wedge \Phi_1(a)\}$
- $d_2 \stackrel{\text{def}}{=} \{a \mid \tau(a) \in L \wedge \Phi_2(a)\}$

$$
\begin{aligned}
& \mathcal{C}_A(d_1 \oplus d_2) \\
= & \qquad\qquad \langle \text{ Definition 8; Definition of set union } \rangle \\
& \{a \mid \tau(a) \in L \wedge \Phi_1(a) \quad \vee \quad \tau(a) \in L \wedge \Phi_2(a)\} \\
= & \qquad\qquad \langle \text{ Definition 10 } \rangle \\
& P_A(\Phi_1 \ \vee \ \Phi_2) \\
= & \qquad\qquad \langle \text{ Claim 1 } \rangle \\
& P_A(\Phi_1) \ \vee \ P_A(\Phi_2) \\
= & \qquad\qquad \langle \text{ Definition 10 } \rangle \\
& \mathcal{C}_A(d_1) \ \vee \ \mathcal{C}_A(d_2)
\end{aligned}
$$

□

In the following, we revisit the example presented in Section 4 to gain a better understanding of the data commitment.

Example 1 Let us consider our interest in reasoning about the datascape concept `Soft_dvlp`, defined as follows:

$$
\begin{aligned}
\texttt{Soft\_dvlp} \stackrel{\text{def}}{=} \{a \mid \tau(a) = \texttt{Occ\_Dep\_Proj} \wedge a.\texttt{unit} = \text{“IT”} \wedge a.\texttt{task} = \text{“Development”} \\
\wedge\ a.\texttt{prof} = \text{“desk-based”}\}
\end{aligned}
$$

Subsequently, we extract the data specializing predicate of `Soft_dvlp` expressed as follows:

$$\Phi(\texttt{Soft\_dvlp}) = (a.\texttt{unit} = \text{“IT”} \wedge a.\texttt{task} = \text{“Development”} \wedge a.\texttt{prof} = \text{“desk-based”})$$

Let us assume that, in our case, the data commitment function $P_A(\Phi(\texttt{Soft\_dvlp}))$ returns `true`, affirming the existence of data supporting the validity of the above formula. Hence, the `Soft_dvlp` concept and its specialized concepts (i.e., `E-learning_software`, `E-commerce_software`) should be included in the ontology. On the other hand, if we assume that in our scenario, there is no data that supports the validity of this formula and $P_A(\Phi(\texttt{Soft\_dvlp}))$ returns `false`. Then, the concept of `Soft_dvlp` and its specialized concepts are excluded from the ontology under consideration.

5.2 Ontological Commitment

Within the DIS framework, the ontological commitment of a concept entails identifying and extracting the smallest module required for effective reasoning on queries and statements. This module encompasses all pertinent concepts and their related concepts, determined by the ontological commitment function $\mathcal{C}_O$. The $\mathcal{C}_O$ function employs a recursive function, `MapToConcept`, to determine the composite concept corresponding to the maximum concept in the lattice that encompasses all the needed objective concepts.

The ontological commitment function $\mathcal{C}_O$ operates based on the composite concept returned from the $\texttt{MapToConcept}(\Phi(a))$function. A formal definition of $\texttt{MapToConcept}(\Phi(a))$ is outlined in Definition 11, followed by a formal definition of the ontological commitment function for a datascape concept.

Definition 11 (Mapping Sorts to Concepts Function) Let $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$ be a domain ontology of a DIS $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$. Let Φ be a data specializing predicate. Let Γ be the set of data specializing predicates.

We define the function $\texttt{MapToConcept}: \Gamma \longrightarrow C$, as follows:

- Base cases:
  (a) $\texttt{MapToConcept}\Big(\Omega_{(i,j)}\Big) = \tau(\texttt{sort\_name}_{(i,j)})$, for $1 \leq i \leq N \wedge 1 \leq j \leq M$;
  (b) $\texttt{MapToConcept}\Big(\lambda\Big) = e_c$, for any λ that is not of the structure of $\Omega_{(i,j)}$;
- Inductive cases:
  (c) $\texttt{MapToConcept}(\Psi_i) = \tau(\texttt{sort\_name}_{(i,M)}) \oplus \texttt{MapToConcept}\Big(\wedge\,(j \mid \; 1 \leq j \leq (M-1) : \Omega_{(i,j)}\,)\Big)$;
  (d) $\texttt{MapToConcept}(\Phi) = \texttt{MapToConcept}(\Psi_N) \oplus \texttt{MapToConcept}\Big(\vee\,(i \mid \; 1 \leq i \leq (N-1) : \Psi_i\,)\Big)$.

$\square$

This function takes a formula that has the structure given in Definition 7 and returns the composite concept that is formed by all atomic concepts relevant to the formula. Its base case (a) deals with elementary literals that do not involve compound operations like conjunction or disjunction, but have the prescribed structure. The base case (b) considers terms that do not abide by the prescribed structure. The cases where $\Phi(a)$ is either true or false are special cases of case (b). Cases (c) and (d) are for handling conjunctive and disjunctive predicates, respectively. Due to the symmetry and transitivity of $\oplus$, the order in which we split a conjunction (or disjunction) into a literal and a smaller (formed with fewer constructing literals) conjunction (or disjunction) does not matter. The idempotence of $\oplus$ ensures that the sort names present in many conjunctions will be present only once in the concept returned by the function `MapToConcept`.

The ontological commitment function of a concept makes usage of the `MapToConcept` function defined above, and the modularization technique. Hence, the utilization of a modularization technique becomes imperative for constructing a sub-ontology or sub-lattice rooted at the target concept. This sub-lattice encompasses all associated sub-concepts and validated datascape concepts to the target concept. The formal definition of the ontological commitment function to a datascape concept is as follows.

Definition 12 (Ontological Commitment of a Datascape Concept) Let $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$ be a domain ontology of a DIS $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$. Let

$$\mathsf{m} : \mathcal{M}_D \times C \to \mathcal{M}_D,$$

be a modularization function, where $\mathcal{M}_D$ is a set of modules (i.e., ontologies). Let d be a datascape concept such that $d \stackrel{\text{def}}{=} \{a \mid \tau(a) \in L \wedge \Phi(a)\}$, and $\forall(a \mid a \in d : \tau(a) \in L)$. The ontological commitment of d, denoted by $\mathcal{C}_O(d)$, is $\mathsf{m}\Big(\mathcal{O},\ \oplus(a \mid a \in d : \tau(a)) \oplus \texttt{MapToConcept}(\Phi)\Big)$. □

In the following, we consider $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$ be a domain ontology of a DIS $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$, and $\mathsf{m} : \mathcal{M}_D \times C \to \mathcal{M}_D$, is a modularization function, where $\mathcal{M}_D$ is a set of modules obtained from $\mathcal{O}$.

Claim 4 Let d be an objective concept such that $d = \{a \mid \tau(a) \in L\}$. The ontological commitment of d is

$$\mathsf{m}\Big(\mathcal{O},\ \oplus(a \mid a \in d : \tau(a))\Big).$$

Proof

$$\begin{aligned}
& \mathcal{C}_O(d) \\
= \quad & \langle \text{ Definition 12 } \rangle \\
& \mathsf{m}\Big(\mathcal{O},\ \oplus(a \mid a \in d : \tau(a)) \oplus \texttt{MapToConcept}(\Phi)\Big) \\
= \quad & \langle \text{ Objective_concept} \stackrel{\text{def}}{=} \{a \mid \tau(a) \in L\}. \text{ Hence } \Phi(a) \equiv \mathsf{true} \ \rangle \\
& \mathsf{m}\Big(\mathcal{O},\ \oplus(a \mid a \in d : \tau(a)) \oplus \texttt{MapToConcept}(\mathsf{true})\Big) \\
= \quad & \langle\ \texttt{MapToConcept}\Big(\lambda(a)\Big) = e_c, \text{ if } \lambda(a) \text{ is not of the structure of } \Omega_{(i,j)}(a). \\
& \quad \text{In this case } \lambda(a) \equiv \mathsf{true} \ \rangle \\
& \mathsf{m}\Big(\mathcal{O},\ \oplus(a \mid a \in d : \tau(a)) \oplus e_c\Big) \\
= \quad & \langle\ e_c \text{ is the zero for } \oplus \ \rangle \\
& \mathsf{m}\Big(\mathcal{O},\ \oplus(a \mid a \in d : \tau(a))\Big)
\end{aligned}$$

□

In Definition 11, when Φ is true or false, it defines an objective concept d. In this case, the ontological commitment of d corresponds to the view travel module originating from the supremum in the lattice L of the ontology, derived from the mapping τ of its values. When $\Phi = \mathsf{true}$, it indicates an objective concept. Conversely, if $\Phi(a) = \mathsf{false}$, the concept is the empty concept e.

Claim 5 Let c_1 and c_2 be concepts from C. We have $\mathcal{C}_O(c_1 \oplus c_2) = \mathsf{m}(\mathcal{O}, c_1 \oplus c_2) = \mathcal{M}_{c_1} + \mathcal{M}_{c_2}$, with $\mathcal{M}_{c_1} = \mathsf{m}(\mathcal{O}, c_1)$ and $\mathcal{M}_{c_2} = \mathsf{m}(\mathcal{O}, c_2)$.

Proof Let c_1 and c_2 be two concepts having as data specializing predicates Φ_1 and Φ_2, respectively.

$$
\begin{aligned}
& \mathcal{C}_O(c_1 \oplus c_2) \\
= & \qquad\qquad \langle \text{ Definition 12 and using Definition 8 } \rangle \\
& \mathsf{m}\Big(\mathcal{O}, \oplus(a \mid \ a \in (c_1 \oplus c_2) : \tau(a)\,) \\
& \qquad \oplus \texttt{MapToConcept}(\Phi_1 \vee \Phi_2)\Big) \\
= & \qquad\qquad \langle \text{ Definitions of } c_1 \text{ and } c_2 \ \rangle \\
& \mathsf{m}\Big(\mathcal{O}, (c_1 \oplus c_2)\Big) \\
= & \qquad\qquad \langle \text{ Equation 1 with } \mathcal{M}_{c_1} = \mathsf{m}(\mathcal{O}, c_1) \text{ and } \mathcal{M}_{c_2} = \mathsf{m}(\mathcal{O}, c_2) \ \rangle \\
& \mathcal{M}_{c_1} + \mathcal{M}_{c_2}
\end{aligned}
$$

□

We recall that when we define a concept c, we identify the sorts that are used to define it. Then, we use the function `MapToConcept` to get the concept k in the ontology that is the supremum of the images of each of the sorts defining c. Finally, the ontological commitment of c concerning the ontology $\mathcal{O}$ is the module obtained by view-traversal on concept k. Due to the strong link between ontological commitment and modularization, several results can be established for ontological commitment that come directly from the properties of modularization. Many of the results obtained for the modularization using view-traversal [4] are transportable to ontological commitment. For instance, the following results are valid for ontological commitment.

Claim 6 Let $\mathcal{C}_O(k) = (\mathcal{C}_k, \mathcal{L}_k, \mathcal{G}_k)$ be the ontological commitment of a concept k with regard to the ontology $\mathcal{O}$. Then we have $\forall(c \mid \ c \in \mathcal{L}_k : \mathrm{c} \sqsubseteq_c \mathrm{k}\,)$.

Proof $\mathcal{C}_O(k)$ returns a module, where its lattice has k as its maximum element. Hence, every concept element of the lattice of the ontological commitment module is below k in the lattice $\mathcal{L}_k$. □

In Claim 6, by abuse of notation, we write $c \in \mathcal{L}_k$ to indicate c in the carrier set L_k of the lattice $\mathcal{L}_k$. The claim indicates that if a concept c is in the lattice of the ontology obtained by the ontological commitment of the concept k, then c must be a part of the concept k. The counter positive of the conclusion of Claim 6 is $\forall(c \mid \ c \in C : \neg(\mathrm{c} \sqsubseteq_c \mathrm{k}) \implies \neg(c \in \mathcal{L}_k)\,)$, which equivalently states that if a concept c is not a part of the Cartesian construction of concept k, then it will not be in the lattice of the obtained module and consequently not in the ontological commitment module.

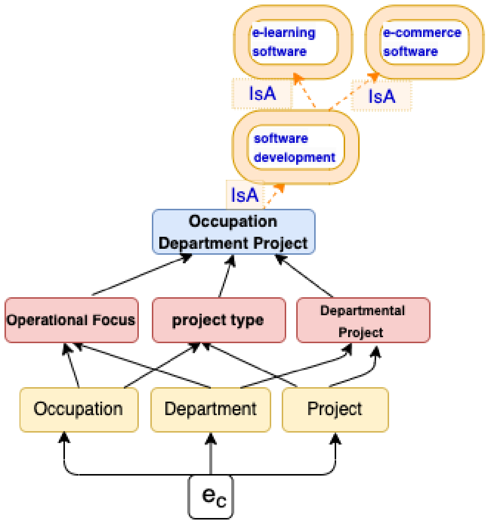


Fig. 3 Constructed Module for `Occ_Dep_Proj` Concept.

Claim 7 (Ontological Commitment Composition) Let $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$ be a domain ontology of a DIS. Let c_1, c_2 be two concepts in $\mathcal{L}$ such that $c_2 \sqsubseteq_c c_1$.

$$\mathcal{C}_O(\mathcal{C}_O(\mathcal{O}, c_1), c_2) = \mathcal{C}_O(\mathcal{O}, c_2)$$

Proof The ontological commitment $\mathcal{C}_O(\mathcal{O}, c_1)$ is $\mathcal{M}_{c_1} = (\mathcal{C}_{c_1}, \mathcal{L}_{c_1}, \mathcal{G}_{c_1})$. The lattice $\mathcal{L}_{c_1}$ has c_1 as its maximum (top) element. Then, the ontological commitment of c_2 considering the ontology $\mathcal{M}_{c_1}$ (i.e., $\mathcal{C}_{\mathcal{M}_{c_1}}(c_2)$) we obtain a module $\mathcal{M}_{c_2}$. Since, we have $c_2 \sqsubseteq_c c_1$ them we have $\mathcal{M}_{c_2}$ is a sub-module of $\mathcal{M}_{c_1}$, which is also a sub-module of $\mathcal{O}$. $\mathcal{M}_{c_2}$ is equal to $\mathcal{C}_O(\mathcal{O}, c_2)$. A detailed proof would invoke minor results from [4]. □

Claim 7 indicates that if c_2 is a constituent of c_1 (via a Cartesian construction), then the ontological commitment of c_2 can be obtained by first having the ontological commitment of c_1, then we use the obtained module as our main ontology in which we look for the ontological commitment of c_2.

In the following, we revisit the example presented in Section 4 to gain a clearer understanding of the ontological commitment function.

Example 2 Let us consider our interest in reasoning about the datascape concept `Soft_dvlp`, which is a rooted graph concept that specializes the lattice concept `Occ_Dep_Proj`. Consequently, the extracted module from $\mathcal{C}_O$ function should be rooted at `Soft_dvlp`, encompassing all its sub-concepts and valid datascape concepts associated within it (i.e., data commitment of datascape concepts have been validated in Example 1). In this example, the ontological commitment function $\mathcal{C}_O$ proceeds to extract the sufficient module to reason with the concept `Soft_dvlp` utilizing the modularization. The extracted module is depicted in Figure 3.

5.3 Using Data Commitment and Ontological Commitment

Ontological commitment and data commitment are distinct notions. A concept that has an ontological commitment different than the trivial empty ontology $\mathcal{O}_0$ does not necessarily imply that its data commitment is valid (i.e., equivalent to $\mathtt{true}$), and vice versa. These two notions pertain to different realms, and asserting that one does not provide evidence for the other. Furthermore, there are cases where two concepts share the same ontological commitment (i.e., they comprise the same set of atomic concepts), yet one may demonstrate data commitment while the other does not, contingent upon the constraints imposed by the formulation of the concepts. For instance, consider the concepts $c_1 = \{a \mid \tau(a) = \mathtt{work\_affiliation} \wedge a.\mathtt{prof} = \text{“manager”}\}$, and $c_2 = \{a \mid \tau(a) = \mathtt{work\_affiliation} \wedge a.\mathtt{prof} = \text{“senior\_Manager”}\}$. Both c_1 and c_2 share the same set of concepts (i.e., $\{\mathtt{emp}, \mathtt{occ}, \mathtt{work\_affiliation}\}$), indicating identical ontological commitments to the domain. However, because c_2 only retrieves the subset *“Senior_manager”* from the concept *“manager”*, it might result in a different data commitment based on the given dataset. If the dataset does not include instances where $\mathtt{occ} = \text{“Senior\_manager”}$, then this concept lacks data commitment.

Definition 13 Let $\mathcal{O}$ be a domain ontology of a DIS $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$. An ontology $\mathcal{O}$ satisfies domain adequacy for a set of concepts A with regards to $\mathcal{D}$ iff for every $c \in A$, we have $\mathcal{C}_A(c)$ is valid, and $\mathcal{C}_O(c)$ is a non-trivial ontological module of $\mathcal{O}$.

Having $\mathcal{C}_A(c)$ for a concept in A indicate that there is data in the dataset under consideration related to it; it means that c is relevant and hence can be kept in the ontology. The second condition indicates that the ontological commitment of c is not the empty concept, and there is a module of $\mathcal{O}$ that includes only the relevant concepts that are needed to reason on c. This module makes as few claims as possible about the world being modelled and sufficient to reason of c.

In Definition 14, we introduce a stronger notion of domain adequacy, which is the notion of optimal domain adequacy.

Definition 14 Let $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$ be a domain ontology of a DIS $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$. An ontology $\mathcal{O}$ satisfies optimal domain adequacy for a set of concepts Q with regards to $\mathcal{D}$ iff it satisfies domain adequacy for Q and $\mathcal{O} = +(c \mid c \in Q : \mathcal{C}_O(c))$.

To satisfy the optimal domain adequacy, in addition to satisfying the domain adequacy for Q, the ontology $\mathcal{O}$ must be exactly the summation of all the ontological commitments of every element of Q. It is optimal in the sense that it contains exactly only the concepts that have data relevance and belong to the ontological commitment of one of the concepts of Q and no more, no less.

If the maximum element $\top$ of the lattice of $\mathcal{O}$ is in Q and $\mathcal{O}$ has the property of domain adequacy, then $\mathcal{O}$ has the optimal domain adequacy for the set of concepts Q with regards to $\mathcal{D}$. This is due to Lemma 1-Case(1) and the idempotence of the summation of modules.

6 Related Work and Discussion

In the literature, views on ontological commitment vary. A prominent perspective is Quine's view [25, 26], known as knowledge by description, which confines ontological commitment to variables bound by extensional quantifiers in bivalent first-order logic, excluding individual constants, (e.g., $\exists(x \mid \; x : x = \text{student} \wedge x = \text{Jhon}))$. Here, the bounded variable x is ontologically committed to its world since $\exists(x)$ admits its existence. In contrast, Barcan Marcus [27] advocates a view where ontological commitment is through direct reference (i.e., direct verifying to an object should be made by an individual), requiring objects to be known by acquaintance rather than description. This divergence is reconciled in [28], which proposes a hybrid formal language integrating both Quine and Marcus perspectives through a first-order logic framework with variables and individuals.

Although Quine and Marcus' views differ, they both consider one reality of the ontology. Our approach distinguishes between objective and datascape realities. This separation of realities offers superior insight into the commitment of a statement or a concept involving data elements. A concept may have existence in the objective reality, but not necessarily in the datascape reality. For instance, considering the example we used for Quine's view, we may have the concept "student" as an object in the domain, but not "John". Thus, x exhibits ontological commitment, but not data commitment.

From an engineering perspective [29], a distinction is made between the definition of ontological commitment from a philosophical perspective and an engineering perspective. A formal definition of ontological commitment from an engineering point of view is given as "A formalized mapping between terms in a knowledge-base and identical or equivalent terms in an ontology" [29]. It is seen that this provided definition is similar to ours regarding ontological and data commitments. Indeed, in the data commitment part, we are mapping the given formula into the set of concepts in the ontology satisfying this formula based on the given domain. For the ontological commitment part, we are mapping the target concept to the appropriate sub-ontology based on the function `MapToConcept`, and the usage of the appropriate modularization technique. The mapping is made easy with the help of the mapping operator (i.e., τ), which we use in both the data commitment function and `MapToConcept` function defined earlier in Definition 9, and Definition 11, respectively.

In [30], the authors introduced a Description Logic (DL) family called Strict Ontological-Committed DL-Lite (STOC-DL-Lite), a variant of the widely known DL-Lite family. STOC-DL-Lite extends DL-Lite by incorporating ontological commitment concepts, providing a richer framework for ontology modelling. A knowledge base K in STOC-DL-Lite consists of a *CBox* (containing global ontological commitment concepts) alongside the standard *TBox* and *ABox* found in any DL-Lite knowledge base. An ontological commitment concept is denoted by $\circ C$, where C is a well-formed concept excluding $(\bot)$. This indicates K's explicit commitment to $\circ C$ within *CBox*. Semantically, every universal entity in K is an explicitly committed concept, with additional implicit commitments derived from the *TBox*, and *ABox*. The strict commitment context *SC-context* of K encompasses all committed concepts, determined inductively through a set of coherent rules applied to the *CBox*, *TBox*, and *ABox*. The approach to ontological commitment in [30] differs from that of this paper. While [30] focuses on identifying a comprehensive set of explicit and implicit commitments across the entire knowledge base, this research identifies the minimal sub-ontology committed to a specific concept or formula.

Concept satisfiability is a fundamental reasoning task in DL-based ontologies [31]. A concept C is satisfiable if there exists an interpretation in the knowledge base where it has instances (i.e., $C^I \neq \emptyset$). This process verifies whether a concept can have instances within the knowledge base. If no instances satisfy a concept, it indicates unsatisfiability. Several algorithms have been developed to check concept satisfiability [32, 33]. In traditional DL-based ontology, concept satisfiability is binary: a concept is either satisfiable (true) or not (false). However, in logics that support uncertainty management, such as fuzzy-DL or probabilistic-DL, a concept might be satisfiable to a certain degree [34, 35]. From our perspective, there are similarities between concept satisfiability and data commitment. Both tasks aim to verify if data supports the existence of a particular concept within a knowledge base. In data commitment, if a concept cannot be satisfied by any instances in the dataset, it implies a logical inconsistency within the ontology, similar to unsatisfiability in traditional DL reasoning.

Gruber's principle of minimal ontological commitment [36] advocates for ontologies to make minimal claims about the modelled world, allowing the parties committed to the ontology freedom to specialize and instantiate the ontology as needed. Our approach refines this principle by extracting a sub-ontology (i.e., a module) containing the essential explicit concepts (from data specializing predicates) and implicit concepts (from rooted graph relationships) along with their relationships. Through Definition 13 and Definition 14, we linked ontological and data commitments to domain adequacy and optimal domain adequacy, respectively. Certainly, there are more results to work out based on these definitions, which makes this paper a step towards a comprehensive theory on domain adequacy.

In [28], we find that ontological commitment and ontological presuppositions are two similar concepts in ontology. However, they refer to different ontological aspects. Ontological commitment brings up the concepts that a theory or statement implies or assumes to hold up. Ontological presuppositions give the underlying assumptions or beliefs about reality assumed within a certain discourse.

7 Conclusion and Future Work

We consider ontologies in a context where the domain consists of atomic concepts or those built from atomic elements using Cartesian constructions. This framework is well-suited to structured or semi-structured data. In big data environments, roughly 20% is structured, 10–20% is semi-structured, and 60–70% is unstructured. While unstructured data dominates, much of it can be transformed into structured form, though proportions vary by source, industry, and context. Our Cartesian approach, though limited in data handling, offers a precise formalization of concepts often treated informally. To our knowledge, this is the first formal treatment of ontological and data commitments in relation to domain adequacy, marking a step toward a comprehensive framework for reasoning about these notions.

We introduced a formal Cartesian framework addressing domain adequacy through its ontological and data commitment aspects, establishing a clear link between ontology modularization and ontological commitment. By leveraging these commitments, we developed a DIS that represents only the essential concepts and relationships of a domain.

This work focuses on DIS ontologies where concepts and relationships are certain, without addressing uncertainty modelling. However, extensive research exists on ontologies capturing uncertain concepts, relationships, and instances, as reviewed in [37]. As future work, we plan to extend our framework to address ontological and data commitments in uncertain ontologies and to further expand the proposed theory to derive additional results. This paper serves as a foundational step toward a comprehensive theory of domain adequacy.

References

[1] Cocchiarella, N.B.: Logic and ontology. Axiomathes **12**(1), 117–150 (2001) https://doi.org/10.1023/A:1012758003706

[2] SNOMED CT - National Center for Biomedical Ontology (01/31/2009). https://bioportal.bioontology.org/ontologies/SNOMEDCT

[3] SNOMED Test Findings - National Center for Biomedical Ontology (07/16/2010). https://bioportal.bioontology.org/ontologies/SNOMED_TEST/?p=summary

[4] LeClair, A., Marinache, A., El Ghalayini, H., MacCaull, W., Khedri, R.: A review on ontology modularization techniques - a multi-dimensional perspective. IEEE Transactions on Knowledge and Data Engineering **35**(5), 4376–4394 (2023) https://doi.org/10.1109/TKDE.2022.3152928

[5] Abbes, S.B., Scheuermann, A., Meilender, T., d'Aquin, M.: Characterizing modular ontologies. In: 7th International Conference on Formal Ontologies in Information Systems-FOIS 2012, pp. 13–25 (2012)

[6] Gries, D., Schneider, F.B.: A Logical Approach to Discrete Math. Springer, Berlin, Heidelberg (1993). https://doi.org/10.1007/978-1-4757-3837-7

[7] Smith, B.: Ontology, pp. 153–166. John Wiley & Sons, Ltd, Chichester, UK (2004). Chap. 11. https://doi.org/10.1002/9780470757017.ch11

[8] Marinache, A., Khedri, R., LeClair, A., MacCaull, W.: DIS: A data-centred knowledge representation formalism. In: 2021 Reconciling Data Analytics, Automation, Privacy, and Security: A Big Data Challenge (RDAAPS), pp. 1–8 (2021). https://doi.org/10.1109/RDAAPS48126.2021.9452007

[9] d'Aquin, M., Schlicht, A., Stuckenschmidt, H., Sabou, M.: In: Stuckenschmidt,

H., Parent, C., Spaccapietra, S. (eds.) Criteria and Evaluation for Ontology Modularization Techniques, pp. 67–89. Springer, Berlin, Heidelberg (2009). https://doi.org/10.1007/978-3-642-01907-4_4

[10] Wache, H., Vögele, T.J., Visser, U., Stuckenschmidt, H., Schuster, G., Neumann, H., Hübner, S.: Ontology-based integration of information - a survey of existing approaches. In: OIS@IJCAI (2001). https://api.semanticscholar.org/CorpusID:5707476

[11] Noy, N.F., Musen, M.A.: In: Stuckenschmidt, H., Parent, C., Spaccapietra, S. (eds.) Traversing Ontologies to Extract Views, pp. 245–260. Springer, Berlin, Heidelberg (2009). https://doi.org/10.1007/978-3-642-01907-4_11

[12] Romero, A.A., Kaminski, M., Grau, B.C., Horrocks, I.: Module extraction in expressive ontology languages via datalog reasoning. Journal of Artificial Intelligence Research **55**(1), 499–564 (2016) https://doi.org/10.1613/jair.4898

[13] Algergawy, A., Babalou, S., Kargar, M.J., Davarpanah, S.H.: Seecont: A new seeding-based clustering approach for ontology matching. In: Advances in Databases and Information Systems, pp. 245–258. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-23135-8_17

[14] LeClair, A., Khedri, R., Marinache, A.: Formalizing graphical modularization approaches for ontologies and the knowledge loss. In: International Joint Conference on Knowledge Discovery, Knowledge Engineering, and Knowledge Management, pp. 388–412 (2019). https://doi.org/10.1007/978-3-030-66196-0_18 . Springer

[15] Marinache, A.: Syntax and semantics of domain information system and its usage in conjecture verification. PhD thesis, School of Graduate Studies, McMaster University, Hamilton, Ontario, Canada (August 2025)

[16] Imieliński, T., Lipski, W.: The relational model of data and cylindric algebras. Journal of Computer and System Sciences **28**(1), 80–102 (1984) https://doi.org/10.1016/0022-0000(84)90077-1

[17] Monk, J.D.: Cylindric Algebras, pp. 219–229. Springer, New York, NY (1976). https://doi.org/10.1007/978-1-4684-9452-5_13

[18] Palmisano, I., Tamma, V., Payne, T., Doran, P.: Task oriented evaluation of module extraction techniques. In: The Semantic Web - ISWC 2009, pp. 130–145. Springer, Berlin, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04930-9_9

[19] Ghafourian, S., Rezaeian, A., Naghibzadeh, M.: Graph-based partitioning of ontology with semantic similarity. In: ICCKE 2013, pp. 80–85 (2013). https://doi.org/10.1109/ICCKE.2013.6682869

[20] Konev, B., Lutz, C., Ponomaryov, D., Wolter, F.: Decomposing description logic ontologies. In: Twelfth International Conference on the Principles of Knowledge Representation and Reasoning (2010)

[21] Kachroudi, M., Zghal, S., Ben Yahia, S.: Ontopart: At the cross-roads of ontology partitioning and scalable ontology alignment systems. International Journal of Metadata, Semantics and Ontologies **8**, 215–225 (2013) https://doi.org/10.1504/IJMSO.2013.057767

[22] Andrew LeClair, R.K.: An algebraic approach to ontology modularization and knowledge refinement. IEEE Transactions on Knowledge and Data Engineering, 1–26 (2025). Under review

[23] Straccia, U., Bobillo, F.: In: Pan, J.Z., Calvanese, D., Eiter, T., Horrocks, I., Kifer, M., Lin, F., Zhao, Y. (eds.) From Fuzzy to Annotated Semantic Web Languages, pp. 203–240. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-49493-7_6

[24] Tannen, V.: In: LIU, L., ÖZSU, M.T. (eds.) Conjunctive Query, pp. 448–449. Springer, Boston, MA (2009). https://doi.org/10.1007/978-0-387-39940-9_1002

[25] Quine, W.V.O.: On what there is. Review of Metaphysics **2**(5), 21–38 (1948)

[26] Smid, J.: The logic behind quine's criterion of ontological commitment. European Journal of Philosophy **28**(3), 789–804 (2020) https://doi.org/10.1111/ejop.12534

[27] Marcus, R.B.: Modalities and intensional languages. Synthese **13**(4), 303–322 (1961) https://doi.org/10.1007/BF00486629

[28] Janssen-Lauret, F.: Committing to an individual: ontological commitment, reference and epistemology. Synthese **193**(2), 583–604 (2016) https://doi.org/10.1007/s11229-015-0763-8

[29] Waterson, A., Preece, A.: Verifying ontological commitment in knowledge-based systems. Knowledge-Based Systems **12**(1), 45–54 (1999) https://doi.org/10.1016/S0950-7051(99)00007-6

[30] Espil, M.M., Ojea, M.G., Ojea, M.A.: Ontological commitments, dl-lite logics and reasoning tractability. In: Advances in Databases and Information Systems, pp. 230–244. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-23135-8_16

[31] Rudolph, S.: Foundations of Description Logics, pp. 76–136. Springer, Berlin, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23032-5_2

[32] Glimm, B.: Classical Algorithms for Reasoning and Explanation in Description Logics, pp. 1–64. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-31423-1_1

[33] Meissner, A.: The ALCN Description Logic Concept Satisfiability as a SAT Problem, pp. 253–263. Springer, Berlin, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23418-7_22

[34] Cerami, M., García-Cerdaña, Esteva, F.: On finitely-valued fuzzy description logics, vol. 55, pp. 1890–1916 (2014). https://doi.org/10.1016/j.ijar.2013.09.021

[35] Stoilos, G., Stamou, G., Pan, J.Z., Tzouvaras, V., Horrocks, I.: Reasoning with very expressive fuzzy description logics. Journal of Artificial Intelligence Research **30**, 273–320 (2007) https://doi.org/10.1613/jair.2279

[36] Gruber, T.R.: Toward principles for the design of ontologies used for knowledge sharing? International Journal of Human-Computer Studies **43**(5), 907–928 (1995) https://doi.org/10.1006/ijhc.1995.1081

[37] Alomair, D., Khedri, R., MacCaull, W.: A comprehensive review of uncertainty modelling in domain ontologies. ACM Computing Surveys, 1–35 (2024). Submitted for publication on Dec. 19, 2024

Chapter 5

Domain-adequacy of Ontologies with Statistically-defined Concepts: A DISEL plug-in

This chapter presents a practical implementation of the adequacy theory introduced earlier, focusing on datascape concepts whose definitions rely on statistical evaluation. It addresses the sixth objective of this thesis by determining the ontology domain adequacy for statistically defined concepts. The proposed system enables empirical verification of concept definitions and demonstrates how statistically grounded ontology engineering can be supported in practice. A case study is included to illustrate the process and benefits of the approach.

Data-adequacy of Ontologies with Statistically-defined Concepts: A DISEL plug-in

Deemah Alomair[a,b,*], Yihai Chen[c], Yijie Wang[c], Ridha Khedri[a]

[a]*McMaster University, 1208 Main Street West, Hamilton, L8S 4L8, Ontario, Canada*
[b]*Imam Abdulrahman bin Faisal University, King Faisal bin Abdulaziz Street, Dammam, 34212, Eastern Province, Kingdom of Saudi Arabia*
[c]*Shanghai University, 99 Shangda Road, Shanghai, 200444, Shanghai, China*

Abstract

An ontology usually captures a domain conceptualization. It involves the concepts within this domain and the relationships among them. This domain is the reality within which a system using the ontology exists and with which it interacts. A core part of this reality is the objective reality, which is formed by the concepts and the relationships that are deemed by the expert as important, noticeable, and relevant. Peripheral to the objective reality is a reality relative to the data to be analyzed. It is referred to as datascape reality. In this paper, we propose an approach to the verification of the relevance of datascape concepts of an ontology that are defined using terms involving data elements and statistical language. The proposed approach aims at improving the data adequacy of a given ontology regarding a specific dataset, leading to a smaller, more adequate ontology. We also propose an automation of data adequacy of ontologies through the generation of R programs run by the Domain Information System Extended Language (DISEL) plug-in to obtain the most adequate sub-ontology for the given data set. The automation system is presented as a DISEL Editor plug-in. Through a case study of a weather ontology, we illustrate the usage of the approach, and we demonstrate the automation of the verification process, which leads to a smaller, more data-adequate ontology.

*Corresponding Author: Deemah Alomair
Email addresses: `alomaird@mcmaster.ca` (Deemah Alomair), `yhchen@shu.edu.cn` (Yihai Chen), `wyj981113@shu.edu.cn` (Yijie Wang), `khedri@mcmaster.ca` (Ridha Khedri)

Preprint submitted to Journal of Web Semantics *September 17, 2025*

Keywords: Knowledge representation, reasoning, Ontologies, Data commitment, Statistical analysis, Data adequacy, Dataset-driven ontologies.

1. Introduction

1.1. Prelude

Ontologies serve as structured representations of a domain, capturing its key concepts and the relationships among them. The domain, in this context, refers to the portion of reality that the ontology is intended to model and reason about. This ontological reality can be understood as comprising two distinct yet interconnected types. The first is the objective reality: the conceptual structure defined by domain experts, grounded in theoretical or abstract knowledge and independent from any underlying data. The second is what we refer to as the datascape reality, a perspective derived from observed data. While objective reality reflects enduring domain understanding, datascape reality captures dynamic patterns and trends revealed through data. This distinction gives rise to two types of concepts within an ontology: Objective concepts, which exist independently of data and are typically defined in the conceptual core of the ontology (e.g., `City`, `Animal`). Datascape concepts, which are defined using data-driven constraints or statistical conditions (e.g., an `AgedCity` where more than 20% of the population is over 65, or an `ActiveAnimal` defined based on activity logs).

We refer to concepts defined using data-driven conditions as datascape concepts [1]. These concepts typically specialize the objective concepts, in some cases, by adding statistical constraints. For example, given the objective concept `City`, a datascape concept like `AgedCity` may be defined as a city where more than 20% of the population is above 65 years old, based on demographic data. Similarly, from the objective concept `Animal`, we can define the datascape concept `ActiveAnimal` as an animal whose recorded average daily movement exceeds a specified threshold based on sensor logs.

The validity of these datascape concepts depends on the underlying dataset. For instance, a dataset describing urban centers in a young population may contain no instances satisfying the `AgedCity` definition. In such cases, retaining the concept adds unnecessary complexity and misaligns the ontology with the actual data context. This highlights the need to distinguish between the two types of adequacies: *Domain adequacy*, where a concept aligns with the objective reality of the domain. The approach used to validate domain adequacy is called ontological commitment. *Data adequacy*, where the concept is supported by the dataset under consideration. The approach used to validate data adequacy is called data commitment.

While domain adequacy is ensured through ontological commitment, as thoroughly discussed in [1], this paper focuses on data adequacy, addressed through data commitment. Building on the foundations in [1], we tackle the verification of data commitment for datascape concepts, those defined statistically based on the dataset, and identify concepts lacking sufficient data support. Unsupported concepts are pruned, yielding a leaner, data-grounded ontology that remains semantically coherent and contextually aligned with the underlying data.

1.2. Motivation

A new challenge has emerged in ontology engineering as ontologies are increasingly applied in data-intensive environments. In such contexts, ontologies are often constructed or adopted in direct alignment with the content of the available datasets, which we refer to as dataset-driven ontologies. These ontologies commonly include datascape concepts: concepts whose definitions are grounded in the dataset under consideration and in some cases they are defined based on statistical properties or data-derived thresholds (e.g., a concept like `Hot-weather`, defined as having an average temperature above 30°C). While this data-centric approach enables ontologies to capture domain-specific patterns observed in data, it also brings a crucial problem related to the fact that not all statistically defined concepts are universally valid across datasets. For instance, a concept such as `Very-Rainy-Weather` may be well-supported in a tropical climate dataset, but becomes irrelevant when applied to data from arid or polar regions. Retaining such unsupported concepts introduces ontological noise. Concepts that lack instantiation in the data do not contribute to reasoning and may degrade the performance or interpretability of ontology-based systems. This issue is particularly acute in dynamic or frequently updated data-driven ontologies, where the underlying datasets evolve, requiring the ontology to remain tightly aligned with the current empirical reality.

The problem becomes even more pronounced in large-scale ontologies, where the cost of unnecessary or unsupported concepts compounds. For example, *SNOMEDCT* [2], a comprehensive medical ontology, contains approximately 367,827 concepts, posing considerable computational and maintenance challenges. Even a focused subset such as the *SNOMED_TEST* view, dictated to medical testing and comprising 4,332 concepts [3], remains substantial. This raises a fundamental question: Is it necessary, or even beneficial, to engage with the full set of concepts in such a large ontology, or within any of its specialized views? The challenge, therefore, is not to develop an entirely new class of ontologies, but rather to ensure the continued adequacy of existing dataset-driven ontologies as their underlying data contexts evolve. Standard techniques for ontology pruning and modularization tend to focus on structural criteria or logical preservation and do not account for the empirical validity of concepts based on the dataset currently in use. As such, there is a critical gap between structural adequacy (logical consistency and modularity) and data adequacy (database relevance and support). Our work addresses this gap by proposing a data-based mechanism for validating and pruning unsupported datascape concepts. This not only ensures that the ontology remains aligned with the empirical dataset but also significantly reduces its complexity and size. This result in more focused and lightweight ontologies that enhances the utility and maintainability of dataset-driven ontologies.

1.3. Contributions

This paper introduces a method to improve the overall ontological data adequacy of dataset-driven ontologies by validating the relevance of statistically defined concepts. Our key contributions are as follows:

− Formalization of data commitment for the datascape concept: We formalize the approach to evaluate the data commitment of the datascape concept, which is defined using statistical conditions over data attributes. The data commitment approach allows us to assess the data-level adequacy of the ontological concepts in a principled way.

− Automated validation framework: We design and implement an automated validation system for Domain Information System (DIS) framework (DIS is a bottom-up, data-centric formalism that constructs ontologies from datasets using a Cartesian construction of concepts. It separates domain knowledge (ontology) from data views (dataset), linking the two via a mapping operator. This structure allows concepts to be grounded in data while preserving ontological structure [4]). The proposed system parses statistical predicates from the ontology definition, generates corresponding R programs, executes them against the dataset, and evaluates whether each datascape concept satisfies its statistical condition.

– Ontology pruning based on data relevance: Concepts that fail the data-commitment test are automatically pruned from the ontology. This results in a smaller, more concise, and data-grounded sub-ontology that preserves only concepts with actual relevance to the dataset under analysis.
– Operational integration with Domain Information System Extended Language (DISEL) and R: We integrate this pruning mechanism into the DISEL ontology editor (a specification language for DIS framework), enabling users to perform statistical adequacy checks during ontology development. Our system bridges logical ontology modelling with statistical data analysis via seamless R engine integration.
– Case study with cross-datasets evaluation: We conduct a case study using a weather ontology applied to multiple datasets representing different geographic climates. This demonstrates the effectiveness of our method in tailoring a dataset-driven ontology to its specific data context and highlights how pruning decisions vary based on underlying data characteristics.

1.4. Structure of the paper

The remainder of the paper is structured as follows. Section 2 introduces the theoretical background. Section 3 presents our framework for validating data commitment of statistical datascape concepts. A case study is discussed in Section 4. Related work is reviewed in Section 5, and conclusions and future directions are provided in Section 6.

2. Preliminaries and Basic Concepts

In this section, we present the essential background for the paper to be self-contained. We introduce DIS, DISEL, and data statistical analysis.

2.1. Domain Information System

The DIS framework [4] is designed to map structured knowledge stored in a dataset with its associated ontology for representation and reasoning purposes. The framework comprises three main components: an ontology (denoted by $\mathcal{O}$), a data view (denoted by $\mathcal{A}$), and a mapping function (denoted by τ) that links data to concepts in the ontology. These elements give the DIS structure $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$.

The construction of the domain ontology structure is based on three elements $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$. The first component is the structure $\mathcal{C} = (C, \oplus, c_e)$, where C carries all the concepts presented in the lattice or rooted graphs, has c_e as the empty concept, and $\oplus$ is an operator for composing concepts of C. $\mathcal{C}$ is a communicative idempotent monoid. The second component is the Boolean lattice structure $\mathcal{L} = (L, \sqsubseteq_c)$. The set of concepts within the lattice is either directly obtained from the database schema (i.e., atomic concepts) by mapping each attribute in the dataset to an atomic concept within the lattice or obtained through the Cartesian composition of the atomic concepts using the $\oplus$ operator. The relationship between a concept and its sub-concepts is the `partOf` relation denoted by $\sqsubseteq_c$ and defined as $a \sqsubseteq_c b \iff a \oplus b = b$. The last component is the rooted graph structure $\mathcal{G}$. A rooted graph $\mathcal{G}_{t_i} = (C_i, R_i, t_i)$ involves concepts in C_i by relating them using relation R_i, and $t_i \in \mathcal{L}$ forms the root. Rooted graphs represent the concepts that are somehow related to a concept in the Boolean lattice and not directly generated by the composition operator. In this way, the Boolean lattice might be enriched with multiple rooted graphs having their roots at different lattice concepts.

The second component of a DIS system is the domain data view $\mathcal{A} = (A, +, \star, -, 0_A, 1_A, \{c_k\}_{k \in U})$, where U represents the finite set of attributes (sorts) of the considered dataset. The cylindrification operators c_k are indexed by the attributes (sorts) used in the data and that correspond to the elements of L, the carrier set of $\mathcal{L}$ [4, 5]. The central concept in this view is the notion of a `sort`, which corresponds to an attribute in the dataset. A pair consisting of a sort and one of its values is referred to as a `sorted value`. A `sorted datum` is then formed as a collection of sorted values, containing at most one sorted value per sort. A collection of such sorted datum forms the `sorted data`. The carrier set A of the cylindric algebra is interpreted as a set of sorted data. For more information about cylindric algebra, we refer the reader to [5, 6].

The last element of the DIS is the mapping operator τ, which maps the data in A to its corresponding concepts in the Boolean lattice L. Thus, the complete structure of DIS is $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$.

2.2. Domain Information System Extended Language

DISEL is a high-level ontology specification language developed based on the DIS framework [7]. It offers a user-friendly, eXtensible Markup Language (XML)-based syntax designed to abstract away the low-level metathetical details of DIS, making the ontology engineering process more accessible to domain experts and practitioners. DISEL inherits the semantics of DIS and is organized around four core constructs: `include`, `AtomDomain`, `concept`, and `graph`, described below.

1. `include` allows the integration of external ontologies, enabling ontology reuse and modular composition.

2. `AtomDomain` defines the atomic concepts of the ontology, typically aligned with dataset attributes. These form the base elements of the Boolean lattice (i.e., level above the smallest element c_e of the lattice). For each atomic concept, DISEL allows declaration of the concept name and an optional concept description.

3. `concept` specifies composite or derived concepts. For each concept, DISEL allows declaring its name, associated sub-lattice (`latticeOfConcepts`), a mathematical definition (in Isabelle-like syntax [8, 9], often wrapped in XML-CDATA [10]), and an optional description.

4. The `graph` construct introduces rooted graphs and consists of four components: a name specifying the graph concept, *RootedAt* indicating the root node, `edge` defining each arc and its endpoints using `from-to` sub-elements, and `relation`, a list of edges belonging to the same relation.

To assist users in building ontologies, the `DISEL Editor` was developed as open, multi-source software implemented in the Qt environment [11]. It supports the full DISEL specification process—from defining atomic concepts and generating the Boolean lattice to adding rooted graphs and importing external ontologies. The main interface is shown in Figure 1. The left panel facilitates the addition and editing of atomic and composite concepts, while the right panel visualizes the generated lattice and any rooted graphs linked to lattice elements. Clicking a lattice node reveals its associated rooted graph and attached concept definitions. Concepts and relations can be added manually or imported from other DISEL ontologies using the `include` construct. A complete example, the *weather ontology*, along with tool downloads, and the schema documentation is available in the official repository [12, 13].

2.3. Statistical Analysis

Descriptive and inferential statistics are widely used in data analysis [14, 15]. Descriptive statistics summarize data using aggregation functions such as the `mean`, `median`, and `mode` (measures of central tendency), as well as

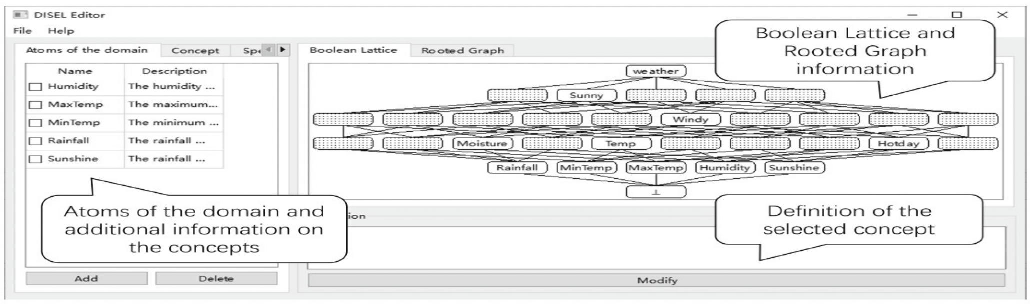


Figure 1: DISEL `Editor` Interface Overview.

range, **`variance`**, and **`standard deviation`** (measures of dispersion) [16]. Several statistical computations fall under descriptive statistics and are classified into subtypes, including types of variables, measures of frequency, central-tendency, dispersion/variation, and position calculation. In essence, descriptive statistics quantitatively capture the characteristics of observed data and can thus describe features of a datascape concept.

The second, inferential statistic, is used to draw conclusions from data through statistical tests. This process typically involves selecting a subset of individuals or data points from a larger population and then making inferences or predictions about the population based on the characteristics of the sample. In other words, it refers to the process of inferring or generalizing findings from a sample to a broader population [15]. **`Estimation`**, **`regression analysis`**, and **`hypothesis testing`** are some of the most common methodologies used in inferential statistics [17, 18].

In data analytics, statistical tools and programming languages play a pivotal role in extracting insights from complex datasets. Prominent examples include *Python* [19] and *R* [20], along with user-friendly programs like *Excel*, Statistical Analysis System (SAS) [21], and Statistical Package for the Social Sciences (SPSS) [22]. Among these, *R* stands out for its robustness and versatility, widely favored by statisticians, data scientists, and analysts. Its extensive ecosystem of statistical packages and libraries supports a broad range of analytical tasks. The language's extensibility fosters collaboration within the active *R* community, allowing users to develop and share custom packages. The use of specialized Integrated Development Environments (IDEs), such as *RStudio*, enhances the user experience for interactive analysis and visualization. Notably, *R* excels in data visualization, particularly with the *ggplot2* package, which provides a powerful framework for creating expressive and informative plots. Its open-source nature ensures ongoing development and support [23]. The adoption of R by major companies like Google and Ford further underscores its relevance in data analytics and statistical modeling [23]. For these reasons, we adopt R in our system.

3. Data Commitment of DIS-based Ontology

The primary objective of the proposed approach is to enable data commitment validation for concepts defined in the DIS framework. This approach facilitates the generation of an instantiated version of the ontology that optimally reflects the domain as represented in the dataset. By enhancing the capabilities of DISEL, the extended version goes beyond static ontology specification; it ensures that the defined concepts are empirically grounded in the data.

As illustrated in Figure 2, the proposed system architecture introduces two integrated reasoning sub-systems: statistical reasoning and formal reasoning. This paper focuses exclusively on the statistical reasoning sub-system, which is responsible for evaluating the empirical support of statistical-based datascape concepts within a given dataset (i.e., data commitment). The formal reasoning sub-system, which will support reasoning over validated concepts, is planned as part of future work. A detailed overview of the statistical reasoning sub-system is presented in the following subsection.

3.1. Statistical Reasoning Sub-System

3.1.1. Overview

In the DIS framework, concepts may originate from two distinct perspectives: Objective concepts, which are independent of any dataset, and datascape concepts, whose existence is conditioned by dataset-specific properties. While objective concepts are typically defined through logical composition within the Boolean lattice, datascape concepts are constructed by adding data restriction to a concept, in some cases, by using statistical predicates. These datascape concepts are typically attached to rooted graph concepts within the ontology.

Given a datascape concept defined by a statistical condition, it is not guaranteed that the dataset under consideration supports any instance that satisfies it. The goal of the statistical reasoning sub-system is to evaluate whether each such concept is data-committed, i.e., whether it has actual

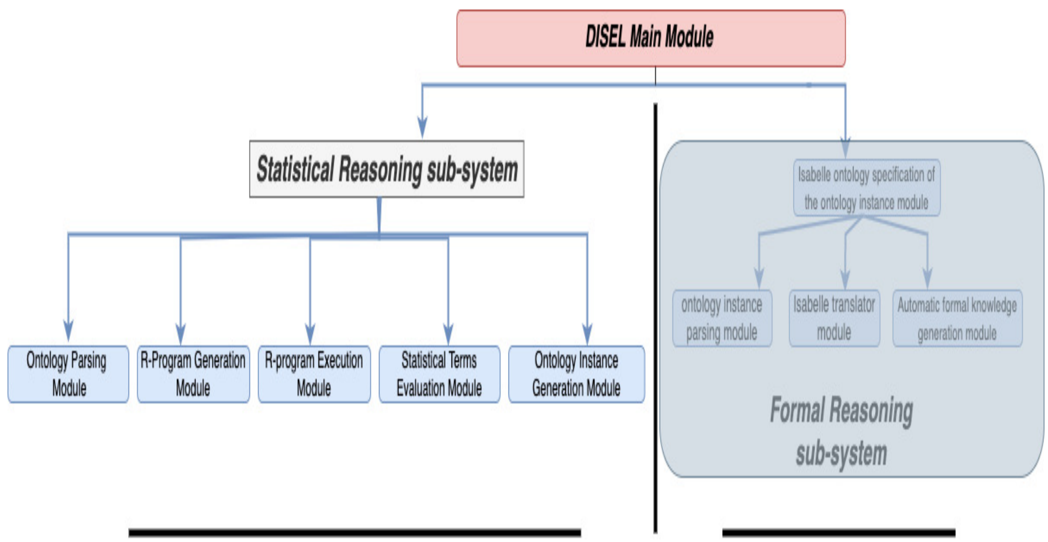


Figure 2: Proposed System Architecture

support in the dataset. If not, it is pruned from the ontology. This process reduces ontological noise and tailors the ontology to reflect domain grounded in data.

3.1.2. Formal Description

Let $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$ be a DIS ontology consisting of a Boolean lattice $\mathcal{L}$ of objective concepts and a set $\mathcal{G}$ of rooted graphs representing datascape concepts. Let $\mathcal{A}$ be a dataset with a set of instances. Each rooted graph concept $C_G \in \mathcal{G}$ is associated with a statistical predicate. Formally, we can define a datascape concept as follows:

Definition 1 (Datascape Concept, borrowed from [1])**.** *Let $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$ be a given DIS. Let A be the carrier set of the data view $\mathcal{A}$, and L be the Boolean lattice in $\mathcal{O}$. A datascape concept is a concept C_G that is defined as follows:*

$$C_G = \{a \mid \tau(a) \in \mathcal{L} \wedge \Phi_{C_G}(a)\},$$

where $a \in A$, and $\Phi(a)$ is a formula given in Disjunctive Normal Form (DNF) as $\Phi_{C_G}(a) = \ \vee(i \ \mid \ 1 \leq i \leq N \ : \ \Psi_i(a)\,)$, with N is a natural number, and $\Psi_i(a) = \ \wedge(j \ \mid \ 1 \leq j \leq M \ : \ \Omega_{(i,j)}(a)\,)$, where M is a natural number and $\Omega_{(i,j)}(a) = (f_{(i,j)}(a \cdot \texttt{sortname}_{(i,j)}), c_{(i,j)}) \in R_{(i,j)}$, where $f_{(i,j)} \in \mathcal{F}$, and

$\mathcal{F} = \{\oplus, e_c, \top_{\mathcal{L}}, +, \star, -, 0, 1, \tau, cyl\}$ is the set of function symbols, $c_{(i,j)}$ is a ground term in DIS language, and $R_{(i,j)}$ is a relator. We refer to Φ as a data specializing predicate, expressed within the formal language of $\mathcal{D}$. □

In the remainder of the paper, we assume that all data-specializing predicates are within the language of the DIS under consideration, augmented with statistical functions. A datascape concept C_G is said to be data-committed if and only if there exists at least one instance $a \in A$ such that the concept's data-specializing predicate $\Phi_{C_G}(a)$ evaluates to true:

$$(\exists a \mid a \in A : \Phi_{C_G}(a) = true),$$

That is, the condition $\Phi_{C_G}(a) = true$ is satisfied if and only if there exists an instance $a \in A$ that fulfills the constraints specified by Φ_{C_G}. Otherwise, C_G is deemed unsupported and excluded from the instantiated ontology. The result of this validation process is a pruned ontology $\mathcal{O}' = (\mathcal{C}', \mathcal{L}, \mathcal{G}')$ where:

$$\mathcal{G}' = \{C_G \mid C_G \in DCC\},$$

where DCC is the set of data committed concepts.

3.1.3. System Architecture

The statistical reasoning sub-system is implemented as a modular architecture that automates the data commitment validation process. It consists of five coordinated modules:

1. Ontology Parsing Module: Identifies statistical predicates in rooted graph concept definitions.

2. R-Program Generation Module: Translates statistical conditions into executable R code.

3. R-program Execution Module: Runs the generated R code and retrieves evaluation results.

4. Statistical Terms Evaluation Module: Evaluates the truth value of predicates based on the R output.

5. Ontology Instance Generation Module: Builds a new ontology containing only data-committed datascape concepts.

These modules form an integrated framework that transforms DISEL ontology into a data-grounded version consistent with the current dataset. The pruning process carried out by the above modules is formalized as follows:

1. For each rooted graph concept C_G, extract the predicate $\Phi(a)$.
2. Identify statistical function calls in $\Phi(a)$ and map them to corresponding R functions.
3. Fetch required data from $\mathcal{D}$ and evaluate each function using R engine.
4. Substitute functions calls in $\Phi(a)$ with their evaluated values, yielding a simplified predicate $\Phi'(a)$.
5. Evaluate the predicate $\Phi'(a)$ defining the datascape concept over all $a \in A$.
6. If $\exists a \in A$ such that $\Phi'(a)$ is true, retain C_G; otherwise, discard it.
7. Construct the new ontology $\mathcal{O}'$ with the updated set of rooted graph concepts.

3.1.4. Ontology Parsing Module

This module performs the initial extraction of statistical predicates from the ontology specification. It operates over the DISEL language's rooted graph construct, specifically focusing on the newly introduced `predicate` element, where statistical conditions are defined. Each `predicate` typically includes a combination of dataset attribute references and statistical operations such as `mean, variance, or distribution-fit`.

The module applies a grammar-aware parser to detect and extract the structure of each `predicate`, isolating the function names and their associated arguments. A predefined keywords dictionary is maintained internally, which maps recognized statistical terms to their corresponding R functions. This ensures compatibility with the statistical engine and prevents syntactic ambiguity. For example, if the term `average` appears, it is mapped to R's `mean()` function, and its argument is resolved against the dataset attributes.

Once a statistical predicate is successfully parsed, the module produces an intermediate representation that encapsulates the function name, argument bindings, and relational references to dataset attributes. This intermediate representation is then passed to the *R-Program Generation Module* for further processing. This parsing stage is essential to decouple the ontology specification from backend execution logic, allowing predicates to be authored in a domain-friendly language while ensuring backend compatibility.

3.1.5. R-program Generation Module

The *R-Program Generation Module* is responsible for transforming parsed statistical predicates into executable *R* code. After the *Ontology Parsing Module* extracts a statistical expression from a datascape concept's predicate, this module maps the identified statistical terms to their corresponding *R* functions, constructs valid *R* expressions, and assembles complete scripts that can be executed later. This module maintains a library of 28 statistical functions, logically organized into four functional categories: *Sampling, descriptive statistics, distribution fitting, and regression analysis*. These functions are defined to operate directly over database-resident data tables, using *RMariaDB* package [24] to establish connections and retrieve attribute values from the relational dataset associated with the ontology.

In the sampling category, the module supports multiple functions: Random sampling, column sampling, conditional sampling, stratified sampling, and cluster sampling. The sampling result is stored as a temporary table in a session of the database to allow the user to perform any further statistical operations on it. As the running session ended, the created sampling data frames were deleted. The descriptive statistics category includes operations like `median, mean, mode, range, maximum value, minimum value, top values, head values, variance, standard deviation, interquartile range, quantile, sum`, and data frame statistical summary functions. The distribution fitting category identifies the best-fitting probability distribution for a given attribute using established statistical tests (e.g., goodness of fit) [25]. The regression category supports both linear and logistic regression models.

To support flexible predicate evaluation, the module generates three distinct types of return values, depending on the structure of the predicate and the kind of statistical reasoning required:

1. Simple type: It returns a single value or a list. Such as `mean, median and distribution type`. They are typically used in predicates that involve direct comparisons (e.g., `mean`(a.`temp`) > 30)

2. Regression type: These return structured results in the form of *R* dataframe. This output is for predicates having a regression model, e.g.,

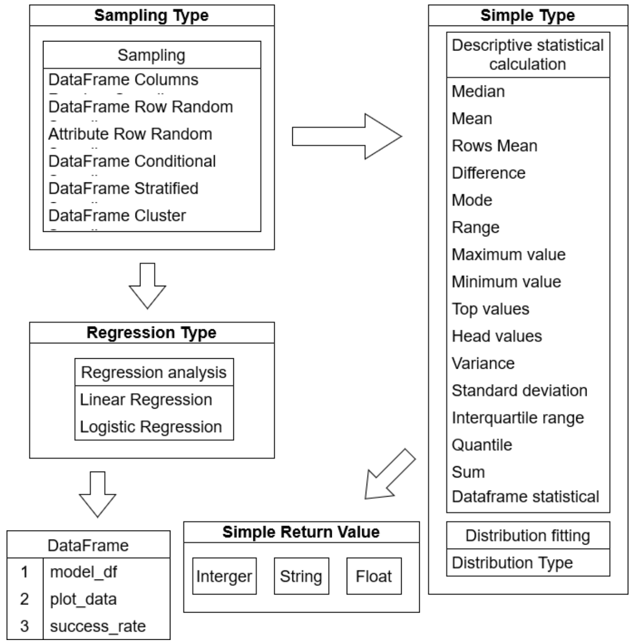


Figure 3: R-program Generation Module

determining whether a specific predictor significantly affects an outcome variable.

3. Sampling type: This represents a two-stage predicate structure. First, a sampling operation is performed (e.g., conditional or cluster Sampling), which generates a temporary subset of the data. Then, a simple or regression operation is applied to that subset. For example, `mean`(a.`temp`) might be calculated only on a stratified sample of high-humidity regions. These predicates require nesting, and the *R* script is generated to reflect this layered logic.

The internal flow of predicate processing is illustrated in Figure 3. As depicted in the diagram, the *sampling-type operations* often initiate the process, generating new data frames as input for further analysis. Depending on the predicate, the output from the sampling stage is then passed to either the *simple-type component* for basic statistical computation or to the *regression-type component* for modelling.

The generated *R* scripts are passed to the *R-Program Execution Module*, where they are run against the live datasets. Their output, whether scalar values, lists, or structured regression objects, is returned for evaluation.

In summary, the *R-Program Generation Module* not only translates high-level statistical predicates into executable form, but also encodes the appropriate computation logic tailored to each predicate's structure. By distinguishing between simple, regression, and sampling-based predicates, the system maintains both flexibility and semantic clarity in data validation, structured to execute independently and return results in a format compatible with the evaluation module. The output of this module is passed to the *R-Program Execution Module* for runtime evaluation.

3.1.6. R-Program Execution Module

The *R-Program Execution Module* enables the evaluation of statistical predicates defined within the ontology by executing corresponding *R* code on real-time data. Once the generated *R* script is received, the module runs it within an embedded *R* environment seamlessly integrated into the Qt-based system. The integration, achieved through the use of the *RInside* and *Rcpp* libraries [26], allows the module to access and retrieve live data from the underlying dataset. These packages provide a bridge between *C++* applications and the *R* interpreter, ensuring that statistical computations can be invoked programmatically and results can be retrieved seamlessly.

This module also uses *ANTLR4* to ensure that statistical expressions, especially those involving nested terms or function compositions, are parsed correctly and transformed into syntactically valid *R* expressions. The output produced by the execution (e.g., numeric values, vectors, data frames, or regression models) is returned to the *Statistical Terms Evaluation Module* for interpretation and decision-making.

This module is responsible not only for executing statistical logic but also for managing error handling, data type consistency, and resource cleanup. It ensures that each predicate is evaluated reliably and the results are correctly formatted for subsequent evaluation, forming a critical bridge between ontology semantics and statistical computation.

3.1.7. Statistical Terms Evaluation Module

The *Statistical Terms Evaluation Module* is responsible for assessing whether each datascape concept, defined by a statistical predicate, is empirically supported by the dataset, i.e., whether it exhibits data commitment. This decision determines whether the concept should be retained or pruned from the final instantiated ontology.

Once the *R-program Execution Module* completes its task and returns the evaluated result of a predicate based on the underlying dataset, this module receives both the computed value and the original predicate expression associated with a datascape concept. The key responsibility of this module is to reconstruct and evaluate the predicate by substituting the statistical term in the predicate with the actual computed result, then checking if the resulting condition is evaluated to true. For example, consider a datascape concept `HumidWeather` with the predicate $\Phi_{C_G}(a) = \{\texttt{average(a.humidity)} \geq 80\%\}$. Let assume that the `average(a.humidity)` evaluates to 76.2%, by the *R-program Execution Module*. Then this module substitutes this value into the predicate and evaluates the expression $76.2 > 80$. Since the condition is false for this evaluation, the concept `HumidWeather` is marked as not data-committed.

This module supports all three types of return values produced by the *R-Program Generation Module*. Once the evaluation is complete, each concept is annotated with a binary status: retain (data-committed) or prune (not-committed). The full set of data-committed concepts is then passed, along with the original ontology structure, to the *Data-Grounded Ontology Instance Generation Module*.

3.1.8. Data Grounded Ontology Instance Generation Module

The *Data Grounded Ontology Instance Generation Module* is the final component in the statistical reasoning sub-system. Its role is to construct an instantiated version of the ontology that includes only those datascape concepts which have been verified as data-committed. In doing so, this module ensures that the ontology remains both semantically consistent and empirically grounded concerning the dataset it was generated from.

This module receives two inputs: the original ontology specification (including all lattice concepts and rooted graph component) and the set of datascape concepts that were evaluated as committed by the *Statistical Terms Evaluation Module*. Using this information, it generates a refined ontology instance $\mathcal{O}' = (L, \mathcal{G}')$, where:

- L is the Boolean lattice of objective concepts, preserved in full.
- $\mathcal{G}' \subseteq \mathcal{G}$ is the set of rooted graph concepts for which $\exists a \in A$ such that $\Phi(a) = true$.

This module processes the ontology by iterating through all rooted graph constructs and copying only those whose predicates were satisfied. Each retained datascape concept $C_G \in \mathcal{G}'$ is attached to its corresponding lattice concept, maintaining the ontological hierarchy and rooted graph structure as defined in the original DISEL specification. The output of this module is a pruned, data-committed ontology that faithfully represents the domain as instantiated in the current dataset. This output can be:

– Exported for use in external tools or downstream reasoning engines;

– Visualized using DISEL `Editor`;

– Passed to the formal reasoning sub-system described in subsection 3.2.

By filtering the ontology based on actual data support, this module reduces complexity, removes irrelevant concepts, and enhances domain specificity, crucial in dynamic, data-intensive applications where datasets frequently change and the ontology must remain aligned. In summary, the *Data-Grounded Ontology Instance Generation Module* completes the final commitment step, transforming an abstract, data-grounded ontology into a concrete, contextually valid domain representation grounded in statistical evidence.

3.2. Formal Reasoning Sub-System

While the *Statistical Reasoning* sub-system validates datascape concepts by assessing their empirical support in a dataset, the *Formal Reasoning* sub-system is designed to enable symbolic knowledge derivation through logical inference. It builds upon the *Data-Grounded Ontology Instance* generated in the previous phase and provides a mechanism for deriving implicit knowledge, checking consistency, and answering domain-specific queries using formal logic.

This sub-system targets objective and committed datascape concepts alike, treating the instantiated ontology as a formal knowledge base over which logical reasoning can be performed. The reasoning engine used in this component is based on *Isabelle/HOL,* a widely used interactive theorem prover that supports higher-order logic [8]. While the implementation of this sub-system is left for future work, its three primary modules are described below.

1. Ontology Instance Parsing Module: This module parses the instantiated ontology, which is encoded in DISEL's XML-based syntax. The parser identifies formal elements such as concept names, their definitions, and graph-based relationships. Each element is extracted and

translated into a corresponding construct in *Isabelle/HOL's* theory language. Specifically, objective concepts are mapped to type-level declarations and definitions, while rooted graphs are translated into relational or set-theoretic assertions.

2. Translation to Isabelle Theory Module: Once the ontology content has been parsed, it is translated into a valid *Isabelle/HOL* theory. This theory serves as a formal specification of the ontology, enabling symbolic manipulation and logical derivation. The translation preserves concept hierarchies, set memberships, and conditions specified in the original ontology. Each concept is represented as a definition or predicate, and inter-concept relationships (e.g., edges in rooted graphs) are formalized as functions, relations, or rules depending on their semantics.

3. Automated Reasoning and Knowledge Derivation Module: With the Isabelle theory in place, the sub-system performs automated reasoning tasks. These include verifying the satisfiability of concepts, detecting inconsistencies, proving subsumption between concepts, or deriving new assertions based on logical entailment. Reasoning is carried out using Isabelle's built-in tactics and automation support, allowing for both human-interactive and fully automated workflows.

Importantly, this sub-system operates over the pruned, validated ontology, ensuring that only empirically grounded and semantically relevant concepts are used during reasoning. This enhances both the tractability and interpretability of the outcomes. The formal reasoning component completes the ontology cycle: beginning with the user-defined specification in DISEL, validating its empirical grounding via statistical analysis, and enabling inference through higher-order logic. By integrating data commitment validation with formal reasoning, the framework supports the development of ontologies that are both data-adequate and logically robust.

4. Case Study

4.1. Data Collection

This case study is adapted from a smart home weather ontology introduced in [27], implemented in Web Ontology Language 2 (OWL 2) using Protégé and the Pallet reasoner [28]. The ontology defines five core categories: *weather phenomenon, weather condition, weather state, weather* *report*, and *weather report source*, each with sub-concepts. For example, *weather phenomenon* includes *temperature, humidity, sun position, wind*, and *atmospheric pressure*. We focus on the `Weather-state` concept and its sub-concepts: `Hot-weather`, `Windy-Weather`, `Cold-weather`, `Dry-weather`, and `Rainy-Weather`. To demonstrate our approach's flexibility and effectiveness, we extend this subset with derived datascape concepts defined by statistical predicates, which evaluate their data commitment. The datasets used come from the *Visual Crossing weather dataset* [29], representing four geographically and climatically diverse locations in 2022:

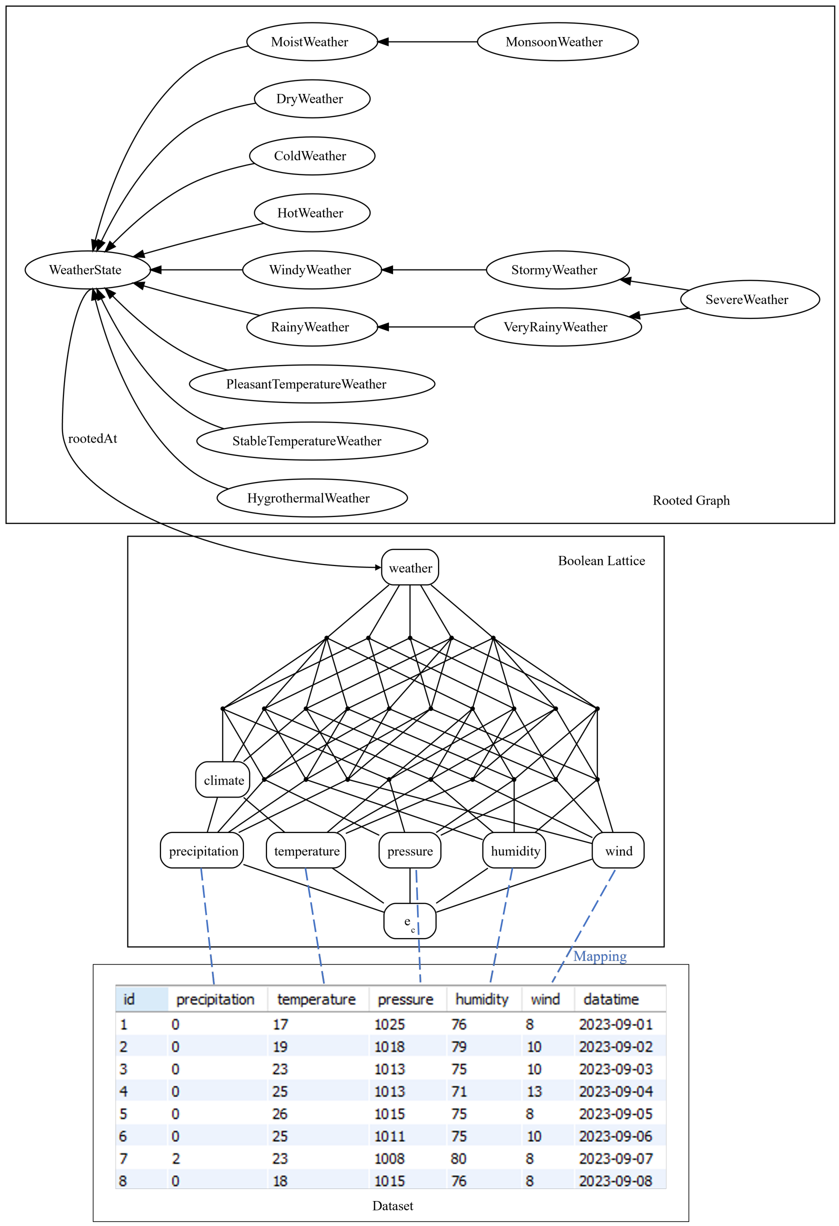
1. Shanghai: Represents a tropical urban climate.
2. Antarctica: An extreme polar climate.
3. Bilma (Sahara Desert): An extremely dry oasis climate.
4. Iquitos (Amazon): A rainforest climate with high humidity.

4.2. Data Analysis

From the dataset, we select five core attributes: `precipitation`, `wind`, `temp`, `pressure`, and `humidity`. These attributes are mapped to atomic concepts in the ontology via the mapping operator τ (as defined in subsection 2.1). Using the $\oplus$ operator, we then generate a Boolean lattice $\mathcal{L}$ that reflects all compositions of atomic concepts. The Boolean lattice with its associated rooted graph concepts is illustrated in Figure 4.

All rooted graph concepts linked to the `Weather-state` concept are defined via statistical predicates, embedded within the `predicate` property of each rooted graph concept. In our study, we selected five such rooted graph concepts. Their formal definitions in the DISEL-XML format are presented in Figure 5, while their statistical conditions are summarized below:

1. `Rainy-Weather`: Average precipitation $> 2mm$.
2. `Very-Rainy-Weather`: Average precipitation $> 5mm$.
3. `Windy-Weather`: Average wind speed $> 10m/s$.
4. `Stormy-Weather`: Average wind speed $> 20m/s$.
5. `Severe-Weather`: Standard deviation of wind or precipitation > 5, valid only if `Stormy-Weather` or `Very-Rainy-Weather` is true.



id	precipitation	temperature	pressure	humidity	wind	datatime
1	0	17	1025	76	8	2023-09-01
2	0	19	1018	79	10	2023-09-02
3	0	23	1013	75	10	2023-09-03
4	0	25	1013	71	13	2023-09-04
5	0	26	1015	75	8	2023-09-05
6	0	25	1011	75	10	2023-09-06
7	2	23	1008	80	8	2023-09-07
8	0	18	1015	76	8	2023-09-08

Figure 4: Constructed Lattice with Its Rooted Graphs.

```xml
<edge>
    <from>VeryRainyWeather</from>
    <to>RainyWeather</to>
    <predicate><![CDATA[ Average(weather.precipitation) > 5 ] ]]></predicate>
</edge>
<edge>
    <from>RainyWeather</from>
    <to>WeatherState</to>
    <predicate><![CDATA[ Average(weather.precipitation) > 2 ] ]]></predicate>
</edge>
<edge>
    <from>WindyWeather</from>
    <to>WeatherState</to>
    <predicate><![CDATA[ Average(weather.wind) > 10 ] ]]></predicate>
</edge>
<edge>
    <from>StormyWeather</from>
    <to>WindyWeather</to>
    <predicate><![CDATA[ Average(weather.wind) > 20 ] ]]></predicate>
</edge>
<edge>
    <from>SevereWeather</from>
    <to>StormyWeather</to>
    <predicate><![CDATA[ standardDeviation(weather.wind) > 5 ] ]]></predicate>
</edge>
<edge>
    <from>SevereWeather</from>
    <to>VeryRainyWeather</to>
    <predicate><![CDATA[ standardDeviation(weather.precipitation) > 5 ] ]]></predicate>
</edge>
```

Figure 5: XML Annotations for Rooted Graphs of Climate State Concept

Table 1: Predicates After Statistical Calculation and Evaluation Results In Shanghai

Predicates	**Edge(from)**	**Edge(to)**	**Predicate Evaluation**
Average(`weather.wind`) $>$ 10	`Windy-Weather`	`Weather-state`	True
Average(`weather.wind`) $>$ 20	`Stormy-Weather`	`Weather-state`	True
Average(`weather.precipitation`) $>$ 2	`Rainy-Weather`	`Weather-state`	False
Average(`weather.precipitation`) $>$ 5	`Very-Rainy-Weather`	`Weather-state`	False
standardDeviation(`weather.wind`) $>$ 5	`Severe-Weather`	`Stormy-Weather`	True
standardDeviation(`weather.precipitation`) $>$ 5	`Severe-Weather`	`Very-Rainy-Weather`	False

4.3. Datascape Concept Validation for Data Adequacy

To assess the adequacy of datascape concepts, we analyzed each `predicate` property in the XML specification using our parsing module. Predicates containing statistical terms (e.g., `Average and standardDeviation`) were evaluated using corresponding functions in the *R* engine. Table 1 shows the statistical calculations associated with each `predicate` and its corresponding evaluation outcome for the Shanghai dataset.

Concepts evaluated as *false* were considered not domain adequate and pruned from the ontology. Conversely, concepts whose `predicate` returns true are retained as rooted graph concepts attached to the ontology.

Table 2: The Comparison of Experiment Results in Data-commitment

Concept	Shanghai	Antarctica	Bilma	Iquitos
Moist Weather	✓			✓
Monsoon Weather	✓			✓
Dry Weather			✓	
Cold Weather	✓	✓		
Hot Weather	✓		✓	✓
Windy Weather	✓	✓	✓	✓
Stormy Weather	✓	✓	✓	
Rainy Weather			✓	✓
Very Rainy Weather				✓
Severe Weather	✓	✓	✓	✓
Pleasant Temperature Weather	✓			
Stable Temperature Weather		✓	✓	✓
Humidity Temperature Correlation Weather		✓		✓

4.4. Data Commitment Across Multiple Datasets

To verify data commitment across diverse domains, we repeated the validation of the datascape concepts for each dataset. Table 2 summarizes the data commitment for each datascape concept across datasets. Figure 6 illustrates rooted graph results per domain. The diversity of datasets helped reveal concept dependencies specific to climate and region. For example, the concept `Severe-Weather` in *Bilma* was supported by wind-related data, whereas in *Iquitos,* precipitation variability emerged as the primary contributing factor. Similarly, the concept `Rainy-Weather` is supported by empirical data from the *Bilma* and *Iquitos* regions, but not from *Shanghai* or *Antarctica*, thereby reinforcing the importance of empirical validation.

5. Related Work and Discussion

5.1. Ontological Commitment and Domain Adequacy

Ontological commitment is a well-established topic in both philosophical and computing literature [30, 31, 32]. Informally, it refers to the requirement that concepts included in the ontology must correspond to entities that genuinely exist in the intended domain. Several philosophical interpretations exist, including those by Quine [33], Marcus [34], Russell [35], and others [30]. These perspectives typically operate under a unified notion of reality and do not distinguish between data-related and objective existence.

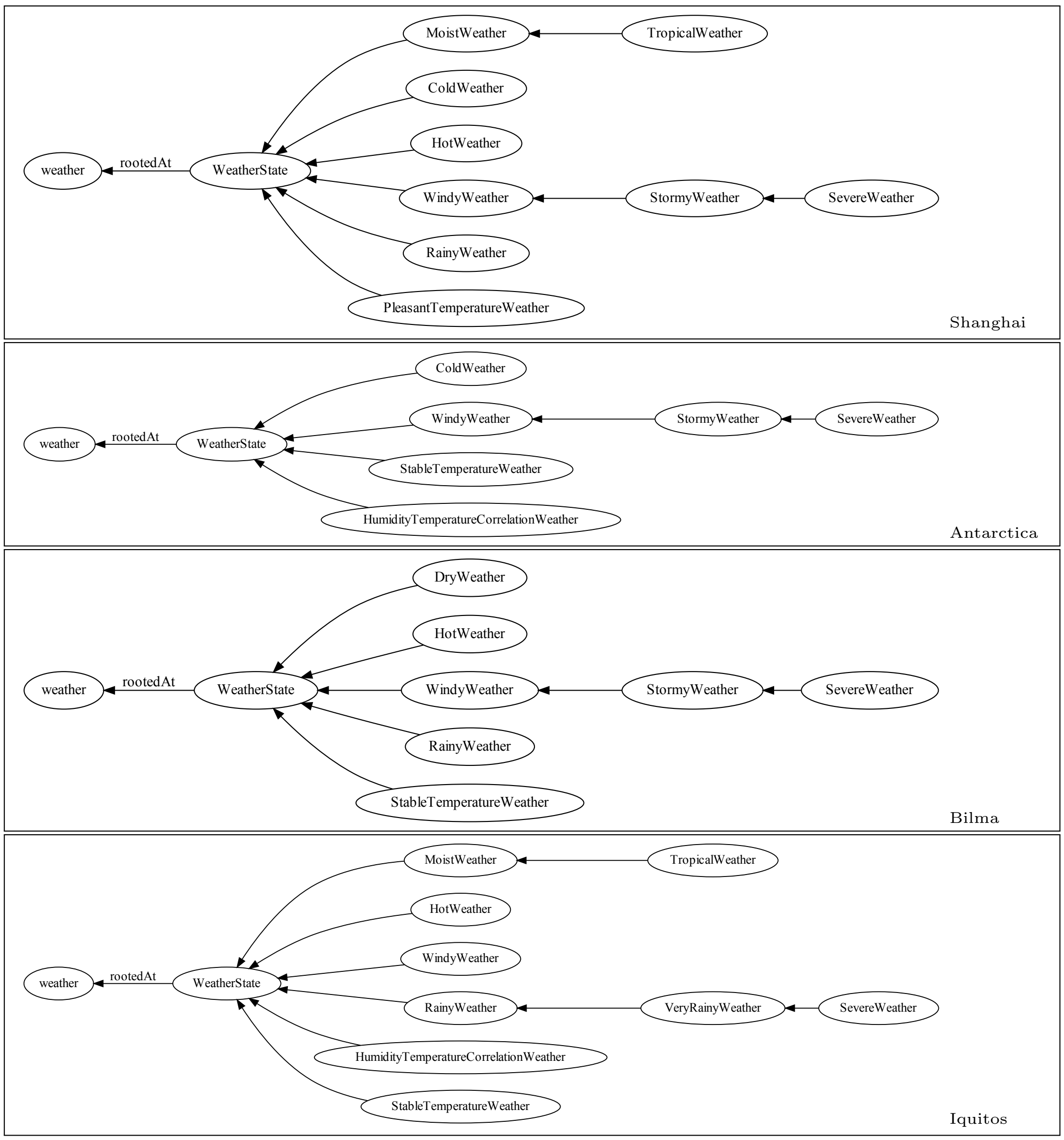


Figure 6: Validated Rooted Graph Concepts of Each Dataset

23

Traditional approaches often conflate these two forms of reality, making it unclear whether a concept's relevance is based on structural semantics or data-driven observation. In contrast, the approach proposed in [1] introduces a dual-layered understanding of reality, distinguishing between objective and datascape realities. Based on this distinction, they identify two complementary forms of commitments: (1) ontological commitment, which ensures conceptual alignment with the abstract domain model, and (2) data commitment, which validates empirical instantiation in a specific dataset. A concept may be ontologically valid yet lack data commitment, and vice versa. Adhering to both ontological and datascape commitments ensures optimal domain adequacy of the ontology.

It is worth noting that while data commitment is an important dimension of domain adequacy, it does not fully capture the semantic or inferential relevance of a concept. Achieving full domain adequacy requires ensuring both forms of commitments: ontological and data commitment. For a more detailed discussion on optimal domain adequacy and the interplay between these forms of commitment, we refer the reader to [1].

Our work contributes to the broader objective of achieving optimal domain adequacy by validating the data commitment of statistically defined concepts. Although data commitment alone is insufficient for full adequacy, it is critical for filtering unsupported concepts and maintaining alignment with evolving data. Future extensions of our work will incorporate deeper ontological validation and formal reasoning to support both dimensions of adequacy in a unified framework.

5.2. Statistical Reasoning in Ontology Engineering

Statistical methods have been increasingly incorporated into ontology-related work, particularly in areas like statistical knowledge representation and intelligent data analytics. Examples include the development of statistical ontologies from domains such as biology [36], smart manufacturing [37], and general statistical processes [38]. These ontologies primarily aim to represent statistical concepts, procedures, and their interrelationships in a formal structure.

Our work departs from this direction. Rather than modelling statistical knowledge, we employ statistical reasoning as a procedural tool to evaluate and validate datascape concepts. This enables us to determine whether such concepts are data-committed, and if not, to remove them from the ontology.

Thus, our goal is not to create a statistical ontology but to use statistical methods to ensure that the ontology remains adequate.

The proposed approach aligns with the emerging field of logical data analytics, where statistical reasoning is used to test the validity of data-driven hypotheses [39]. In our framework, dataset $\mathcal{A}$ is treated as a population of domain instances, and datascape concepts are defined through statistical predicates over these instances. These definitions are interpreted as hypotheses that must be tested to confirm whether they represent valid characteristics of the population. For example, the concept of `Humid Weather` may be defined as a set of instances where the average humidity exceeds a threshold. The existence of such a concept in the ontology depends on whether this statistical pattern is actually present in the data.

5.3. Ontology Size Reduction and Modularization

Reducing ontology size while preserving reasoning capabilities is a long-standing goal in knowledge representation. Various modularization techniques have been proposed to extract minimal sub-ontologies relevant to specific reasoning tasks [40]. One such approach, known as view traversal [41], builds a view-dependent module by incorporating a target concept, its sub-concepts, and their relationships.

However, most modularization techniques focus on preserving logical structure, without explicitly considering whether the retained concepts are relevant in the context of a given dataset. In contrast, our approach introduces data-ware modularization, where the inclusion of a datascape concept is contingent upon statistical validation against the dataset. This leads to an ontology that is not only logically coherent but also data-committed, making it particularly suitable for data-driven or dynamic applications.

6. Conclusion and Future Work

In this paper, we presented an approach for verifying the relevance of ontology concepts that are defined using terms derived from data elements, concepts that frequently incorporate constructs from statistical languages. Specifically, we introduced an automated verification process using a statistical reasoning system employed to DISEL specification. This verification aims to reduce the ontology to include only concepts that are either objectively needed (i.e., objective concepts) or whose existence is empirically supported by the given dataset (i.e., committed datascape concepts). The resulting ontology satisfies the data adequacy, enabling more efficient use in downstream tasks such as reasoning or modularization. A case study was used to demonstrate the practical utility of our approach and to validate the automation of the verification process.

As part of future work, we plan to extend this framework to interoperate with logical reasoning, either through formalization in *Isabelle/HOL* for DIS ontologies or by interfacing with a decidable Description Logic (DL) reasoner when the DIS language is restricted to an appropriate sub-language of a first-order logic. Additionally, we aim to conduct an empirical evaluation on large-scale ontologies to assess the impact of data commitment on ontology size reduction. For instance, ontologies like *SNOMEDCT* [3], which contains over 367,827 concepts, present a major bottleneck for reasoning tasks due to their size. This empirical study will bring a better understanding of the gain achieved in ensuring that the concepts of an ontology have data relevance.

References

[1] D. Alomair, R. Khedri, Towards a Cartesian theory of ontology domain adequacy, Synthese. Submitted for publication (Under Review) (2025).

[2] SNOMED CT - National Center for Biomedical Ontology (01/31/2009). URL https://bioportal.bioontology.org/ontologies/SNOMEDCT

[3] SNOMED Test Findings - National Center for Biomedical Ontology (07/16/2010). URL https://bioportal.bioontology.org/ontologies/SNOMED_TEST/?p=summary

[4] A. Marinache, R. Khedri, A. LeClair, W. MacCaull, DIS: A data-centred knowledge representation formalism, in: 2021 Reconciling Data Analytics, Automation, Privacy, and Security: A Big Data Challenge (RDAAPS), 2021, pp. 1–8. doi:10.1109/RDAAPS48126.2021.9452007.

[5] T. Imieliński, W. Lipski, The relational model of data and cylindric algebras, Journal of Computer and System Sciences 28 (1) (1984) 80–102. doi:https://doi.org/10.1016/0022-0000(84)90077-1.

[6] J. D. Monk, Cylindric Algebras, Springer New York, New York, NY, 1976, pp. 219–229. doi:https://doi.org/10.1007/978-1-4684-9452-5_13.

[7] Y. Wang, Y. Chen, D. Alomair, R. Khedri, DISEL: A language for specifying dis-based ontologies, in: The 15th International Conference on Knowledge Science, Engineering and Management (KSEM 2022), Lecture Notes in Artificial Intelligence, Springer, Singapore, 2022, pp. 1–16. doi:10.1007/978-3-031-10986-7_13.

[8] Isabelle, http://isabelle.in.tum.de/index.html (2 2024).

[9] L. Bulwahn, A. Krauss, F. Haftmann, L. Erkök, J. Matthews, Imperative functional programming with isabelle/hol, in: Theorem Proving in Higher Order Logics, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 134–149. doi:https://doi.org/10.1007/978-3-540-71067-7_14.

[10] XML - CDATA sections, https://www.tutorialspoint.com/xml/xml_cdata_sections.htm (2024).

[11] Qt:cross-platform software design and development tools, https://www.qt.io/ (2024).

[12] Y. Wang, The github repository of DISEL editor, https://github.com/YbJerry/DISEL_Editor (2022).

[13] Y. Wang, [dataset] a weather example specified using DISEL, https://github.com/YbJerry/DISEL_file/blob/master/dis.xml (2022).

[14] P. Mishra, C. M. Pandey, U. Singh, A. Keshri, M. Sabaretnam, Selection of appropriate statistical methods for data analysis, Annals of Cardiac Anaesthesia 22 (3) (2019) 297–301. doi:10.4103/aca.ACA_248_18.

[15] B. Cohen, Essentials of statistics for the social and behavioral sciences, 2003.

[16] P. Kaur, J. Stoltzfus, V. Yellapu, Descriptive statistics, International Journal of Academic Medicine 4 (1) (2018) 60–63. doi:10.4103/IJAM.IJAM_7_18.

[17] S. Barnes, C. Lewin, An introduction to inferential statistics: Testing for differences and relationships, Research methods in the social sciences (2005) 226–235.

[18] K. Stapor, Descriptive and Inferential Statistics, in: Introduction to Probabilistic and Statistical Methods with Examples in R, Springer International Publishing, Cham, 2020, pp. 63–131. doi:10.1007/978-3-030-45799-0_2.

[19] T. Haslwanter, An introduction to statistics with Python, With Applications in the Life Sciences; Springer International Publishing: Cham, Switzerland (2016). doi:https://doi.org/10.1007/978-3-030-97371-1.

[20] M. Gardener, Beginning R. The statistical programming language, 2012.

[21] R. N. Rodriguez, SAS, WIREs Computational Statistics 3 (1) (2011) 1–11. doi:https://doi.org/10.1002/wics.131.

[22] V. Beddo, F. Kreuter, A handbook of statistical analyses using SPSS, Journal of Statistical Software 11 (07 2004). doi:10.18637/jss.v011.b02.

[23] G. Kamingu, Python vs. R for data scientists: An introduction, Tech. rep., Optimization and Machine Learning Laborato (June 2023). doi:10.13140/RG.2.2.17253.37602.

[24] RMariaDB: Database interface and MariaDB driver, https://cran.r-project.org/web/packages/RMariaDB/index.html (10 2023).

[25] S. P. Millard, distChoose, https://www.rdocumentation.org/packages/EnvStats/versions/2.7.0/topics/distChoose (2022).

[26] RInside: Easy embedding of R inside C++ (and C), https://github.com/eddelbuettel/rinside (2024).

[27] P. Staroch, A weather ontology for predictive control in smart homes, Thesis, Technische Universität Wien (2013). doi:10.34726/hss.2013.21706.

[28] OWL2, https://www.w3.org/TR/owl2-overview (2012).

[29] Weather data & API global forecast & history data, https://www.visualcrossing.com (2024).

[30] F. Janssen-Lauret, Committing to an individual: ontological commitment, reference and epistemology, Synthese 193 (2) (2016) 583–604. doi:10.1007/s11229-015-0763-8.

[31] A. Waterson, A. Preece, Verifying ontological commitment in knowledge-based systems, Knowledge-Based Systems 12 (1) (1999) 45–54. doi:10.1016/S0950-7051(99)00007-6.

[32] M. M. Espil, M. G. Ojea, M. A. Ojea, Ontological commitments, DL-Lite logics and reasoning tractability, in: Advances in Databases and Information Systems, Springer International Publishing, Cham, 2015, pp. 230–244. doi:https://doi.org/10.1007/978-3-319-23135-8_16.

[33] J. Smid, The logic behind quine's criterion of ontological commitment, European Journal of Philosophy 28 (3) (2020) 789–804. doi:https://doi.org/10.1111/ejop.12534.

[34] R. B. Marcus, Modalities and intensional languages, Synthese 13 (4) (1961) 303–322. doi:https://doi.org/10.1007/BF00486629.

[35] B. Russell, V.—knowledge by acquaintance and knowledge by description, Proceedings of the Aristotelian Society 11 (1) (2015) 108–128. doi:10.1093/aristotelian/11.1.108.

[36] J. Zheng, M. R. Harris, A. M. Masci, Y. Lin, A. Hero, B. Smith, Y. He, The Ontology of Biological and Clinical Statistics (OBCS) for standardized and reproducible statistical analysis, Journal of Biomedical Semantics 7 (1) (2016) 53. doi:10.1186/s13326-016-0100-2.

[37] Z. Zheng, B. Zhou, D. Zhou, A. Q. Khan, A. Soylu, E. Kharlamov, Towards a statistic ontology for data analysis in smart manufacturing, in: International Workshop on the Semantic Web, 2022.

[38] C. Franck, V. Florian, R. Flavio, V. Zoltán, M. Juan, F. Thomas, COOS – Core Ontology for Official Statistics, https://linked-statistics.github.io/COOS/coos.html (2023).

[39] J.-W. Romeijn, Philosophy of Statistics, 2014, pp. 1–86.

[40] A. LeClair, A. Marinache, H. El Ghalayini, W. MacCaull, R. Khedri, A review on ontology modularization techniques - a multi-dimensional

perspective, IEEE Transactions on Knowledge and Data Engineering 35 (5) (2023) 4376–4394. doi:10.1109/TKDE.2022.3152928.

[41] A. LeClair, R. Khedri, A. Marinache, Formalizing graphical modularization approaches for ontologies and the knowledge loss, in: International Joint Conference on Knowledge Discovery, Knowledge Engineering, and Knowledge Management, Springer, 2019, pp. 388– 412. doi:10.1007/978-3-030-66196-0_18.

Chapter 6

Conclusion and Future Works

This thesis investigated the issue of ontological uncertainty modelling either related to information or to the relevance of concepts and their relations. Central to this work is the development of formal frameworks that enable ontology-based systems to operate reliably in the presence of uncertainty. The research addressed multiple dimensions of uncertainty, distinguishing between classical information imperfections such as incompleteness and the more nuanced challenge of relevance uncertainty, which concerns the contextual appropriateness of concepts within a specific domain and dataset. To tackle these challenges, the thesis introduced a set of formal models, reasoning procedures, and supporting tools that ensure both logical soundness and data alignment, thereby enhancing the expressive power and practical utility of semantic systems.

6.1 Interpretation and Value of the Contributions

The contributions of this thesis offer a comprehensive and layered advancement in the design and reasoning capabilities of ontology-based systems operating under uncertainty. Taken together, they address a longstanding gap in semantic technologies: the lack of formal, data-aligned, and uncertainty-aware reasoning mechanisms that reflect the needs of real-world domains. The value of the work can be interpreted across several dimensions:

- Clarifying the ontology-uncertainty relationship. The developed taxonomy of uncertainty and its application through a structured survey provided the field with a much-needed conceptual framework. It clarifies how different uncertainty types (e.g., incompleteness, vagueness, inconsistencies) interact with ontology modelling, and it identifies the appropriate formalism for each. This framework strengthens the theoretical foundation of uncertainty-aware ontology engineering and offers a roadmap for selecting or designing reasoning tools fit for specific applications.

- Advancing uncertainty reasoning. Through the formal integration of possibility theory, the thesis provides a fine-grained reasoning mechanism capable of operating over incomplete and uncertain knowledge. The use of necessity measures allows logical assertions to be qualified with degrees of certainty, enabling nuanced inference that more closely mirrors real-world decision-making processes. This is particularly valuable in domains where information is usually incomplete or partially available.

- Operationalizing reasoning under incompleteness. The DISEL tool transforms the formal contributions of the thesis into a usable, automated reasoning environment that supports necessity-weighted logic. DISEL demonstrates that reasoning under uncertainty can be made tractable and practical. It facilitates key ontology services, such as subsumption and concept satisfiability, in environments where classical assumptions of completeness and determinism do not hold.

- Enhancing domain adequacy through commitment. The formalization of domain adequacy, defined through the dual lens of ontological (structural alignment with conceptual needs) and data commitment (empirical alignment with observed data), offers a principled way to assess and improve the fitness of ontologies for their intended use. This concept goes beyond traditional ontology evaluation metrics by embedding both logical and data-oriented criteria, making it particularly relevant in applied settings where ontologies must be tailored to actual domain behavior.

- Bridging data statistical semantics to logical reasoning. A central achievement of this thesis lies in reconciling statistical data analysis with logical ontological reasoning. By incorporating statistical validation mechanisms and defining data-committed concepts, the thesis makes it possible to anchor ontological definitions in empirical reality. This ensures that modelled concepts are not only theoretically sound but also grounded in observable patterns, a critical requirement for domain-specific intelligent systems.

6.2 Future Work

While this thesis lays a comprehensive foundation for modelling and reasoning with uncertainty in ontology-based systems, it also opens several avenues for future research and development. These directions build upon the theoretical results, practical tools, and insights into domain adequacy, aiming to further extend the applicability, expressiveness, and performance of the proposed frameworks.

1. Formal verification of reasoning procedures and domain adequacy. The formal reasoning mechanisms developed in this thesis, particularly those grounded in necessity-based possibilistic logic, would benefit from machine-checked verification using tools such as *Isabelle/HOL*. Formalizing the DIS framework and domain adequacy definitions in a proof assistant would strengthen the soundness and trustworthiness of the reasoning procedures and enable certifiable inference in safety-critical applications.

2. Integration of more expressive statistical and machine learning models. The current statistical plug-in supports hypothesis-driven reasoning based on predefined predicates. Future extensions could incorporate:

    - Richer statistical models, including multivariate analysis, and time-series inference.

    - Interpretable machine learning models, such as decision trees or rule-based learners, to discover or refine data-specializing predicates.

3. Support for evolving and streaming data. To extend applicability in dynamic environments, the framework could be adopted to handle evolving datasets and streaming data. This would require:

    - Incremental validation of datascape concepts.

    - Temporal extensions to the DIS framework to represent time-sensitive uncertainty.

    - Monitoring mechanisms to track and explain changes in necessity values over time.

4. Expanded evaluation across multiple domains. While the case studies in this thesis demonstrated feasibility, more diverse empirical evaluations are needed to generalize the results. Future work should:

    - Apply the framework to domains such as smart healthcare, intelligent education systems, or cybersecurity.

- Evaluate the performance of reasoning tasks under varying levels of incompleteness.
- Access usability, ontology quality improvements, and user trust in real deployment settings.

5. Enhancing interoperability and user accessibility. For broader adoption, future versions of the DISEL tool and the statistical plug-in can be enhanced by:
   - Developing a graphical interface for defining concepts, necessity weights, and statistical predicates.
   - Providing interoperability with OWL, RDF, and *Protégé*, enabling users to import/export standard ontologies whole, leveraging the extended reasoning capabilities.
   - Building support for explainable reasoning visualizations, allowing users to understand how necessity values propagate and impact reasoning outcomes.
6. Broadening the scope of uncertainty handling within DIS framework. This thesis focused primarily on incompleteness and relevance uncertainty. Future work could expand the framework to address additional uncertainty types, including vagueness and inconsistency.

Appendix A

DISEL: A Language for Specifying DIS-based Ontologies

This appendix includes the full version of the published manuscript titled: *DISEL: A Language for Specifying DIS-based Ontologies*, by Yijie Wang, Yihai Chen, Deemah Alomair, and Ridha Khedri. This manuscript present the main syntactical constructs of DISEL language, which is designed as a language for specifying DIS-based ontologies.

The content of this chapter is co-authored with Dr. Yihai Chen, Yijie Wang, and Dr. Ridha Khedri, and has been published in the *Lecture Notes in Computer Science* (LNAI, volume 13369), as part of the proceedings of the *15th International Conference on Knowledge Science, Engineering and Management (KSEM 2022)*. The work is cited as follows, with copyright held by the authors and *Springer Nature*:

Wang, Y., Chen, Y., Alomair, D., Khedri, R. (2022). DISEL: A Language for Specifying DIS-Based Ontologies. In: Memmi, G., Yang, B., Kong, L., Zhang, T., Qiu, M. (eds) Knowledge Science, Engineering and Management. KSEM 2022. Lecture Notes in Computer Science(), vol 13369. Springer, Cham. https://doi.org/10.1007/978-3-031-10986-7_13

DISEL: A Language for Specifying DIS-Based Ontologies

Yijie Wang[1], Yihai Chen[1], Deemah Alomair[2,3], and Ridha Khedri[2]( )

[1] Shanghai University, Shanghai, China
{wyj981113,yhchen}@shu.edu.cn
[2] McMaster University, Hamilton, ON, Canada
{alomaird,khedri}@mcmaster.ca
[3] Imam Abdulrahman Bin Faisal University, Dammam, Kingdom of Saudi Arabia

Abstract. We present the main syntactical constructs of *DISEL* language, which is designed as a language for specifying DIS-based ontologies. The adoption of this language would enable the creation of shareable ontologies for the development of ontology-based systems. We give the main constructs of the language and we illustrate the specification of the main components of a *DISEL* ontology using a simplified example of a weather ontology. DIS formalism, on which the proposed language is based, enables the modelling of an ontology in a bottom-up approach. The syntax of *DISEL* language is based on XML, which eases the translation of its ontologies to other ontology languages. We also introduce *DISEL Editor* tool, which has several capabilities such as editing and visualising ontologies. It can guide the specifier in providing the essential elements of the ontology, then it automatically produces the full *DISEL* ontology specification.

Keywords: Ontology · Ontology language · Ontology specification · Knowledge engineering · DIS formalism · Ontology visualization

1 Introduction

In the last two decades, we have seen an expansion in the volume and complexity of organized data sets ranging from databases, log files, and transformation of unorganized data to organized data [16,43]. To extract significant knowledge from this data, the use of ontologies allows to connect the dots between information directly inferred from the data to concepts in the domain of the data. Ontology refers to a branch of metaphysics about the study of concepts in a world (i.e., a reality). This world is commonly referred to in information science as domain. The literature abounds with works related to the creation of ontologies, and to the characterisation of the best methods to create representations of reality [20]. Ontologies have been used in several areas of knowledge. For instance, in biological sciences and with the development of the GeneOntology and the creation of the community of the Open Biological and Biomedical Ontology (OBO) Foundry a wide interest rose to use of ontologies in this area [23].

G. Memmi et al. (Eds.): KSEM 2022, LNAI 13369, pp. 155–171, 2022.
https://doi.org/10.1007/978-3-031-10986-7_13

Another area where ontologies are widely used is that of information sciences. A wide literature has been published related to ontologies and their use in basic reasoning problems (e.g., [46]), learning (ontologies and queries) (e.g., [41]), privacy management (e.g., [40]), interactive query formulation and answering (e.g., [29]), or data cleaning [28]. These are only a few example of the areas of application of ontologies in information sciences.

Ontologies are used to represent and reason about a domain. There are several formalisms with their specific languages that are used to specify domain knowledge, such as graphs (e.g., [22]), mathematical structures (e.g., [49]), and logics (e.g., [17]). We found in [32] that there is no consensus to whether an ontology only includes the concepts and relations of the domain, or if it also includes the instances of concepts. This aspect led researchers to propose several ontology languages based on various formalisms, which only a few of them are widely being used. Two classes of languages are widely used and they are related to the family of logic languages, Common Logic (CL) [3], and Ontology Web Language (OWL) [13,39]. CL is developed from Knowledge Interchange Format (KIF) [21], which is a variant of Lisp, and its dialects [38]. Hence, the main language of CL has a functional programming style. There are indications that CL's main language is becoming a language based on Extensible Markup Language (XML), which is called eXtended Common Logic Markup Language (XCL). On the other hand, OWL has a wider usage and is based on the XML. Several tools, such as Protégé-OWL editor, support the usage OWL.

Starting from a novel data-centered knowledge representation, called Domain Information System (DIS) [31], we propose a high level language for ontology specification. DIS is a formalism that offers a modular structure, which separates the domain knowledge (i.e., the ontology) from the domain data view (i.e., the data or instances of the concepts). It is specific for dealing with Cartesian domains where concepts are formed from atomic concepts through a Cartesian construction. Hence, it is a formalism that takes advantage of a Cartesian perspective on information. The bulk of the data in what is referred to as big data is structured data. It is essentially formed of machine-generated structured data that include databases, sensor data, web log data, security log data, machines logs, point-of-sale data, demographic information, star ratings by customers, location data from smart phones and smart devices, or financial and accounting data. The size of this data is increasing significantly every second. The need for better data analytic techniques that go beyond the capabilities of a Database Management System (DBMS) by connecting the data to concepts that are in the domain but that cannot be defined in a DBMS. In a DIS, the core component of an ontology is a Boolean lattice built from atomic concepts that are imported from the schema of the dataset to be analysed. In [36], we found that DIS enables the integration of several datasets and their respective ontologies for reasoning tasks requiring data-grounded and domain-related answers to user queries. Currently, the language of DIS is a low level language as it is based on a set theory, lattice theory, and graph theory. In this paper, we present a high-level language, called Domain Information System Extended Language (DISEL), that is structured and based on XML. It is built on the top of that of current DIS mathematical language. DISEL uses a mixed structure of directed graphs and trees to precisely capture DIS specifications. It is also based on XML so that it can be easily integrated to many software systems. It aims to make specifying ontologies as easy as possible and without any mathematical complications.

In Sect. 2, we introduce DIS on which DISEL is build. Then, we present the example that is used to illustrate the usage of the constructs of the language. In Sect. 3, we review ontology languages that we found in the literature. In Sect. 4, we introduce the main elements of DISEL and we illustrate their usage using the weather ontology introduced in Sect. 2. In Sect. 5, we give the main design decisions of DISEL. In Sect. 6, we discuss the main features of DISEL, its strengths, and its weaknesses. In Sect. 7, we present our concluding remarks and point to the highlights of our future work.

2 Background

2.1 Domain Information System

DIS is a novel formalism for data-centered knowledge representation that addresses the mapping problem between a set of structured data and its associated ontology [31]. It separates the data structure from the ontology structure and automatically performs the mapping process between the two. A DIS is formed by a domain ontology, a domain data view, and a function that maps the latter to the first. The construction of the domain ontology structure is based on three elements $\mathcal{O} = (\mathcal{C}, \mathcal{L}, \mathcal{G})$, where $\mathcal{C}$ is a concepts structure, $\mathcal{L}$ is a lattice structure, and $\mathcal{G}$ is a set of rooted-graphs. A concept might be atomic/elementary, or composite. A concept that does not divide into sub-concepts is known as atomic concept. In DIS and when for example the data is coming from a database, atomic concepts are the attributes in the database schema. A composite concept is formed by several concepts. The composition of concepts is an operation $\oplus$ defined on C. The set of concepts making the structure $(C, \oplus, c_e)$ is an idempotent communicative monoid, that we denote by $\mathcal{C}$.

The second component of the domain ontology structure is a Boolean lattice $\mathcal{L} = (L, \sqsubseteq_c)$. It is a free lattice generated from the set of atomic concepts. The composition between these concepts make up the remaining concepts of the lattice. The set of concepts of the lattice are either directly obtained from the database schema (i.e., atomic concepts), or obtained through the Cartesian composition of the atomic concepts using the operator $\oplus$. The relationship between a concept and its a composed concept that involves it is the relation `partOf` denoted by $\sqsubseteq_c$.

The last element of the domain ontology is the set of rooted graphs, where an element $G_{t_i} = (C_i, R_i, t_i)$ is a graph on the set of vertices C_i, having the set of edges R_i, and is rooted at the vertex t_i. Rooted graphs represent the concepts that are somehow related to an atomic or composite concept from the Boolean lattice. In other terms, they introduce the concepts that are not directly generated from the database schemes and are related to an atomic or composite concept of the lattice through a relationship from the domain. In this way, the Boolean lattice might be enriched with multiple rooted-graphs having their roots at different lattice concepts.

Using the previous components, we get the full construction of the domain ontology O = ($\mathcal{C}$, $\mathcal{L}$, $\mathcal{G}$). The second component of DIS is the domain data view. Domain data view is abstracted as diagonal-free cylindric algebra $\mathcal{A}$ = (A, +, *, $-$, 0_A, 1_A, $\{c_k\})_{k\in L}$ [25]. It is a Boolean algebra with cylindrification operators $\{c_k\}$ [31]. The cylindrification operators are indexed over the elements of the carrier set L of the Boolean lattice. The last element of the DIS is the mapping operator τ, which maps the data in $\mathcal{A}$ to its related element in the Boolean lattice of O structure. The mapping τ is a function that takes a datum from the data set and returns its corresponding concept in the lattice. Thus, the complete structure of DIS is $\mathcal{D} = (\mathcal{O}, \mathcal{A}, \tau)$. An illustrative representation of the system is shown in the Fig. 1, where the atoms of the lattice are *Attr1*, *Attr2*, and *Attr3*. We notice two rooted graphs that are attached to two elements of the lattice. The reader can find in [36] a comprehensive case study related to a film and TV domain which illustrates the usage of DIS.

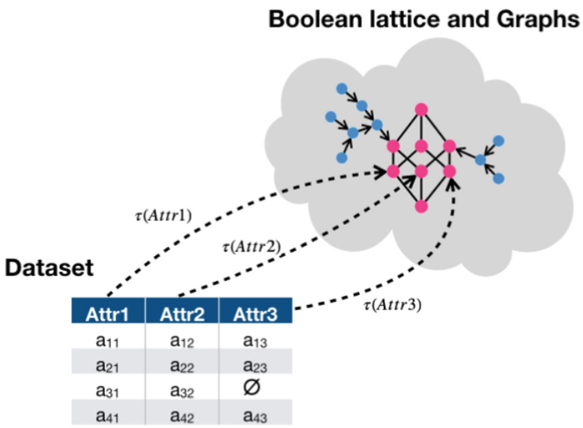


Fig. 1. Abstract view of domain information system.

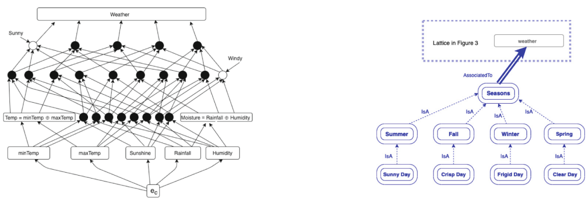
Hence, the formalism of DIS uses a basic mathematical language which is that of lattices, relations, sets, and graphs. The need for a high level language based on the current mathematical language of DIS would extend its use to users who are not well versed in mathematics.

2.2 Illustrative Example

In this paper, we use an example of a weather ontology. Figure 2a gives the dataset that to considered for the example. Figure 2b gives the Boolean lattice associated with the dataset of Fig. 2a (which is borrowed from [7]). We notice that the first level of the Boolean lattice consists of all attributes of the dataset, which are known as atomic concepts. Then, the upper levels are just compositions between these atomic concepts. The top element of the obtained lattice is called "Weather", which is a composition between all atomic concepts. The lattice alone, while it brings all the concepts obtained from the dataset, is not enough to capture all concepts of the weather domain. For that reason, the need to construct rooted-graphs is essential. More concepts are added to the rooted-graphs, then linked to the Boolean lattice specific concepts by the relation of the graph. For example, the rooted-graph *Seasons* is linked to the top of the lattice *Weather* using *IsAssociated* relation as shown in Fig. 2c.

MinTemp	MaxTemp	Rainfall	Sunshine	Humidity
8	24.3	0	6.3	68
14	26.9	3.6	9.7	80
13.7	23.4	3.6	3.3	82
13.3	15.5	39.8	9.1	62
7.6	16.1	2.8	10.6	68
6.2	16.9	0	8.2	70
6.1	18.2	0.2	8.4	63
8.3	17	0	4.6	65
8.8	19.5	0	4.1	70

(a) Weather dataset (Borrowed from [7])



(b) The Boolean lattice (c) One of the Rooted graphs

Fig. 2. A Dataset, its Boolean lattice, and one example of a rooted graph

3 Literature Review on Languages for Ontologies

The literature (e.g., [35]) reveals that there are two trends in developing ontology languages. The first is based on functional languages. The second is XML-based. For more exhaustive surveys on the ontology languages, we refer the reader to [35]. In the following section, we present a brief review to reflect the latest developments in this area.

3.1 Functional Languages

The main functional languages are KIF and CL. KIF [21] was proposed by Stanford AI Lab in 1992. It can be considered as a Domain-specific Language (DSL) based on Lisp. Although v language is capable of articulating specifications of ontologies, it is designed to interchange the knowledge among different programs. Many organizations designed dialects of KIF or proposed extensions to KIF (e.g., IDEF5 [42] or OBKC [12]) to specify ontologies.

CL [3] is a framework that contains a family of languages, which are called dialects. Currently, CL has three dialects: Common Logic Interchange Format (CLIF), Conceptual Graphs Interchange Format (CGIF), and XCL. The dialect CLIF can be considered a simplified form of KIF 3.0. Therefore, its syntax is very similar to KIF; they have Lisp-like syntax [3]. The dialect CGIF is to describe conceptual graphs. It has two versions: Core CGIF and Extended CGIF. The dialect XCL has an XML-based syntax. Actually, XCL is currently the recommended syntax in CL, and the latter can be converted into XCL directly.

3.2 XML-Based Languages

We explore the two main XML-based languages: DARPA Agent Markup Language + Ontology Interchange Language (DAML+OIL) and OWL. DARPA Agent Markup Language (DAML) [15] was merged with Ontology Interchange Language (OIL), which lead to the name DAML+OIL. The language DAML+OIL is a markup language for web resources and it is based on Resource Description Framework (RDF) and Resource Description Framework Schema (RDFS). It gave a strong base for OWL [39], which is designed by W3C.

OWL is a markup language and it is based on DAML+OIL. That means OWL has a relatively fixed format so that it can be easily reasoned on ontologies written with it. Actually, OWL has become the most used language in ontology modeling. The second version of OWL is called OWL2 [13]. In [13], the reader can find a table comparing the different types of OWL grammars.

OWL2 has two semantics for reasoning. One is the Direct Semantics [24], the other is the RDF-Based Semantics [45]. Using the Direct Semantics, ontologies in OWL2 Description Logic (DL) can be translated to $\mathcal{SROIQ}$ which is a particular DL, so that OWL2 could use some DL reasoner. Using the RDF-Based Semantics, ontologies can keep the original meaning. In that way, we can say OWL2 DL is a subset of OWL2 Full. We provide a comparison between all previously mentioned languages in Table 1.

Table 1. Comparison of ontology languages

<table>
<tr><th>Language</th><th>Original grammar</th><th>Based theory</th><th>Extension</th></tr>
<tr><td>KIF</td><td>Lisp</td><td>First-order logic</td><td>IDEF5</td></tr>
<tr><td>CL</td><td>CFIL, CGIL, XCL</td><td>First-order logic</td><td>–</td></tr>
<tr><td>DAML+OIL</td><td>XML</td><td>RDF, RDFS</td><td>–</td></tr>
<tr><td>OWL2 DL</td><td rowspan="2">RDF/XML, OWL/XML, Turtle, Functional Syntax, Manchester Syntax</td><td>Direct Semantics - DL</td><td rowspan="2">MOWL, tOWL</td></tr>
<tr><td>OWL2 Full</td><td>RDF-based semantics</td></tr>
</table>

3.3 Other Ontology Languages

The area of knowledge representation and reasoning involves an indispensable amount of uncertain, and vague information. However, ontology languages are known to be definite (i.e., a concept or relationship exists or not). By default, no support to uncertainty or vagueness is involved in the existing languages [19]. We started observing the rise in exploring languages for probabilistic or fuzzy ontologies, which are ontology languages that are extended with probabilistic or fuzzy formalisms in order to represent uncertainty and vagueness in the domain knowledge. Several efforts to explore such an extension are ongoing. Existing formalisms to handle uncertainty/vagueness in ontology languages are fuzzy logic [51], Bayesian Network (BN) [19], and Dempster Shafer Theory (DST) [37]. Some of the recent approaches that apply probabilistic ontology using BN framework are: BayesOWL [18] and PR-OWL [14]. BN framework is a well-known model to handle uncertainty. However, it was noticed that extending ontology languages with such structure is not an easy task. Moreover, it produces a complex framework [11]. On the other hand, fuzzy ontology approaches include fuzzyOWL [47], and fuzzyOWL2 [10]. Combining fuzzy logic and BN is another approach to get probabilistic ontologies. An example of DST-based ontology is presented in [9]. Several reviews regarding uncertainty management in ontology modelling are available like (e.g., [34,44]).

3.4 Summary

In [35], we find that the most of the ontology languages have the same source and have been sponsored by Defense Advanced Research Projects Agency (DARPA). This limits the diversity among these languages despite their large number. If not disrupted by new ideas, ontology languages tend to have a relatively predictable direction towards a family of similar languages. There is a trend leading to having a family of languages where its members are simple dialects. This has the risk of limiting the extent of usage of these dialects to a class of ontologies. The diversity of languages based on different formalisms is the way to eventually converge towards the objective of designing more simple and expressive ontology languages. Our efforts to design a new language, DISEL, that is based on DIS formalism is towards reaching this objective.

4 DISEL Syntax and Support Tool

In this section, we present the syntax of DISEL language and the features of DISEL Editor (DISEL) tool, which can be downloaded from [2]. The syntax of DISEL is based on XML language. The main component of DISEL is the part for specifying ontologies. Following the DIS formalism for defining ontologies, DISEL's ontology structure is formed of five elements as indicated in Table 2. We allocate a subsection for the first and second elements and then a subsection is dedicated for each of the three remaining elements of this structure. The complete DISEL specification of the *Weather* ontology of the example given in Subsect. 2.2 is available in [5]. An XML schema of DISEL is given in [6]. We also developed a tool named DISEL Editor. It enables ontology specifiers to use a graphical user interface to enter the necessary elements of an ontology, then the system automatically generates the full specification within DISEL language. The current features of DISEL Editor include visual editing ontology, displaying the Boolean lattice of the ontology and visualizing the rooted graph attached to the concepts of the lattice. In the following subsections, we guide the reader through the specification of the weather ontology to illustrate the usage of DISEL Editor tool to specify the *Weather* ontology.

Table 2. Main constructs of an *ontology*

Element name	Description	Type	Cardinality
include	A language construct for including the needed domain information or ontologies from other files	includeType	$[0, +\infty)$
name	The name of ontology	string	only one
atomDomain	The atoms of the domain	atomDomainType	only one
concept	The definition of the concept	conceptType	$[0, +\infty)$
graph	The relation between concepts	graphType	$[0, +\infty)$

4.1 DISEL Editor Interface Overview

In Subsect. 2.1, we indicated that a DIS ontology is formed by a lattice that is obtained from a given set of atomic concepts. The lattice of concepts is then systematically generated as a free Boolean complete lattice from the given set of atoms. The interface of DISEL Editor, which is shown in Fig. 3, enables users to obtain the lattice of concepts from the set of given atomic concepts. The left dock window of DISEL Editor mainly shows the *Atoms of the domain* and *Concept* information. The right upper tab window shows the Boolean lattice generated from the set of atomic concepts that are provided in the *Atoms of the Domain*. If the user clicks a concept in the shown Boolean lattice, DISEL Editor changes the tab to show the rooted graph attached to (rooted at) this concept. The bottom right text edit shows the definition of the selected concept.

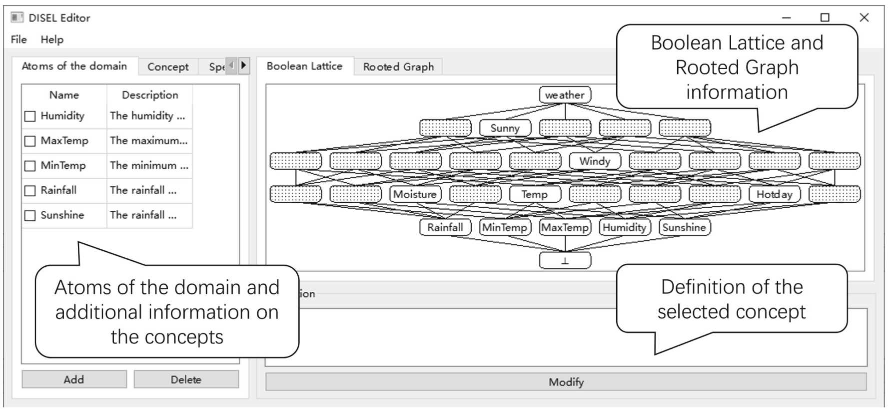


Fig. 3. Interface overview

4.2 Name and Include Constructs

The *name* construct is to be used for giving a unique name to the ontology. As illustrated in [36], DIS allows the integration of several ontologies for a more complex reasoning task than on a single ontology. To allow the usage of concepts and relationships of an external ontology within the specified one, we use the construct *include* to bring the sought external ontology. The *include* construct uses as parameters a *name* attribute and a *filePath* attribute. The *name* element specifies the name of the ontology to be included. The *FilePath* gives the path of included ontology file. Listing 1.1 illustrates the usage of the *include* construct. Sometimes we do not need to include the whole ontology. Sometimes, all what is needed is a vertex (or a set of vertices) that is shared between the constructed ontology and an external ontology. In this case, the usage of that vertex is as illustrated in Listing 1.2, where the vertex *summer* is from an ontology named "OntExample" that has been included using the *include* construct.

Listing 1.1. Code of *include*

```
<include name="OntExample" filePath="OntExample.xml" />
```

Listing 1.2. Code of *include*

```
<edge>
    <from DIS="OntExample">Summer</from>
    <to>Seasons</to>
</edge>
```

4.3 AtomDomain Construct

AtomDomainType can only be included by an ontology with an alias *atomDomain*. It defines all the atoms of the ontology being specified. An atom is the elementary concept in the sense that it cannot be obtained by the composition of other concepts (for further mathematical details on the notion of atoms, we refer the reader to the theoretical foundation of DIS [31,36]). Several types, such as *concept* or *atomSet*, make use of *atom*. The XML schema of *AtomDomainType* and *AtomDomain* is illustrated in [6]. The main constructs of an *atom* are given in Table 3. Using DISEL Editor, the user can click the "Add" Button (see Fig. 3) to create a new *atom* and show the information. While displaying the details of AtomDomain, DISEL Editor shows the Boolean lattice of the current ontology built by AtomDomain as illustrated in Fig. 3. Hence, the user can have a clear understanding of the structure of the ontology.

4.4 Concept

The construct *concept* is the core element in DISEL language. It is used to specify the concepts that are used in the ontology. The main constructs used in *concept* are given in Table 4. Any concept, other than the atomic ones and the ones generated from them to build the Boolean lattice, is introduced using the construct *concept*. In Listing 1.3, Line 2 introduces the concept *Temp*.

Table 3. Main constructs of an *atom*

Element name	Description	Type	Cardinality
name	The name of the atomic concept	string	only one
description	The description of the *atom* concept	string	at most one

Table 4. Main constructs of a *concept*

Element name	Description	Type	Cardinality
name	The name of ontology	string	only one
latticeOfConcepts	The set of atoms from which the lattice is built	latticeOfConceptsType	only one
fefinition	The formulas used by *concept*	string	at most one
fescription	The description of *concept*	string	at most one

Listing 1.3. Example of latticeOfConcepts

```
<latticeOfConcepts>
    <concept>Temp</concept>
    <atom>minTemp</atom>
    <atom>maxTemp</atom>
</latticeOfConcepts>
```

The construct *LatticeOfConcepts* is a set of atoms from which the lattice of concepts is constructed using a Cartesian product construction. Hence, all the concepts of the lattice are tuples of these atoms. If we have n atoms, then the top element of the lattice is an n-tuple of the atomic concepts. It is possible that under *latticeOfConcepts* we find only the atoms and the concepts obtained from them. Also, not all the concepts obtained from the atom have a meaning in the domain of application. The ones that do have a meaning are given the names used for them in the domain of application. In Listing 1.3, we provide an example of the latticeOfConecpts for the concept *Temp*, which is the name in the weather domain for the tuple formed by atoms *minTemp*, and *maxTemp*. We also provide a syntactic sugar called *atoms* and its usage is illustrated in Listing 1.4. It has been shown that the number of attributes in a well normalized dataset schema rarely exceeds ten [48]. Since each attribute would give an atomic concept, then the number of atomic concepts, associated to a normalized database, rarely exceed ten concepts.

Listing 1.4. Syntactic sugar of *atom* in *LatticeOfConcepts*

```
<latticeOfConcepts>
    <atoms>Rainfall Humidity</atoms>
</latticeOfConcepts>
```

It is common that in a domain related to a data set we have more concepts than the ones obtained from the atomic concepts. For instance, we might need new concepts related to concepts in the lattice through a rooted graph. In this case, DISEL Editor enables the ontology specifier to add a new concept as shown in Fig. 4.

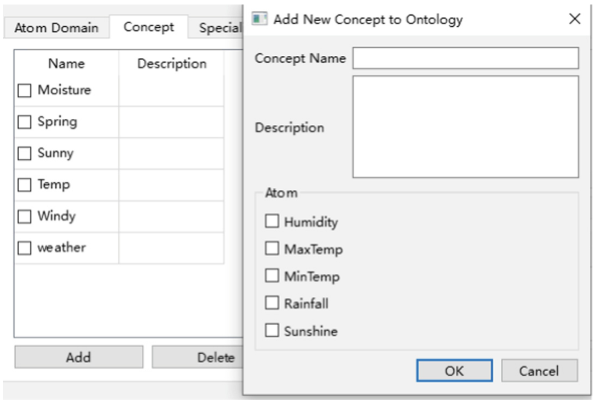


Fig. 4. The concept tab window and dialog

The construct *definition* brings additional information about the concept. For instance, the mathematical definition of the concept can be presented using *definition*. For the convenience of writing, we adopted an Isabelle-like syntax to be used to mathematically define concepts that we omit in this paper for conciseness. At this stage, we can simply say that the type of *definition* is a string. Specific syntactical analytic work is entrusted to a lexer and parser. Operators that contain $<$ and $>$ may cause ambiguity, regarding whether they are part of a markup text or simply part of a textual definition. We recommend to the user to adopt XML CDATA[1] to write the content that goes into *definition*. An example of using CDATA to write the a definition is shown in [6]. The Backus-Naur Form (BNF) of *definition* is shown in [1].

Finally, the construct *description* is to enable the ontology specifier to document the ontology by adding descriptions of the introduced concepts.

4.5 Graph

The part *graph* of the specification of an ontology is to introduce the rooted graphs that bring other concepts than what we have in the lattice of concepts. It also brings relationships among these newly introduced concepts. This part of the specification is for what is called rooted graphs in DIS, which are graphs that must have their roots in the lattice of concepts. The main constructs used to

[1] CDATA stands for Character Data and it indicates that the data in between these strings should not be interpreted as XML markup.

introduce rooted graphs are given in Table 5. The constructs *name* and *rootedAt* are self-explanatory. The construct *edge* is to introduce an edge in the specified graph and it uses two sub-constructs, *from* and *to*, to give the endpoints of an edge. The construct *relation* gives the set of edges of the graph. It also uses sub-construct *properties* to introduce the properties of the relations such as symmetry, transitivity, or being an order. The reader can find in [6] the specification of the rooted graph that is presented in Fig. 2c.

Table 5. The main constructs of a *graph*

Element name	Description	Type	Cardinality
name	The name of rooted graph	string	only one
rootedAt	The root node of rooted graph	string	only one
edge	The structure of rooted graph	edgeType	$[0, +\infty)$
relation	The list of edge	relationType	$[0, +\infty)$

5 Design Decisions

In the following, we present the main design decision for DISEL language. The first design decision is related to basing the language on XML. The latter is currently a widely used markup descriptive language. Many tools provide application programming interfaces for XML to store and exchange information. XML has a stable grammar that can be parsed easily. Basing a high level DIS language on XML would increase the usage of DIS-ontologies by a wider community of users. The exchange between the many ontologies languages that are based on XML will be enhanced. We have considered basing DISEL language on functional languages. Most popular languages are imperative languages. These languages focus on how to process the problem. Functional languages recognize function as first-class citizen. In some cases, functional language can regard its code as information and modify its code. Traditional ontology languages, such as KIF, expand Lisp's grammar in this way. But the grammar of functional language is too free and powerful to control. These are the reasons for taking the decision to not base DISEL on functional languages.

6 Discussion

This paper presents the main syntactical constructs of DISEL language, which was designed as a language for specifying DIS-based ontologies. The adoption of this language would enable the creation of shareable ontologies for the development of ontology-based systems such as the ones characterised in [27]. The language DISEL is based on DIS formalism that is conceived for the specification of information system. Therefore, it inherits the strengths and the shortcomings of this formalist. As it is illustrated in [36], the DIS enables the construction of an information system in a bottom-up approach. To model the domain knowledge, which is captured by the ontology element in DIS, we start from the attributes found in the schema of the dataset to form the set of atomic concepts. The latter is then used to form a free Boolean lattice that its elements are constructed from the atomic concepts using a Cartesian product construction. Then, we bring additional concepts to the ontology by adding the rooted graphs. Each of these graphs brings new concepts that are related to the concepts generated from the datasets. Hence, any concept in the ontology is related to the dataset under consideration. That is why the rooted graphs must have roots in the lattice of concepts. As it is indicated in [36], formalisms such as DL [8], Ontology-Based Data Access (OBDA) [50], and Developing Ontology-Grounded Methods and Applications (DOGMA) [26] offer a clear separation of the domain ontology from the data view. Then they need to match the data schema to the concepts in the domain. This activity is called the mapping activity, which introduces non-trivial challenges [50]. To avoid these challenges, the designers of DIS adopted a data-guided approach for the construction of the ontology. The DIS formalism is limited to a Cartesian world: the concepts of the lattice are tuples of concepts. This means that concepts that are composite are formed by the Cartesian product of atomic concepts. This is the case for the realities/domains where data is structured (e.g., relational data bases or log files). DIS puts the data in the data view and the domain knowledge in the ontology. This separation of concern in the design of DIS enables an ease in the evolution and aging of its specification.

As previously indicated, the current language of DIS is formed by basic mathematical languages of lattices, relations, graphs, and relations. We propose DISEL language as a high-level language that eases the specification by guiding the specifier through the specification process and by hiding the above low level mathematical language as much as possible. The DISEL Editor enables the ontology specifier to interactively specify and change the ontology with ease. It constitutes a visualisation system of the specified ontology. We aim at further enhancing this system as we discuss in Sect. 7. The tools that compare to DISEL Editor are Protégé and Eddy [33]. The first supports standard OWL and enables the translation to a family languages. The second, Eddy, has an important feature, which enables the use of Graphol language to graphically handle ontologies.

An ontology specified using DISEL is intended to be compiled into a formal *Isabelle* [4] specification. Then we use the proof assistant *Isabelle/HOL* to automatically generate the DIS theory, as an extension of the HOL-Algebra theories with explicit carrier sets. We explicitly representing the carrier sets of the several algebraic structure used in their corresponding classes and locales. This approach allows us to use a less complex approach for building the free Boolean lattice component, and for reasoning on a DIS [36]. Then, the obtained DIS theory is used to reason on the dataset and to generate fact-grounded knowledge from the given data.

7 Conclusion and Future Work

In this paper, we introduced DISEL and presented its main syntactical constructs. We used an example of a simplified weather ontology to illustrated its usage. We also presented DISEL Editor, which enables us to write, modify, and visualize a DISEL specification. Our future work aims at further enhancing the capabilities of DISEL Editor. First, we plan to add to DISEL Editor capabilities for compiling DISEL specification into a formal *Isabelle* [4] specification. This work is at the publication stage and its software modules are at the testing stage. Second, we aim at enhancing the visualization system of DISEL Editor by adding capabilities related to ontology module visualization. If an ontology is large and there is a need to automatically visualize or extract a module from it, then DISEL Editor should be able to carry the extraction and the visualization of the obtained module. We will consider the modularization techniques presented in [30,32] and that are appropriate for DIS ontologies. Third, we aim at adding to DISEL Editor an API to ease the link between the considered dataset and its ontology.

In [19], we found that there is no support to uncertainty or vagueness is involved in existing ontology languages [19]. We aim at extending the syntax of DISEL to enable the specification of uncertainty in ontologies. This would give DISEL capabilities for data and statistical analysis. In such a way, we would be able to use statistical reasoning on dataset or define probabilistic or fuzzy concepts. We also aim at using ontologies specified using DISEL to tackle data cleaning in continuation of the work given in [28].

References

1. The BNF of DISEL concept definition. https://github.com/YbJerry/DISEL_file/blob/master/BNF.txt. Accessed 15 Mar 2022
2. The GitHub repository of DISEL editor. https://github.com/YbJerry/DISEL_Editor. Accessed 15 Mar 2022
3. Information technology - Common Logic (CL) - a framework for a family of logic-based languages. https://www.iso.org/standard/66249.html. Accessed 15 Mar 2022
4. Isabelle. http://isabelle.in.tum.de/index.html. Accessed 9 Jan 2021
5. A weather example specified using DISEL. https://github.com/YbJerry/DISEL_file/blob/master/dis.xml. Accessed 15 Mar 2022
6. XML schema of DISEL. https://github.com/YbJerry/DISEL_file/blob/master/dis.xsd. Accessed 15 Mar 2022
7. Avagyan, Z.: Weather.csv. www.kaggle.com/zaraavagyan/weathercsv/metadata
8. Baader, F., Calvanese, D., McGuinness, D., Patel-Schneider, P., Nardi, D.: The Description Logic Handbook: Theory, Implementation and Applications. Cambridge University Press, Cambridge (2003)
9. Bellenger, A., Gatepaille, S.: Uncertainty in ontologies: Dempster-Shafer theory for data fusion applications. In: Workshop on Theory of Belief Functions, pp. 1–6. Brest, France, April 2010

10. Bobillo, F., Straccia, U.: An OWL ontology for fuzzy OWL 2. In: Rauch, J., Raś, Z.W., Berka, P., Elomaa, T. (eds.) ISMIS 2009. LNCS (LNAI), vol. 5722, pp. 151–160. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04125-9_18
11. Carvalho, R.N., dos Santos, L.L., Ladeira, M., da Rocha, H.A., Mendes, G.L.: UMP-ST plug-in: documenting, maintaining and evolving probabilistic ontologies using UnBBayes framework. In: Uncertainty Reasoning for the Semantic Web III. LNCS, vol. 8816. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-13413-0_1
12. Chaudhri, V.K., et al.: OKBC: a programmatic foundation for knowledge base interoperability. In: AAAI/IAAI, pp. 600–607 (1998)
13. Consortium, W.: OWL 2 Web Ontology Language document overview (2012)
14. Costa, P., Laskey, K.: PR-OWL: a framework for probabilistic ontologies. In: Proceedings of the Conference on Formal Ontologies and Information Systems, vol. 150, pp. 237–249, May 2006
15. Dan, C.: DAML+OIL (March 2001) reference description. In: World Wide Web Consortium (2001)
16. De Mauro, A., Greco, M., Grimaldi, M.: A formal definition of Big Data based on its essential features. Libr. Rev. **65**(3), 122–135 (2016). https://doi.org/10.1108/LR-06-2015-0061
17. Del Vescovo, C., Parsia, B., Sattler, U., Schneider, T.: The modular structure of an ontology: atomic decomposition and module count. In: WoMO, pp. 25–39 (2011)
18. Ding, Z., Peng, Y., Pan, R.: A Bayesian approach to uncertainty modelling in OWL ontology. Technical report, Maryland and University Baltimore, Department of Computer Science and Electrical Engineering (2006)
19. Fareh, M.: Modeling incomplete knowledge of semantic web using Bayesian networks. Appl. Artif. Intell. **33**(11), 1022–1034 (2019)
20. Fonseca, F.: The double role of ontologies in information science research. J. Am. Soc. Inf. Sci. Technol. - JASIS **58**, 786–793 (2007)
21. Genesereth, M.R., Fikes, R.E., Brachman, R., Gruber, T., Schubert, L.: Knowledge Interchange Format Version 3.0 Reference Manual. Morgan Kaufmann Publishers, Burlington (1992)
22. Ghafourian, S., Rezaeian, A., Naghibzadeh, M.: Graph-based partitioning of ontology with semantic similarity. In: 2013 3th International eConference on Computer and Knowledge Engineering (ICCKE), pp. 80–85. IEEE (2013)
23. Hoehndorf, R., Schofield, P.N., Gkoutos, G.V.: The role of ontologies in biological and biomedical research: a functional perspective. Brief. Bioinform. **16**(6), 1069–1080 (2015)
24. Horrocks, I., Parsia, B., Sattler, U.: OWL 2 web ontology language direct semantics. In: World Wide Web Consortium, pp. 42–65 (2012)
25. Imielinski, T.W.L., Jr.: The relational model of data and cylindric algebras. J. Comput. Syst. Sci. **28**(1), 80–102 (1984)
26. Jarrar, M., Meersman, R.: Ontology engineering – the DOGMA approach. In: Dillon, T.S., Chang, E., Meersman, R., Sycara, K. (eds.) Advances in Web Semantics I. LNCS, vol. 4891, pp. 7–34. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-89784-2_2
27. Jaskolka, J., Leclair, A., Maccaull, W., Khedri, R.: Architecture for ontology-supported multi-context reasoning systems. Data Knowl. Eng. **140**, 102044:1–14 (2022). ISSN 0169-023X. https://doi.org/10.1016/j.datak.2022.102044, https://www.sciencedirect.com/science/article/pii/S0169023X22000453
28. Khedri, R., Chiang, F., Sabri, K.E.: An algebraic approach for data cleansing. In: the 4th EUSPN, vol. 21, pp. 50–59 (2013). Procedia Computer Science

29. Koopmann, P.: Ontology-based query answering for probabilistic temporal data. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, no. 01, pp. 2903–2910, July 2019
30. LeClair, A.: A formal approach to ontology modularization and to the assessment of its related knowledge transformation. Ph.D. thesis, Faculty of Engineering, McMaster University, Hamilton, Ontario, Canada, November 2021
31. LeClair, A., Khedri, R., Marinache, A.: Formalizing graphical modularization approaches for ontologies and the knowledge loss. In: Fred, A., Salgado, A., Aveiro, D., Dietz, J., Bernardino, J., Filipe, J. (eds.) IC3K 2019. CCIS, vol. 1297, pp. 388–412. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-66196-0_18
32. LeClair, A., Marinache, A., Ghalayini, H.E., MacCaull, W., Khedri, R.: A review on ontology modularization techniques - a multi-dimensional perspective. IEEE Trans. Knowl. Data Eng., 1–28 (2022). https://doi.org/10.1109/TKDE.2022.3152928
33. Lembo, D., Pantaleone, D., Santarelli, V., Savo, D.F.: Eddy: a graphical editor for OWL 2 ontologies. In: IJCAI (2016)
34. Lukasiewicz, T., Straccia, U.: Managing uncertainty and vagueness in description logics for the semantic web. J. Web Semant. **6**(4), 291–308 (2008)
35. Maniraj, V., Sivakumar, R.: Ontology languages - a review. Int. J. Comput. Theory Eng. **2**(6), 887–891 (2010)
36. Marinache, A., Khedri, R., Leclair, A., MacCaull, W.: DIS: a data-centred knowledge representation formalism. In: IEEE RDAAPS21, Hamilton, Ontario, Canada, pp. 1–8. IEEE Computer Society, May 2021
37. Mathon, B.R., Ozbek, M.M., Pinder, G.F.: Dempster-Shafer theory applied to uncertainty surrounding permeability. Math. Geosci. **42**(3), 293–307 (2010)
38. McCarthy, J.: Recursive functions of symbolic expressions and their computation by machine, part I. Commun. ACM **3**(4), 184–195 (1960)
39. Mcguinness, D.L., Ed, F.H.: OWL web ontology language: overview (2004)
40. Nuradiansyah, A.: Reasoning in description logic ontologies for privacy management. KI - Künstliche Intelligenz **34**(3), 411–415 (2020)
41. Ozaki, A.: Learning description logic ontologies: five approaches. Where do they stand? KI - Künstliche Intelligenz (German J. Artif. Intell.) **34**(3), 317–327 (2020)
42. Peraketh, B., Menzel, C.P., Mayer, R.J., Fillion, F., Futrell, M.T.: Ontology capture method (IDEF5). Technical report, Knowledge Based Systems Inc., College Station, TX (1994)
43. Sabarmathi, G., Chinnaiyan, R.: Investigations on big data features research challenges and applications. In: 2017 International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 782–786, June 2017
44. Samani, Z.R., Shamsfard, M.: The state of the art in developing fuzzy ontologies-a survey, pp. 1–46. arXiv preprint arXiv:1805.02290 (2018)
45. Schneider, M., Carroll, J., Herman, I., Patel-Schneider, P.F.: OWL 2 web ontology language RDF-based semantics. In: W3C Recommendation, December 2012
46. Schneider, T., Simkus, M.: Ontologies and data management: a brief survey. KI - Künstliche Intelligenz (German J. Artif. Intell.) **34**(3), 329–353 (2020)
47. Stoilos, G., Stamou, G.B., Tzouvaras, V., Pan, J.Z., Horrocks, I.: Fuzzy OWL: uncertainty and the semantic web. In: OWLED, vol. 188, pp. 1–10 (2005)
48. Vassiliadis, P., Zarras, A.V., Skoulis, I.: How is life for a table in an evolving relational schema? Birth, death and everything in between. In: Johannesson, P., Lee, M.L., Liddle, S.W., Opdahl, A.L., López, Ó.P. (eds.) ER 2015. LNCS, vol. 9381, pp. 453–466. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-25264-3_34
49. Wang, L., Liu, X., Cao, J.: A new algebraic structure for formal concept analysis. Inf. Sci. **180**(24), 4865–4876 (2010)

50. Xiao, G., et al.: Ontology-based data access: a survey. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence, pp. 5511–5519. AAAI Press (2018)
51. Zadeh, L.A.: Fuzzy sets. In: Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi A Zadeh, pp. 394–432. World Scientific (1996)